AN INTRODUCTION TO STATISTICS

WITH

DATA ANALYSIS

by

SHELLEY RASMUSSEN

Department of Mathematical Sciences
Olney 428T
University of Massachusetts/Lowell
Lowell, MA 01854

Originally published by Brooks/Cole Publishing Company,
Division of Wadsworth, Inc.

ISBN 0-534-13578-1

Shelley_Rasmussen@uml.edu

Minitab is a statistical package, a computer program that performs many statistical procedures. The versions of Minitab now available for use on personal computers are menu-driven and much easier to use than the main-frame version originally discussed in this text. Those sections are not included in this online edition of the text. At this time, the most recent version is Minitab 15, available at very reasonable prices for purchase or rental from:

www.e-academy.com/minitab

System Requirements

Processor:	PC with a 1 GHz 32- or 64-bit processor
Memory:	512 MB or more of available RAM
Disk Space:	125 MB free space available
Operating System:	Microsoft Windows 2000, XP, or Vista.
Display:	A display capable of 1024 X 768 or higher resolution
Software:	Adobe Acrobat Reader 5.0 or higher for Meet Minitab

CHAPTER 6

Some Ideas in Probability Needed for Statistical Inference

IN THIS CHAPTER

Experiment, outcome, sample space, event
Probability function
Conditional probability, Bayes' rule
Independence, independent events, dependent events
Random variable, probability distribution of a random variable
Mean or expected value, variance, standard deviation of a random variable

So far we have used tools of data analysis to learn about a collection of information. In particular, we applied these tools in an exploratory study of the World Bank indicators. For such an analysis to make any sense, we have to make some assumptions about the data set. We must assume that the variables (such as the World Bank indicators) are measured and recorded correctly, for instance. When we know or suspect that these assumptions are not met (as in possible deliberate misreporting of World Bank indicators by member countries), we must be careful in our interpretations of results.

In formal statistical inference, we go beyond the goals of data analysis. We want to use samples to learn about larger populations. The tools of data analysis can help us do this. In addition, however, we want to use our observations to make probability statements about the larger populations. To make such probability statements, we must assume more about the observations than is necessary for data analysis.

Suppose, for example, that we are medical researchers. We want to compare a new treatment with a standard treatment in curing patients with a disease. We treat ten patients with the new treatment and ten patients with the standard treatment. Eight patients receiving the new treatment are cured, while three patients receiving the standard treatment are cured. How can we use the results of this experiment to make a statement comparing the cure rates for the two treatments? (By cure rate, we mean the proportion of patients cured.)

Since the numerical results of this experiment are so simple, data analysis amounts to reporting that eight of ten patients were cured under the new treatment and three of ten patients were cured under the standard treatment. (We could present this information in a frequency table.) In formal statistical inference, we might ask this question: If there were really no difference in cure rates between the new and standard treatments, would we be surprised by our observed experimental results? Or: How *likely* is it that we would *by chance* have observed a difference at least as extreme as eight out of ten cured versus three out of ten cured, if the two treatments really had the same cure rate? The answer to this question helps us make a statement comparing the two treatments. (Formal statistical analysis of this type of experiment is the subject of Sections 16-4 and 16-5.)

When we use the words *likely* and *by chance,* as in the question above, we are dealing with probabilities. Even the term *cure rate* refers to the likelihood or probability of cure in a group of patients.

We must make some assumptions about our experiment for any statements involving probabilities to make sense. We must assume, for instance, that the patients receiving the new treatment are similar to the patients receiving the standard treatment. Then, if there were really no difference in cure rates between the two treatments, we would expect a similar number of patients cured in each group. If the experimental results show a striking difference between proportions cured in the two groups, we might doubt that the two treatments are really similar.

In general, statistical inference involves making probability statements about populations based on what we observe in our samples. All formal statis-

tical analysis depends on making assumptions that justify these inferences. When we design an experiment to meet the assumptions needed later for formal analysis of the results, we are in the realm of experimental design, an important part of statistics that we will address in Part III. The ideas in probability that we need for formal statistical inference are the subject of discussion here in Part II.

6-1 Probability as Chances

Probability is a part of everyday life. Buying a state lottery ticket, we like to know our chance of winning. We may guess the odds that the hometown team will win the next game. If a screening test comes up positive for the acquired immune deficiency syndrome (AIDS), we would like to know the probability that we really do have AIDS.

On radio and television weather reports, we hear statements such as "partly cloudy today with a 50% chance of showers" or "the probability of rain today is 90%." What does it mean to say the chance of rain is 90%? We address this question in Example 6-1.

EXAMPLE 6-1

When weather forecasters say there is a 90% probability of rain, they mean that the chances are 9 in 10 that at least .01 inch (or .2 millimeter) of rain will fall during the forecast period (usually 12 hours) at any given point in the forecast area (National Weather Service *Operations Manual,* 1984; kindly provided by Joe Bocchieri and Keith Seitter). If you live in an area receiving a 90% probability of rain forecast, then over the next 12 hours there are 9 chances in 10 that your location will receive at least .01 inch of rainfall.

A 50% chance of snow means that over the forecast period, any given point in the forecast area has 1 chance in 2 of receiving at least .01 inch water-equivalent of snow.

Ranges for probability of precipitation forecasts are from near 100% down to 20%. With probabilities below 20%, precipitation is usually not mentioned in the main part of a weather forecast.

The *probability* of an event is the chance or likelihood of the event occurring. A weather forecaster estimates the chance of precipitation. A lottery player wants to know the likelihood that her selection will win. A cancer researcher may estimate the probability that a male cigarette smoker will develop lung cancer. Figuring the chance of a serious nuclear reactor accident might be the task of a government regulatory commission. An oddsmaker assesses the likelihood that the Lakers will win four of a possible seven games in the National Basketball Association championship finals.

> The **probability** of an event is the chance or likelihood of the event occurring.

Some definitions that help us discuss probability more formally are given in Section 6-2. We will illustrate the terms defined with some simple examples in probability.

6-2 Experiment, Outcome, Sample Space, Events

By an *experiment* we will mean a process leading to a well-defined observation, called the experimental *outcome*. The *sample space* is the set of all possible outcomes of the experiment.

An **experiment** is a process leading to a well-defined observation or outcome.

The **sample space** is the set of all possible outcomes of the experiment.

EXAMPLE 6-2 In one version of a state lottery called Megabucks, a player purchases a ticket after selecting six different integers from 1 to 36. Part of the ticket price goes into the state treasury and the other part into the pot. On drawing day, state officials select six integers from 1 to 36. Players holding tickets with the winning six numbers share the pot. This experiment involves selection of six different integers from 1 to 36. An observation or outcome is a particular combination of six numbers. The sample space is the set of all possible combinations of six integers from 1 to 36.

EXAMPLE 6-3 A researcher carries out this experiment as part of a study of cigarette smoking and lung cancer. She selects a male cigarette smoker at random from among all male cigarette smokers. She then keeps in touch with him until he either develops lung cancer or dies with no evidence of lung cancer. The sample space for this experiment contains two possible outcomes: either the man develops lung cancer or the man dies with no evidence of lung cancer. (The idea of random selection is one we will encounter when we discuss experimental design and statistical inference. *Random selection* of an individual from a population means that each member of the population has an equal chance of being selected.)

EXAMPLE 6-4 The New York Knicks and the Phoenix Suns are meeting in the National Basketball Association championship series. The team that wins four of a possible seven games wins the championship. The experiment involves the two teams playing until one team has won four games. An observation or outcome is a particular sequence of wins by the two teams, with one team winning four games. The sample space consists of all possible outcomes of the series.

We will consider two types of sample spaces: finite sample spaces and continuous sample spaces. A *finite sample space* contains a finite number of

outcomes. The sample spaces in Examples 6-2, 6-3, and 6-4 are all finite sample spaces.

A **finite sample space** is a sample space that contains a finite number of outcomes.

A *continuous sample space* equals an interval of values. The next three examples describe some continuous sample spaces.

A **continuous sample space** is a sample space that equals an interval of values.

EXAMPLE 6-5 For this experiment, we time the length of a Friday morning statistics class. The class is scheduled to meet for 50 minutes on Fridays beginning at 10:30 A.M. The class must be out of the room by 11:25 A.M. at the latest. Therefore, the outcome or length of the class meeting can be from 0 to 55 minutes. The sample space is the interval [0, 55] consisting of all real numbers from 0 to 55, a continuous sample space. (In interval notation, brackets indicate that the endpoints are included in the interval.)

EXAMPLE 6-6 In a life-testing experiment, an engineer tests a computer chip until failure. The outcome is the time in hours from start of testing until failure. In theory, any number greater than or equal to 0 is possible. Therefore, the sample space consists of all real numbers greater than or equal to 0, denoted by the interval $[0, \infty)$. (In interval notation, a parenthesis indicates that the endpoint is not included in the interval.) This is a continuous sample space.

EXAMPLE 6-1 *(continued)* You measure the rainfall at your home over a 12-hour period. The outcome is the amount of rainfall, in inches. The sample space consists of numbers greater than or equal to 0, a continuous sample space.

Some sample spaces are neither finite nor continuous according to our definitions. Suppose, for instance, that our experiment consists of counting the number of defects on a copper sheet for use in personal computer hardware. The sample space contains all possible outcomes of this experiment: 0, 1, 2, 3, and so on. Since this sample space contains all the nonnegative integers, it is a discrete, countably infinite sample space. Although very useful for many applications, we will not consider such sample spaces in this book.

An *event* is a subset of the sample space. One event in the Megabucks lottery, Example 6-2, consists of the outcomes with all six numbers greater than 3. Another event consists of all outcomes that include the number 30. In the basketball championship, Example 6-4, one event consists of outcomes with the series ending in five games. Another event consists of all outcomes in which the Knicks win the series.

An **event** is a subset of the sample space.

When the sample space is continuous, we generally consider events that are continuous, or intervals, as well. For the statistics class in Example 6-5, one event includes all outcomes when the class runs long. This event includes class times longer than 50 minutes, represented by the interval (50, 55]. The interval [49, 51] denotes the event that the class lasts from 49 to 51 minutes. (Recall, in interval notation, a parenthesis indicates that the endpoint is not included in the interval; a bracket, that the endpoint is included in the interval.)

One event in the life-testing experiment in Example 6-6 is that the computer chip lasts more than 2,000 hours, denoted by the interval (2,000, ∞). The event that the computer chip fails within the first 1,000 hours is represented by the interval [0, 1,000].

If you measure less than half an inch of rainfall in Example 6-1, the event is denoted by the interval [0, .5).

In Section 6-3, we define a probability function for a finite sample space. We consider probability functions for continuous sample spaces in Chapter 8 and in Part III. Probability functions formalize our intuitive notions about probability, and help us build the probability models we need for statistical inference.

6 - 3

Probability Functions

A *probability function* is a rule that describes how we assign probabilities or chances to events.

> A **probability function** assigns a unique number or probability to each outcome in a finite sample space S. The probability of each outcome is greater than or equal to 0. The sum of the probabilities of all the outcomes in S equals 1. The probability of an event E, denoted by $P(E)$, is the sum of the probabilities of all the outcomes in E.

From the definition of a probability function, we see that the probability of the entire sample space is 1:

$$\text{If } S \text{ denotes the sample space, then } P(S) = 1.$$

The probability that an event E does not occur equals 1 minus the probability that the event E does occur:

$$\text{For any event } E,\ P(\text{not } E) = 1 - P(E).$$

We sometimes find it useful to consider the probability that at least one of several events occur. The probability that at least one of several events occur equals 1 minus the probability that none of the events occur.

For any collection of events:
P(at least one of the events occur) = 1 − P(none of the events occur)

Also, the probability that at least one of several events occur is less than or equal to the sum of the individual probabilities of the events:

For any collection of events:
P(at least one of the events occur)
$\leq$ sum of the individual probabilities of the events

Suppose that several events have no outcomes in common; the experiment can result in at most one of these events. Then the probability that at least one of the events occur equals the sum of the individual probabilities of the events:

If several events have no outcomes in common, then:
P(at least one of the events occur)
$=$ sum of the individual probabilities of the events

Note that if several events have no outcomes in common, then *at least one occurs* is the same as *exactly one occurs.*

A sample space S and a probability function P for S provide a *probability model* for an experiment. Examples 6-7 and 6-8 illustrate two different probability models for the same experiment.

A sample space S and a probability function P defined for S provide a **probability model** for an experiment.

EXAMPLE 6-7

The experiment consists of tossing a coin three times and noting for each toss whether the coin comes up heads or tails. For a single toss, let H indicate that the coin lands heads up and T that the coin lands tails up. Assume that when the coin is tossed it will come up either heads or tails and not, for example, land on its side and roll away where we cannot see it.

We can denote an outcome of the three coin tosses by an ordered triple. The first, second, and third elements of the ordered triple show the results of the first, second, and third tosses, respectively. For instance, the ordered triple (H, T, H) denotes one possible outcome of the experiment: heads on the first toss, tails on the second toss, heads on the third toss.

The sample space S contains eight possible outcomes, listed in the first column of Table 6-1. Suppose we think of these outcomes as resulting from independent tosses of a fair coin. (By a *fair coin,* we mean heads and tails are equally likely. By *independent tosses,* we mean the outcome of one toss does not in any way affect the outcome of another toss.) Then under the resulting probability model, each outcome in S is equally likely. For this model, the probability function P assigns the number $\frac{1}{8}$ to each of the eight outcomes in S.

Let E denote the event that all three tosses result in heads. Then E contains exactly one outcome and we write $E = \{(\text{H, H, H})\}$, using braces to enclose the outcomes in the event. According to the probability model in Table 6-1, $P(E) = \frac{1}{8}$.

If event E does not occur, then there is at least one tail in the three tosses. From the relation $P(\text{not } E) = 1 - P(E)$, we see that

$$P(\text{at least one tail}) = 1 - P(\text{no tails}) = 1 - P(E) = 1 - \frac{1}{8} = \frac{7}{8}$$

Suppose A_1 denotes the event that heads come up on the first two tosses, A_2 the event that heads come up on the second and third tosses, and A_3 the event that heads come up on the first and third tosses. Table 6-2 shows the outcomes in these events. Under the probability model in Table 6-1, each of events A_1, A_2, and A_3 has probability $\frac{1}{4}$. The probability that at least one of these three events occur is less than or equal to the sum of the individual probabilities of the three events:

$$P(\text{at least one of events } A_1, A_2, \text{ and } A_3 \text{ occur}) \leq P(A_1) + P(A_2) + P(A_3)$$

The event that at least one of A_1, A_2, and A_3 occur is the same as the event that at least two tosses result in heads. Therefore,

$$P(\text{at least two heads}) \leq P(A_1) + P(A_2) + P(A_3) = \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = \frac{3}{4}$$

In this simple example, it is easy to find the exact probability of at least two heads to be $\frac{1}{2}$.

TABLE 6-1 With this probability model, each possible outcome of three coin tosses is equally likely.

Outcome	Probability of the outcome
(H, H, H)	$\frac{1}{8}$
(H, H, T)	$\frac{1}{8}$
(H, T, H)	$\frac{1}{8}$
(T, H, H)	$\frac{1}{8}$
(H, T, T)	$\frac{1}{8}$
(T, H, T)	$\frac{1}{8}$
(T, T, H)	$\frac{1}{8}$
(T, T, T)	$\frac{1}{8}$

TABLE 6-2 Some possible events resulting from three coin tosses and their probabilities based on the probability model in Table 6-1

Description	Event	Probability of the event
Heads on the first two tosses	$A_1 = \{(H, H, H), (H, H, T)\}$	$P(A_1) = \frac{1}{4}$
Heads on the second and third tosses	$A_2 = \{(H, H, H), (T, H, H)\}$	$P(A_2) = \frac{1}{4}$
Heads on the first and third tosses	$A_3 = \{(H, H, H), (H, T, H)\}$	$P(A_3) = \frac{1}{4}$
Exactly one tail	$B_1 = \{(H, H, T), (H, T, H), (T, H, H)\}$	$P(B_1) = \frac{3}{8}$
Exactly two tails	$B_2 = \{(T, T, H), (T, H, T), (H, T, T)\}$	$P(B_2) = \frac{3}{8}$
Three tails	$B_3 = \{(T, T, T)\}$	$P(B_3) = \frac{1}{8}$

Let B_1 denote the event that exactly one tail comes up, B_2 the event that exactly two tails come up, and B_3 the event that exactly three tails come up. The outcomes in B_1, B_2, and B_3 are listed in Table 6-2. Under the probability model in Table 6-1, each of B_1 and B_2 has probability $\frac{3}{8}$ and B_3 has probability $\frac{1}{8}$.

The three events B_1, B_2, and B_3 have no outcomes in common. Therefore, the probability that at least one of these three events occur is equal to the sum of the individual probabilities of the three events:

$$P(\text{at least one of } B_1, B_2, \text{ and } B_3 \text{ occur}) = P(B_1) + P(B_2) + P(B_3)$$

The event that at least one of B_1, B_2, and B_3 occur is the same as the event that there is at least one tail. Therefore,

$$P(\text{at least one tail}) = P(B_1) + P(B_2) + P(B_3) = \frac{3}{8} + \frac{3}{8} + \frac{1}{8} = \frac{7}{8}$$

This is the same answer we obtained for the probability of at least one tail in our earlier calculation.

Let's consider now an alternative probability model for the three-coin-toss experiment.

EXAMPLE 6-8

We toss a coin three times. This time, however, the tosses are not all independent of one another. If the first two tosses result in one head and one tail, then the three-toss outcome has probability $\frac{1}{8}$. However, if heads come up on the first two tosses, then heads are sure to come up on the third toss. Similarly, if tails come up on the first two tosses, then tails will come up on the third. The probability model for this experiment is shown in Table 6-3.

Two of the eight outcomes in Table 6-3 have probability 0. We can define the sample space by including only outcomes with probabilities greater than 0, as in Table 6-4.

TABLE 6-3 The probability model for the three-coin-toss experiment described in Example 6-8

Outcome	Probability of the outcome
(H, H, H)	$\frac{1}{4}$
(H, H, T)	0
(H, T, H)	$\frac{1}{8}$
(H, T, T)	$\frac{1}{8}$
(T, H, H)	$\frac{1}{8}$
(T, H, T)	$\frac{1}{8}$
(T, T, H)	0
(T, T, T)	$\frac{1}{4}$

TABLE 6-4 An alternative description of the probability model for the three-coin-toss experiment in Example 6-8

Outcome	Probability of the outcome
(H, H, H)	$\frac{1}{4}$
(H, T, H)	$\frac{1}{8}$
(H, T, T)	$\frac{1}{8}$
(T, H, H)	$\frac{1}{8}$
(T, H, T)	$\frac{1}{8}$
(T, T, T)	$\frac{1}{4}$

As Example 6-8 illustrates, we can often write the sample space for an experiment in more than one way. Also, the outcomes in a finite sample space need not be equally likely.

Odds or odds ratios are used in a variety of fields, including sports, gambling, and public health. In Section 6-4, we define what we mean by the odds of an event occurring. This section is optional; it is not required to understand any other concepts we will discuss.

6-4

The Odds of an Event (Optional)

Odds ratios depend on the probability model for an experiment. The *odds of an event* is the probability that the event occurs divided by the probability that the event does not occur:

$$\text{For any event } E, \textbf{ odds of event } \boldsymbol{E} = \frac{P(E)}{P(\text{not } E)} = \frac{P(E)}{1 - P(E)}$$

We illustrate the use of the odds ratio with two examples.

EXAMPLE 6-7 *(continued)*

We toss a fair coin three times, with the probability model shown in Table 6-1. Let A denote the event that at least one head comes up in the three tosses. Then A contains seven outcomes and $P(A) = \frac{7}{8}$. The odds of at least one head appearing is the odds of event A:

$$\text{Odds of at least one head} = \frac{P(A)}{1 - P(A)} = \frac{7/8}{1/8} = 7$$

We can denote this odds ratio by 7:1. The interpretation is that over many repetitions of the three-coin-toss experiment, we expect at least one head about seven times for every one time that no heads come up.

If B is the event that a head appears on the first toss, then B contains four

outcomes and $P(B) = \frac{1}{2}$. The odds of a head on the first toss is the odds of event B:

$$\text{Odds of a head on the first toss} = \frac{P(B)}{1 - P(B)} = \frac{1/2}{1/2} = 1$$

This 1:1 ratio indicates that in three independent tosses of a fair coin, a head on the first toss is as likely as a tail on the first toss.

EXAMPLE 6-4 *(continued)*

The Houston Rockets and the Chicago Bulls are meeting in the National Basketball Association championship finals. An oddsmaker lays the odds of the Bulls winning the series as 3:2 or 3 to 2 in favor of the Bulls. According to this oddsmaker, what is the probability that the Bulls will win the championship? We can write the 3:2 odds ratio as

$$\frac{P(\text{Bulls win})}{P(\text{Rockets win})} = \frac{3}{2}$$

The oddsmaker is predicting that the Bulls have 3 chances of winning, compared with 2 chances for the Rockets. Since the Bulls are given 3 chances out of a total of 5 (three chances for the Bulls plus two chances for the Rockets), he believes that

$$P(\text{Bulls win}) = \frac{3}{5} = .6 \quad \text{and} \quad P(\text{Rockets win}) = \frac{2}{5} = .4$$

The oddsmaker believes there is a 60% chance that the Bulls will win the championship.

6-5 Conditional Probability, Independent and Dependent Events

Sometimes we want to revise a probability based on new information. A conditional probability is an updated probability given this new information. Related to conditional probability is the idea of independence (independence of events and independence of observations). We use independence to develop probability models for many experimental situations.

Consider updating probabilities with new information. Suppose you are a project leader at work and your boss has promised to send you a new assistant. In your experience, about 75% of the assistants your boss sends are women. So, if asked, you are likely to say that there is a 75% chance or probability $\frac{3}{4}$ that your new assistant is a woman. At lunch the day before your assistant is to arrive, a friend tells you that your new assistant is bald. With this additional information, you may revise your guess. Since many more bald people are male than female, you might decide the probability of a female assistant is a good deal less than $\frac{3}{4}$. This revised probability is the conditional probability that your assistant is a woman, given that the assistant is bald. We

say the event that the assistant is a woman and the event that the assistant is bald are dependent events, because knowing that one of the events occurred alters the estimated probability of the other event.

The *conditional probability* of an event A given an event B is the revised probability of event A occurring, based on the information that event B has occurred. We define such a conditional or revised probability this way:

> The **conditional probability of event A given event B**, denoted $P(A|B)$, is the probability that events A and B occur together, divided by the probability of event B:
>
> $$P(A|B) = \frac{P(A \text{ and } B)}{P(B)}$$
>
> provided that $P(B)$ is greater than 0.

The conditional probability of event A given event B is defined only when event B is possible, so $P(B)$ must be greater than 0. Sometimes we find it useful to write the definition of conditional probability in the following way:

For any events A and B with $P(B)$ greater than 0,
$$P(A \text{ and } B) = P(B) \times P(A|B).$$

The probability that both A and B occur equals the probability of B times the conditional probability of A given B. This relation gives us a way to find the probability of both A and B occurring when we know the probability of B and the conditional probability of A given B. We will find this useful in some examples that follow.

Two events are *dependent* if knowing that one of the events occurred changes the calculated probability that the other event occurred. Formally, we say:

> Two events C and D are **dependent events** if $P(C|D)$ does not equal $P(C)$. This is the same as saying that $P(D|C)$ does not equal $P(D)$, and the same as saying that $P(C \text{ and } D)$ does not equal $P(C) \times P(D)$.

Let's illustrate the ideas of conditional probability and dependent events with an example.

EXAMPLE 6-9

The following frequency table classifies each of 2,475 people serving time in Georgia prisons for murder, according to race of the convict and race of the convict's victim:

Race of the convict	Race of the victim		
	Black	White	Total
Black	1,438	228	1,666
White	64	745	809
Total	1,502	973	2,475

These data appear, along with additional information, in an article discussing possible racial biases in the Georgia death penalty system ("Supreme Court Ruling on Death Penalty," *Chance,* 1988).

We might use this information to construct a probability model for classifying a convicted murderer in Georgia by race and race of the victim:

Outcome	Probability of the outcome
(prisoner black, victim black)	$\frac{1,438}{2,475} = .58$
(prisoner black, victim white)	$\frac{228}{2,475} = .09$
(prisoner white, victim black)	$\frac{64}{2,475} = .03$
(prisoner white, victim white)	$\frac{745}{2,475} = .30$

Let B denote the event that a convicted murderer is white. Listing the outcomes in B within braces, we can write: B = {(prisoner white, victim black), (prisoner white, victim white)}. According to this probability model, the probability that a convicted murderer is white equals

$$P(B) = \frac{64}{2,475} + \frac{745}{2,475} = \frac{809}{2,475} = .33$$

Let C be the event that the murder victim was white: C = {(prisoner black, victim white), (prisoner white, victim white)}. According to this probability model,

$$P(C) = \frac{228}{2,475} + \frac{745}{2,475} = \frac{973}{2,475} = .39$$

If events B and C both occur, then the convicted murderer is white and the victim was white. Therefore,

$$P(B \text{ and } C) = P(\text{prisoner white, victim white}) = \frac{745}{2,475} = .30$$

The conditional probability that the convicted murderer is white given that the victim was white is

$$P(\text{prisoner is white}|\text{victim was white}) = \frac{P(\text{prisoner white and victim white})}{P(\text{victim white})}$$

and we can write this as

$$P(B|C) = \frac{P(B \text{ and } C)}{P(C)} = \frac{745/2,475}{973/2,475} = \frac{745}{973} = .77$$

Without additional information, we would say there is about a 33% chance that a convicted murderer in Georgia is white. Knowing that the victim was white, we revise this estimate and say there is about a 77% chance that the convicted murderer is white.

The event B that the convicted murderer is white and the event C that the victim was white are dependent events, since $P(B|C)$ does not equal $P(B)$. The race of the convicted murderer and the race of the victim are *dependent* or *associated characteristics*. (We will see this idea again in Sections 16-3 and 16-5 when we discuss formal statistical procedures to test for independence of two qualitative variables.)

Two events are *independent* if knowing that one of the events occurred does not change the calculated probability that the other event occurred. The probability that the two independent events both occur is the product of the individual probabilities of the two events.

> Two events A and B are **independent events** if $P(A|B) = P(A)$. This is the same as saying that $P(B|A) = P(B)$, and the same as saying that $P(A \text{ and } B)$ equals $P(A) \times P(B)$.

We can use the idea of independence to construct probability models for experiments, as illustrated in Example 6-10.

EXAMPLE 6-10

A person is Rh positive (Rh+) if a substance called the Rh+ antigen is on the surface of the red blood cells. If this Rh+ antigen is not there, a person is Rh negative (Rh−).

Each year at City Hospital, about 51% of the babies born are boys. About 85% of the babies are Rh+, regardless of sex, the rest being Rh−. Sex and presence of the Rh+ antigen are *independent* or *unassociated characteristics* since males are as likely as females to be Rh positive. (In Sections 16-3 and 16-5 we will address the problem of deciding whether two qualitative variables are independent or unassociated.)

Suppose we classify a baby born at City Hospital by sex and by presence of the Rh+ antigen. Then there are four possible outcomes in the sample space. We can use the given information to construct a probability function for this sample space.

The outcome (girl, Rh+) represents the result that the baby is a girl and is Rh+. Since sex and presence of the Rh+ antigen are independent characteristics, we calculate the probability of this outcome as

$$P(\text{girl, Rh+}) = P(\text{girl}) \times P(\text{Rh+}) = .49 \times .85 = .4165$$

We calculate the probabilities of the other outcomes in a similar way (Exercise 6-5).

We can extend the definition of independence to any finite number of events. Consider events A_1 through A_k. It is not enough to say that any pair

of these events are independent (see Exercise 6-28). Instead, we say the following:

> Events A_1 through A_k are **independent events** if for any subcollection of two or more of these events, the probability that each of the events in the subcollection occurs is equal to the product of the probabilities of these events taken separately.

This definition of independence tells us that if A_1 through A_k are independent events, then the probability that A_1 through A_k all occur is the product of the probabilities $P(A_1)$ through $P(A_k)$.

EXAMPLE 6-7 *(continued)*

We toss a fair coin three times and the tosses are independent of one another. That is, the result of one toss does not affect the result of another toss. What is the probability of three heads in a row?

Let E_1 be the event that a head appears on the first toss, E_2 the event that a head appears on the second toss, and E_3 the event that a head appears on the third toss. Because the coin is fair, the probability of each of these events is $\frac{1}{2}$. From the description of the experiment, we know that E_1, E_2, and E_3 are independent events.

The probability of three heads in a row is the probability that each of the events E_1, E_2, and E_3 occurs. Therefore, the probability of three heads in three tosses equals

$$P(\text{three heads in a row}) = P(E_1, E_2, \text{and } E_3 \text{ each occur})$$
$$= P(E_1) \times P(E_2) \times P(E_3) = \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{8}$$

Similar calculations give the entire probability model for this experiment, shown in Table 6-1.

In Example 6-7, we use the independence of the coin tosses to build a probability model for the experiment. In general, we use the concept of independence to develop probability models for experiments in formal statistical inference. An important assumption for building these probability models is that the observations are independent, as we will see in Part III.

We can use a formula called Bayes' rule for calculating a conditional probability when we know certain other probabilities and conditional probabilities. We discuss Bayes' rule in Section 6-6. That material is optional; the information on Bayes' rule is not necessary for understanding any topics that follow.

6-6 Bayes' Rule (Optional)

Bayes' rule is a formula for calculating a conditional probability when certain other probabilities and conditional probabilities are known.

For any events A and E with positive probability, the following relationship (called **Bayes' rule**) is true:

$$P(E|A) = \frac{P(E) \times P(A|E)}{P(E) \times P(A|E) + P(\text{not } E) \times P(A|\text{not } E)}$$

Among other applications, we can use Bayes' rule to study the effectiveness of medical screening tests, as illustrated in the next example.

EXAMPLE 6-11

Some people have proposed large-scale screening programs for the acquired immune deficiency syndrome (AIDS). Is such a program a good idea? In particular, what is the likelihood that a person testing positive for AIDS really has AIDS? If a person tests negative for AIDS, what is the chance he or she really does not have the disease?

To answer these questions, we need an estimate of the proportion of people affected with the AIDS virus in the population to be tested. We also need some information on how good the screening test is at identifying presence and absence of the virus.

Suppose that 1.5 million of roughly 250 million Americans are infected with the AIDS virus. Then our estimate of the proportion infected is 1.5 million divided by 250 million, or .006 ("Random Testing for AIDS?" *Chance,* 1988).

Assume the probability that a person with AIDS is correctly diagnosed (called the *sensitivity* of the screening test) is .98. Assume also the probability that a person without AIDS is correctly diagnosed (called the *specificity* of the screening test) is .93. (These values are based on results reported in Weiss et al., 1985.)

The **sensitivity** of a diagnostic test is the probability that a person with the condition under study will test positive.

The **specificity** of a diagnostic test is the probability that a person without the condition under study will test negative.

Suppose that a single person is randomly selected from the United States population and tested for AIDS. The sample space for this experiment has four outcomes, corresponding to the four possible combinations of disease status (infected with the AIDS virus or not infected) and test result (positive or negative).

Let A denote the event that the person has the AIDS virus. Let T denote the event that the person tests positive for AIDS. Then the probabilities we have estimated above can be written as

$$\begin{aligned} P(A) &= .006 \\ P(T|A) &= .98 \\ P(\text{not } T|\text{not } A) &= .93 \end{aligned}$$

We would like to find $P(A|T)$, the probability that a person testing positive really has AIDS. We can find this conditional probability using Bayes' rule:

$$P(A|T) = \frac{P(A) \times P(T|A)}{P(A) \times P(T|A) + P(\text{not } A) \times P(T|\text{not } A)}$$
$$= \frac{.006 \times .98}{(.006 \times .98) + (.994 \times .07)} = .0779$$

Here, $P(\text{not } A) = .994 = 1 - .006$ is the probability the person does not have AIDS and $P(T|\text{not } A) = .07 = 1 - .93 = 1 - P(\text{not } T|\text{not } A)$ is the probability that a person without AIDS tests positive.

What have we shown here? $P(A|T)$ is the probability that a person testing positive for AIDS is really infected with the virus. This conditional probability is .0779, meaning that of every 10,000 people testing positive for AIDS, we would expect only 779 to actually have the disease! The other 9,221 people are *false positives:* they test positive for AIDS but really do not have the disease.

We can use Bayes' rule again to find the probability that a person testing negative really does not have AIDS, $P(\text{not } A|\text{not } T)$:

$$P(\text{not } A|\text{not } T) = \frac{P(\text{not } A) \times P(\text{not } T|\text{not } A)}{P(\text{not } A) \times P(\text{not } T|\text{not } A) + P(A) \times P(\text{not } T|A)}$$
$$= \frac{.994 \times .93}{(.994 \times .93) + (.006 \times .02)} = .99987$$

This conditional probability tells us that of every 100,000 people testing negative for AIDS, we would expect 99,987 to really be free of the virus. The other 13 people are *false negatives:* they test negative for AIDS but really have the disease. They are infected people missed by the screening procedure.

We can use a tree diagram such as the one in Figure 6-1 to illustrate how Bayes' rule works. Imagine that 100,000 people selected at random from the general United States population are tested for AIDS. We assume the proportion of people in the general United States population with AIDS is .006. Therefore, as the tree diagram in Figure 6-1 illustrates, we expect about 600 of the 100,000 people tested to have AIDS and about 99,400 to be free of the disease.

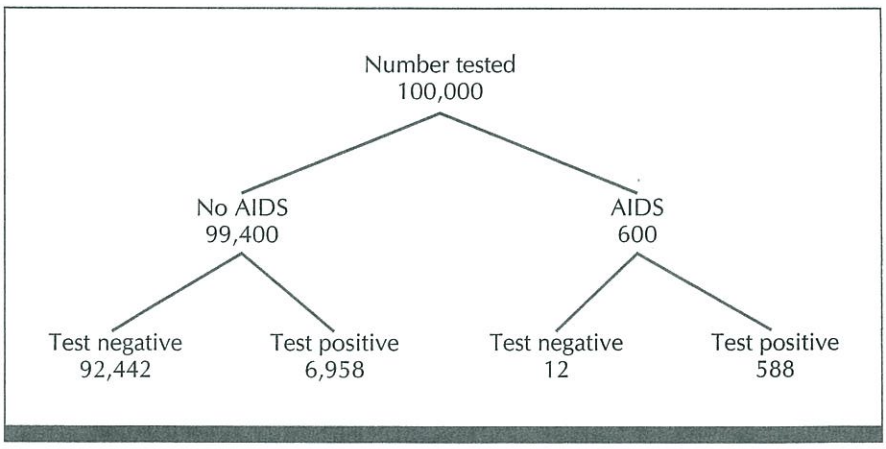


FIGURE 6-1 Tree diagram illustrating the use of Bayes' rule in Example 6-11

We assume a 98% chance that a person with AIDS tests positive for the disease. Therefore, of the 600 people we expect with AIDS, about $.98 \times 600 = 588$ will test positive for AIDS; about $.02 \times 600 = 12$ will test negative.

Similarly, we assume a 93% chance that a person without AIDS will test negative for the disease. Therefore, of the 99,400 people we expect to be free of the AIDS virus, about $.93 \times 99{,}400 = 92{,}442$ will test negative for AIDS; about $.07 \times 99{,}400 = 6{,}958$ will test positive. These expected numbers are shown along the bottom of the diagram in Figure 6-1.

To find the probability that a person testing positive for AIDS really has the disease, we look at the expected numbers in our diagram. We expect 6,958 people without AIDS and 588 people with AIDS to test positive for the disease. Therefore, we calculate

$$P(A|T) = \frac{588}{6{,}958 + 588} = .0779$$

which is the same answer we found using Bayes' rule directly.

In a similar fashion we can find the probability that a person testing negative really is free of the AIDS virus. We expect 92,442 people without AIDS and 12 people with AIDS to test negative for the disease. Therefore,

$$P(\text{not } A|\text{not } T) = \frac{92{,}442}{92{,}442 + 12} = .99987$$

which is again the same answer we found using Bayes' rule directly.

What do these calculations tell us? If the screening procedure comes up negative, there is a very large likelihood the tested individual is in fact free of AIDS. However, if the test results are positive, there is less than an 8% chance that the person really has AIDS! Used as a general screening device, this test would make a lot of people without AIDS very nervous, at least until they could be rechecked. In practice, a screening procedure might not do as well as we have described (Barnes, 1987). What do these calculations suggest to you about the practicality of a large-scale screening procedure such as this?

The probability of a false positive and the probability of a false negative depend on how well the test identifies presence and absence of the condition under study. These probabilities also depend on the proportion of people with the condition in the population. (See Exercises 6-9 through 6-12.) Note that for this type of application, it is *conditional probabilities* (of disease given a positive test result, for instance) that are of main public interest.

6-7 Random Variables

A random variable represents a numerical observation resulting from an experiment. We will be using random variables in our discussion of statistical inference.

A *random variable* is a rule that assigns a number to each outcome in the sample space. A *finite random variable* is a random variable that takes on

a finite number of values. A *continuous random variable* is a random variable that takes values in an interval of numbers.

> A **random variable** is a rule that assigns a number to each outcome in the sample space.
>
> A **finite random variable** is a random variable that takes on a finite number of values.
>
> A **continuous random variable** is a random variable that takes on values in an interval of numbers.

It is possible to have a random variable that is neither finite nor continuous. For instance, if a random variable X denotes the number of defects on a sheet of copper used in personal computer hardware, then X has a countably infinite set of possible values: 0, 1, 2, and so on. Although very useful, we will not consider such random variables in this book.

Examples 6-12, 6-13, and 6-14 illustrate three finite random variables.

EXAMPLE 6-12 You apply for a job and you are either offered the job or not. We can write a sample space as S = {success, failure}, where success indicates you are offered the job and failure indicates that you are not. Let the random variable X represent the rule that assigns the number 1 to success and the number 0 to failure. We can write this rule as

$$X(\text{success}) = 1$$
$$X(\text{failure}) = 0$$

X is a finite random variable that takes on two values, 0 and 1. The random variable X codes the outcomes success and failure as numbers, 1 and 0, respectively. This kind of coding is common. For instance, a worker may code the answer to a survey question as a number before entering it on a computer file.

EXAMPLE 6-13 A pollster interviews a couple with two children. He asks the couple the gender of each child. A possible sample space is S = {(G, G), (G, B), (B, G), (B, B)}. The outcome (G, B) indicates, for example, that the firstborn child is a girl and the secondborn child is a boy.

Perhaps the pollster is really interested only in the number of boys, not birth order. Then he might record their answer using the following rule, denoted by W:

$$W(\text{G, G}) = 0$$
$$W(\text{G, B}) = 1$$
$$W(\text{B, G}) = 1$$
$$W(\text{B, B}) = 2$$

W is a finite random variable that takes on three values: 0, 1, and 2. The pollster uses the rule W to count the number of male children the couple has.

EXAMPLE 6-14

You are interested in your letter grade for an introductory statistics course. We might write a sample space as $S = \{A, B, C, D, F\}$. To see how your grade contributes to your grade point average, you must convert your grade to a number. The random variable Z is a rule assigning a number to each letter grade:

$$Z(A) = 4$$
$$Z(B) = 3$$
$$Z(C) = 2$$
$$Z(D) = 1$$
$$Z(F) = 0$$

Z is a finite random variable with five possible values: 0, 1, 2, 3, and 4.

Example 6-15 illustrates two continuous random variables and one finite random variable.

EXAMPLE 6-15

We select one freshman at random from among all freshmen at State University. Each freshman represents a possible outcome of the experiment; the sample space consists of all freshmen at State University.

We record height, weight, and sex for the freshman selected. Let the random variable X be the height in inches of the student. We can think of X as a rule that assigns to each student in the sample space the number corresponding to the student's height in inches. Let the random variable Y be the weight in kilograms of the student selected. Then Y is a rule that assigns to each student in the sample space the number that is the student's weight in kilograms. X and Y are continuous random variables, since height and weight can take values over intervals of numbers.

Let the random variable Z code the sex of the student selected. Z equals 1 if the student is a woman and 2 if the student is a man. Since it has two possible values, Z is a finite random variable.

As this example illustrates, we can define more than one random variable on the same sample space. This makes sense because in real data collection situations, many numerical variables may be of interest for each individual or outcome observed.

We are often more interested in the numbers assigned by a random variable than in the actual outcomes in the sample space. Then we want to assign probabilities to events defined by the random variable.

For instance, the pollster in Example 6-13 may be interested in the probability that a couple with two children has one boy and one girl. We write this probability as $P(W = 1)$, where W is the random variable that counts the number of male children.

In Example 6-15 we may be interested in the probability that the height of the freshman selected is from 66 to 72 inches. We write this probability as $P(66 \leq X \leq 72)$, where X is the random variable corresponding to the height

in inches of the student selected. Because a student is selected at random from among the freshmen at State University, the probability that X takes a value from 66 to 72 is the same as the proportion of freshmen at State University who are from 66 to 72 inches tall.

We denote the probability that the student selected is female by $P(Z = 1)$, where Z is 1 for a female and 2 for a male. Because the freshman was selected at random, $P(Z = 1)$ is the same as the proportion of women in the freshman class at State University.

We can write the probability that the student selected is female and from 66 to 72 inches in height as $P(Z = 1, 66 \leq X \leq 72)$. Again, because of the random selection, this probability corresponds to the proportion of freshmen at State University who are females from 66 to 72 inches in height.

The probability that the student selected weighs more than 100 kilograms is $P(Y > 100)$, where Y represents the student's weight in kilograms. We denote the probability that the freshman selected is male and weighs more than 100 kilograms by $P(Z = 2, Y > 100)$. We can write $P(Z = 2, 66 \leq X \leq 72, Y > 100)$ to denote the probability that the freshman selected is male, from 66 to 72 inches tall, and weighs more than 100 kilograms.

The Probability Distribution of a Random Variable

A probability function for the sample space determines the probability of an event defined by a random variable. The *probability distribution* of a random variable X refers to the collection of probabilities assigned to events defined by X.

> The **probability distribution** of a random variable is the collection of probabilities assigned to events defined by the random variable.

We can use the probability distribution of a random variable to find the mean or expected value of the random variable, as well as its variance and standard deviation. Many problems in statistical inference involve asking questions about the mean and standard deviation of a random variable.

We define expected value, variance, and standard deviation for finite random variables in Section 6-8. We can give similar definitions for continuous random variables. In Chapter 8 and Part III, we discuss some continuous random variables.

6-8 Mean, Variance, and Standard Deviation of a Finite Random Variable

The mean of a random variable is a measure of location, a measure of the center of the probability distribution of the random variable. The *mean* of a finite random variable X is a weighted average of the values the random variable takes on, each value weighted by the probability that X equals that value.

> The **mean** of a finite random variable X equals $\Sigma\, xP(X = x)$, where the sum is over all numbers x with $P(X = x)$ greater than 0.

Some symbols for the mean of a random variable X are μ and μ_X.

The mean of a random variable X is also called the *expected value of X*, sometimes denoted by EX or $E(X)$. The term *expected value* refers to the average value of X we *expect* to see over many observations. The notation EX or $E(X)$ reminds us that the mean is an expected average in this sense.

The variance of a random variable is a measure of the variation of the random variable about its mean. The *variance* of a random variable X equals $E(X - \mu)^2$, the expected squared deviation of X about its mean.

> The **variance** of a finite random variable X equals $\Sigma\, (x - \mu)^2 P(X = x)$, where μ denotes the mean of X and the sum is over all numbers x with $P(X = x)$ greater than 0.

If a random variable X takes on only one value, then the variance of X equals 0. Otherwise, the variance of X will always be greater than 0 (see Exercise 6-25). Some symbols for the variance of X are σ^2, σ_X^2, and $\text{Var}(X)$.

The standard deviation is also a measure of variation in a random variable. The *standard deviation* of a random variable X is the positive square root of the variance of X. We denote the standard deviation of X by σ, σ_X, or $\sqrt{\text{Var}(X)}$.

> The **standard deviation** of a random variable is the positive square root of the variance of the random variable.

The units of the mean and standard deviation of a random variable X are the same as the units of X. The variance of X is in squared units.

In a sampling situation, we have a sample of observed values of a random variable. We use the sample mean and sample standard deviation to estimate the mean and standard deviation, respectively, of the random variable. When we discuss statistical inference, we see how to use the (observed) sample mean and sample standard deviation to make inferences about the (unknown) mean and standard deviation of a random variable.

Example 6-16 illustrates how to find the mean, variance, and standard deviation of a finite random variable. We will consider a real taste-test experiment similar to the one described in Example 6-16 when we introduce the main ideas of statistical inference in Chapter 9.

EXAMPLE 6-16 Three statistics students volunteer for a taste test comparing Coke and Pepsi. Each student tastes samples in two identical-looking cups and decides which beverage he or she prefers. How many students do we expect to pick Pepsi? We can answer this question only after we specify a probability model for the problem. We will describe one possible probability model.

Suppose the students make selections independently of one another. Suppose also that the probability of picking Pepsi is $\frac{3}{5}$ and the probability of picking Coke is $\frac{2}{5}$ for all three students. We will refer to the tasters as the first,

second, and third students. One possible outcome of the experiment is that the first and second students pick Pepsi and the third student picks Coke. We denote this outcome by (Pepsi, Pepsi, Coke). There are eight possible outcomes in the sample space, listed in Table 6-5.

Let the random variable Y equal the number of Pepsi selections for any given outcome of the experiment. The values of Y for the eight outcomes in the sample space are shown in column 2 of Table 6-5.

We assume the students make independent selections. Thus, the probability for each outcome is calculated as shown in column 3 of the table. For example,

$$P(\text{Pepsi, Pepsi, Coke}) = \frac{3}{5} \times \frac{3}{5} \times \frac{2}{5} = \frac{18}{125} = .144$$

Under our probability model, the probability that the first two students pick Pepsi and the third student picks Coke is $\frac{18}{125}$ or .144.

We see that the random variable Y takes on four values: 0, 1, 2, and 3. The probability distribution for Y is given by:

$$P(Y = 3) = \frac{27}{125}$$

$$P(Y = 2) = \frac{18}{125} + \frac{18}{125} + \frac{18}{125} = \frac{54}{125}$$

$$P(Y = 1) = \frac{12}{125} + \frac{12}{125} + \frac{12}{125} = \frac{36}{125}$$

$$P(Y = 0) = \frac{8}{125}$$

The mean or expected value of Y is

$$\begin{aligned} E(Y) &= 3 \times P(Y = 3) + 2 \times P(Y = 2) \\ &\quad + 1 \times P(Y = 1) + 0 \times P(Y = 0) \\ &= 3 \times \frac{27}{125} + 2 \times \frac{54}{125} + 1 \times \frac{36}{125} + 0 \times \frac{8}{125} \\ &= \frac{225}{125} = 1.8 \text{ Pepsi selections} \end{aligned}$$

The expected value of Y is 1.8, the expected number of Pepsi selections among the three students. This expected value does not equal any of the values Y can take on. How should we interpret this number? We can think of the expected value as the average number of Pepsi selections among three students, if we were able to do the experiment many times under the same conditions. Suppose many groups of three students participate and the probability model above applies to each group. If we divide the total number of Pepsi selections by the number of student groups, the result will be close to 1.8. In

TABLE 6-5 The value of the random variable Y and the probability of each outcome in the sample space for Example 6-16

Outcome	Value of Y (number of Pepsi selections)	Probability of the outcome
(Pepsi, Pepsi, Pepsi)	3	$\frac{3}{5} \times \frac{3}{5} \times \frac{3}{5} = \frac{27}{125} = .216$
(Pepsi, Pepsi, Coke)	2	$\frac{3}{5} \times \frac{3}{5} \times \frac{2}{5} = \frac{18}{125} = .144$
(Pepsi, Coke, Pepsi)	2	$\frac{3}{5} \times \frac{2}{5} \times \frac{3}{5} = \frac{18}{125} = .144$
(Coke, Pepsi, Pepsi)	2	$\frac{2}{5} \times \frac{3}{5} \times \frac{3}{5} = \frac{18}{125} = .144$
(Coke, Coke, Pepsi)	1	$\frac{2}{5} \times \frac{2}{5} \times \frac{3}{5} = \frac{12}{125} = .096$
(Coke, Pepsi, Coke)	1	$\frac{2}{5} \times \frac{3}{5} \times \frac{2}{5} = \frac{12}{125} = .096$
(Pepsi, Coke, Coke)	1	$\frac{3}{5} \times \frac{2}{5} \times \frac{2}{5} = \frac{12}{125} = .096$
(Coke, Coke, Coke)	0	$\frac{2}{5} \times \frac{2}{5} \times \frac{2}{5} = \frac{8}{125} = .064$

some groups, all three students may pick Pepsi, in other groups one or two students may pick Pepsi, and in other groups no students may pick Pepsi. But the average number of Pepsi selections per group will be about 1.8.

The variance of the random variable Y in this example is

$$\begin{aligned}\text{Var}(Y) &= (3 - 1.8)^2P(Y = 3) + (2 - 1.8)^2P(Y = 2) \\ &\quad + (1 - 1.8)^2P(Y = 1) + (0 - 1.8)^2P(Y = 0) \\ &= 1.44 \times \frac{27}{125} + .04 \times \frac{54}{125} + .64 \times \frac{36}{125} + 3.24 \times \frac{8}{125} \\ &= .72 \text{ (Pepsi selections)}^2\end{aligned}$$

The standard deviation of Y is $\sqrt{\text{Var}(Y)} = \sqrt{.72} = .85$ Pepsi selection.

Note that the sample space for the taste test in Example 6-16 (the eight outcomes are listed in Table 6-5) looks like the sample space for a three-coin-toss experiment such as we considered in Example 6-7. Each student in the taste test has two possible choices (Pepsi or Coke), while each coin toss has two possibilities (head or tail). When we have independent repetitions of a two-outcome experiment and the probabilities of the two outcomes are the same for each repetition, we say we have a *binomial experiment.* We will discuss probability models for binomial experiments in Chapter 7. Many real experiments involve repetitions having two possible outcomes (for example, cured versus not cured in a medical experiment, alive versus dead in an animal toxicity study, product A versus product B in a market research survey, acceptable versus defective in an industrial setting). We will discuss formal statistical analysis of such binomial experiments in Sections 10-2, 10-5, 16-1, and 16-2.

Example 6-17 is a simple application of probabilities and expected values.

EXAMPLE 6-17 How do we make decisions in situations involving risk, when we are not certain of the consequences of possible choices or courses of action? Knowledge,

prejudice, and assessment of costs and benefits may come into play. But our decisions in risky situations may also be influenced by the way we perceive these situations (Allman, 1985), as the following experiment illustrates.

Experimenters presented two groups of physicians with a hypothetical problem regarding an impending outbreak of a rare disease (Kahneman and Tversky, 1982). If untreated, the disease will kill 600 people.

The experimenters offered physicians in the first group two possible programs for fighting the disease. With program A, 200 of the disease victims will be saved. With program B, there is a $\frac{1}{3}$ chance that all 600 disease victims will be saved and a $\frac{2}{3}$ chance that none of them will be saved. The expected number of disease victims saved under program B is 200 (Exercise 6-26). Therefore, as far as *expected* number of victims saved is concerned, programs A and B are equivalent. A majority of the physicians in this group chose program A; they preferred to save 200 disease victims for sure rather than risk saving none of the victims.

The experimenters also presented physicians in the second group with two programs for dealing with the disease. With program C, 400 disease victims will die. With program D, there is a $\frac{2}{3}$ chance that all 600 disease victims will die and a $\frac{1}{3}$ chance that none of them will die. The expected number of deaths among disease victims is 400 for program D (Exercise 6-27), the same as for program C. However, a majority of physicians in this group chose program D. These physicians preferred to take a chance and try to prevent all the disease victims from dying rather than accept the certainty of 400 dying.

The results of the programs offered the two groups of physicians are the same. Programs A and C both guarantee 200 victims saved, with 400 deaths. Programs B and D offer a $\frac{1}{3}$ chance that all disease victims are saved (none die) and a $\frac{2}{3}$ chance that none are saved (all the disease victims die). The choices offered the two groups were similar in content but not in presentation. The physicians in the first group chose between two gains. These physicians tended to prefer a certain gain (in saved lives) to an uncertain situation with possible large gain and possible no gain. The physicians in the second group chose between two losses. These physicians tended to reject the certain loss (deaths) and choose the uncertain course with possible no loss as well as possible large loss.

How we evaluate and react to a situation involving uncertainty and risk seems to depend not only on our background, experience, knowledge, and so on, but also on whether we consider the problem in terms of gains or losses, benefits or costs. (When making decisions in the face of uncertainty, it matters whether we think of the proverbial glass as half full or half empty!)

Summary of Chapter 6

A probability function defined on a sample space formalizes the way we assign probabilities or chances to events.

The conditional probability of event A given event B is the probability

that events A and B both occur divided by the probability of event B. Bayes' rule provides a way of calculating a conditional probability when certain other probabilities and conditional probabilities are known. A useful application of Bayes' rule is in evaluation of medical screening procedures.

Two events are independent if knowing that one of the events occurred does not change the calculated probability that the other event occurred. The concept of independence is useful in building many probability models. Two events are dependent if knowing that one of the events occurred does change the calculated probability that the other event occurred.

A random variable is a rule that assigns a number to each outcome in a sample space, representing numerical observations made during an experiment. The probability distribution of a random variable specifies the probability of each event defined by the random variable. The method of analysis we select in statistical inference depends on the probability distribution of the variable observed.

The mean or expected value of a random variable is a measure of the center of the values taken on by the random variable. The variance and standard deviation measure the spread or variation in the values taken on by the random variable. Much of statistical inference involves asking questions about the mean and variance of a random variable. In a classical analysis, we base our inferences about the unknown mean and variance of a random variable on the sample mean and sample variance calculated from the observations.

Exercises for Chapter 6

EXERCISE 6-1

In 1984, researchers asked a group of 1,060 Vermont 4th and 5th graders about their experiences with cigarette smoking (Haugh et al., 1986). Of the 515 girls, 379 said they had never smoked, 114 said they had tried a few cigarettes, 7 reported they were regular smokers, and 15 said they used to smoke. Among the 545 boys, 322 said they had never smoked, 187 said they had tried a few cigarettes, 20 described themselves as regular smokers, and 16 as former smokers.

a. Arrange these findings in a two-way frequency table. Use this information to construct a probability model for classifying a Vermont 4th or 5th grader by sex and smoking history.

b. Write down an appropriate sample space and probability function for this problem.

c. Find the probability of the following events: the child is a girl; the child has never tried cigarettes; the child has tried a few cigarettes; the child is a regular smoker; the child is a boy who smokes regularly; the child is a girl who has never smoked.

d. Find the following conditional probabilities: the probability the child is a girl, given the child is a regular smoker; the probability the child is a regular smoker, given the child is a girl; the probability the child is a boy, given the

child has tried a few cigarettes; the probability the child has tried a few cigarettes, given the child is a boy.

e. Are sex and smoking status independent or dependent under this probability model?

EXERCISE 6-2

In a 1984 survey for the Veterans Administration, researchers interviewed 2,033 women veterans with recent wartime service (World War II, Korean War, Vietnam War). The researchers classified these women by exposure to combat (exposed or not) and by job (nurse or not). Of the 396 nurses, 94 were exposed to combat and 302 were not. Of the 1,637 women veterans who were not nurses, 67 were exposed to combat and 1,570 were not (Dienstfrey, 1986).

a. Arrange these results in a two-way frequency table. Use this information to construct a probability model for classifying women veterans by exposure to combat and job.

b. Write down an appropriate sample space and probability function for this problem.

c. Are exposure to combat and job status independent or dependent under this probability model?

EXERCISE 6-3

An obstetrical nurse has observed over long experience that 12.75% of the babies she has helped deliver were boys with no hair, 25.5% were boys with very little hair, 12.75% were boys with lots of hair, 12.25% were girls with no hair, 24.5% were girls with very little hair, and 12.25% were girls with lots of hair.

a. If a newborn baby is classified as above by sex and by quantity of hair, construct an appropriate sample space. Use the nurse's observations to define a probability function for this sample space.

b. Under this probability model, are sex and quantity of hair of a newborn baby independent?

EXERCISE 6-4

At a toy factory, employees work around the clock producing miniature cars. Standards for the toys have been set, so that a quality control inspector can classify any given miniature car as ready for shipment or not acceptable for shipment. One week 9,000 miniature cars are produced. These cars can be classified by acceptability for shipment and by the shift that produced them, as follows:

Shift	Not acceptable	Ready for shipment	Total
Day shift, 7 A.M.–3 P.M.	100	2,900	3,000
Evening shift, 3 P.M.–11 P.M.	300	2,700	3,000
Night shift, 11 P.M.–7 A.M.	600	2,400	3,000
Total	1,000	8,000	9,000

The quality control inspector selects at random one miniature car for inspection from the 9,000 cars and classifies it according to the shift that produced it and its acceptability for shipment.

a. Write down a sample space and probability function for this experiment.

b. Is the chance that a toy car is not acceptable independent of the shift producing it?

EXERCISE 6-5

Refer to the experiment in Example 6-10 to answer parts (a)–(e). Note that part (c) uses optional material on odds ratios.

a. Write down a sample space and probability function for this experiment.

b. Find the following probabilities: the probability the baby is a boy; the probability the baby is Rh+; the probability the baby selected is a boy who is Rh+.

c. Find the following odds: the odds the baby is a girl; the odds the baby is Rh+; the odds the baby is a girl who is Rh+.

d. Find the conditional probability that the baby is a girl, given the baby is Rh+. Find the probability the baby is a girl. Why are these two probabilities equal?

e. Find the conditional probability that the baby is Rh+, given the baby is a girl. Why does this conditional probability equal the probability that the baby selected is Rh+?

EXERCISE 6-6

Last year at City Hospital, 510 of the babies born were boys and 490 were girls. Of the 510 boys, 31 were born with a condition affecting color vision called red–green colorblindness; 2 of the girls were born with this condition. Use this information to construct a probability model for classifying a newborn baby by sex and presence or absence of red–green colorblindness. [Parts (c) and (e) use optional material on odds ratios.]

a. Write down a sample space for this experiment. Find a probability function for this sample space based on the given information.

b. Find the following probabilities: the probability that the baby selected is a girl; the probability that the baby selected has red–green colorblindness; the probability that the baby selected is a girl and has red–green colorblindness.

c. Find the following odds: the odds the baby has red–green colorblindness; the odds the baby is a boy; the odds the baby is a boy with red–green colorblindness.

d. Find the following conditional probabilities: the probability the baby is a boy, given the baby has red–green colorblindness; the probability the baby has red–green colorblindness, given the baby is a boy.

e. Find the following odds: for a girl baby, the odds she has red–green colorblindness; for a baby boy, the odds he has red–green colorblindness; for a

baby with red–green colorblindness, the odds the baby is a boy; for a baby without red–green colorblindness, the odds the baby is a boy.

f. Are the event that the baby is a boy and the event that the baby has red–green colorblindness independent or dependent events?

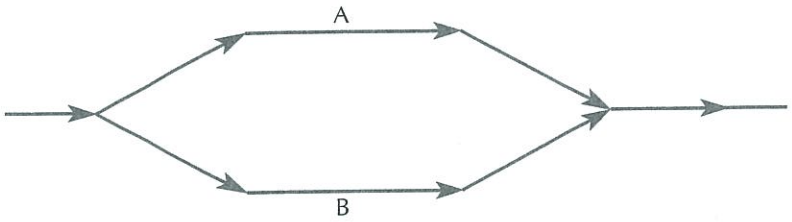
g. Red–green colorblindness is known in genetics as a sex-linked characteristic. Discuss what this means in light of your answers to parts (a)–(f) of this problem.

EXERCISE 6-7

A unit of a large machine consists of two components A and B. The probability that component A fails during operation is .01 and the probability that component B fails during operation is .05. The components A and B operate independently of one another in the sense that the functioning or failure of one of the components does not affect the likelihood of failure for the other component. Suppose that during a test run, components A and B are each tested and classified as success if functional, failure if nonfunctional.

a. Write down a sample space for this experiment. Find the probability function for this sample space. That is, based on the assumptions above, write down the probability of each outcome in the sample space.

b. Suppose the components A and B are wired in parallel, as shown here:



The unit of the machine functions if either component A or component B is functional. What is the probability that the unit is functional during operation? What is the probability that the unit fails during operation?

c. Suppose that instead components A and B are wired in series, as shown here:

The unit of the machine functions if and only if both components A and B are functional. What is the probability that the unit is functional during operation? What is the probability the unit fails during operation?

EXERCISE 6-8

Three Boston sports teams played on May 4, 1988. From the $2\frac{1}{2}$-point line favoring the Celtics over the Knicks, a basketball fan estimated that the Celtics

had probability .55 of winning that day. (This exercise uses optional material on odds ratios.)

a. *The Boston Globe* gave the odds of the Bruins beating the Devils in the Stanley Cup hockey playoffs as $8\frac{1}{2}$ to 10 in favor of the Bruins. Based on these odds, what is the estimated probability of the Bruins winning the game?

b. *The Boston Globe* favored the Red Sox with 10:12 odds over the White Sox. Based on this odds ratio, what is the estimated probability of the Red Sox winning the baseball game?

c. Let an outcome record a win or loss that day for each of the three Boston teams. Write down an appropriate sample space.

d. What does it mean to assume that the results for the three Boston teams that day were independent of one another? Assuming independence and using the estimated probabilities from above, write down a probability function for this sample space.

e. Find the probability of each of the following events: all three Boston teams win; at least one Boston team wins; no Boston team wins.

Note: That day, all three Boston teams lost. (Odds and point line from *The Boston Globe,* May 4, 1988, pages 52, 57.)

EXERCISE 6-9

Suppose a screening procedure for AIDS is not as good as the one described in Example 6-11. This screening test is positive for 90% of people with the disease and negative for 85% of people without the disease. Suppose that in the population tested, the proportion infected with the AIDS virus is .006. (This exercise uses optional material on Bayes' rule.)

a. Find the probability that a person testing positive really has AIDS. What is the probability of a false positive?

b. Find the probability that a person testing negative really does not have AIDS. What is the probability of a false negative?

c. Compare your results with those we found in Example 6-11.

EXERCISE 6-10

Suppose a screening procedure for AIDS is better than the one described in Example 6-11. This screening test is positive for 99% of people with the disease and negative for 99% of people without the disease. Suppose that in the population tested, the proportion infected with the AIDS virus is .006. (This exercise uses optional material on Bayes' rule.)

a. Find the probability that a person testing positive really has AIDS. What is the probability of a false positive?

b. Find the probability that a person testing negative really does not have AIDS. What is the probability of a false negative?

c. Compare your results with those we found in Example 6-11 and Exercise 6-9.

EXERCISE 6-11 Suppose the screening procedure described in Example 6-11 is applied to a population at high risk for AIDS—say, 10% of people in this population are infected with the AIDS virus. Suppose, as in Example 6-11, that the screening test is positive for 98% of people with the disease and negative for 93% of people without the disease. (This exercise uses optional material on Bayes' rule.)

a. Find the probability that a person testing positive really has AIDS. What is the probability of a false positive?

b. Find the probability that a person testing negative really does not have AIDS. What is the probability of a false negative?

c. Compare your results with those we found in Example 6-11.

EXERCISE 6-12 A screening procedure for AIDS is applied to a population at high risk for AIDS; 10% of this population have AIDS. This screening test is positive for 90% of people with the disease and negative for 85% of people without the disease. (This exercise uses optional material on Bayes' rule.)

a. Find the probability that a person testing positive really has AIDS. What is the probability of a false positive?

b. Find the probability that a person testing negative really does not have AIDS. What is the probability of a false negative?

c. Compare your results with those in Exercise 6-9 and Exercise 6-11.

EXERCISE 6-13 Sometimes people are automatically tested more than once in a screening program. Suppose in Example 6-11, each person is tested twice and a person is said to be positive if and only if both test results come up positive. (The same result would be achieved by retesting only people who are positive on the first test.) How does this change the probability of a false positive and the probability of a false negative? (This exercise uses optional material on Bayes' rule.)

a. Find the probability that a person with a positive result really does have AIDS. What is the probability of a false positive?

b. Find the probability that a person with a negative result really does not have AIDS. What is the probability of a false negative?

c. Compare your results with those in Example 6-11.

EXERCISE 6-14 In a country with a stable population, about .8% of all adult men develop lung cancer and 30% of all adult men are smokers. Studies of lung cancer patients have shown that 56.3% of all adult male lung cancer patients are cigarette smokers. It is also known that of the adult men who will never develop lung

cancer, 29.8% are smokers. Answer the following questions for this country. (This exercise uses optional material on odds ratios and Bayes' rule.)

a. What is the probability that an adult male cigarette smoker will develop lung cancer? What are the odds of an adult male cigarette smoker developing lung cancer?

b. What is the probability of an adult male nonsmoker developing lung cancer? What are the odds of an adult male nonsmoker developing lung cancer?

c. Compare your answers to parts (a) and (b).

d. Why is it important for these calculations to assume that the population of the country is stable? (That is, very few people are moving in and out of the country.)

e. What other simplifying assumptions are necessary for these calculations to be valid?

EXERCISE 6-15

A new *in utero* diagnostic procedure has been developed to test for a particular type of birth defect. A preliminary investigation showed the test results to be positive for 95% of the pregnant women who subsequently gave birth to babies having the birth defect under study. The results were negative for 97% of the pregnant women who later gave birth to babies not having the birth defect.

Suppose that the incidence of the birth defect in the United States is 100 per 100,000 births. That is, for every 100,000 babies born in the United States, 100 are found to have the birth defect. (This exercise uses optional material on Bayes' rule.)

a. If a woman has a positive test result, what is the probability her baby will have the birth defect? What is the probability of a false positive?

b. If a woman has a negative test result, what is the probability her baby will not have the birth defect? What is the probability of a false negative?

c. Discuss the advantages and disadvantages of using this diagnostic procedure.

EXERCISE 6-16

How useful is the lie detector or polygraph test in assessing a person's guilt or innocence? Based on studies involving 120 guilty and 120 innocent people (Gastwirth, 1987, page 217), we will estimate the probability that a guilty person is detected as guilty to be .88. We will estimate the probability that an innocent person is classified as innocent to be .86. (This exercise uses optional material on Bayes' rule.)

a. Suppose the polygraph is used among people indicted of a felony. We will assume that past experience has shown that about 90% of people indicted of a felony are in fact guilty of the felony.

(i) Find the probability that a person with a positive (guilty) lie detector result is really guilty. What is the probability of a false positive?

(ii) Find the probability that a person with a negative (not guilty) lie detector result is really innocent. What is the probability of a false negative?

b. Suppose the polygraph is used to screen for illicit drug use among employees of a large corporation. We will assume that experience with similar companies has shown that about 3% of employees are involved in illicit drug use.

(i) Find the probability that an employee with a positive (drug use) lie detector result really is an illicit drug user. What is the probability of a false positive?

(ii) Find the probability that a person with a negative (no drug problem) lie detector result is really not an illicit drug user. What is the probability of a false negative?

c. Taking all the probabilities in parts (a) and (b) at face value, discuss the importance of the proportion guilty in the test population to interpretation of lie detector results.

d. We applied the lie detector test to two different populations: people indicted of a felony and people working for a large corporation. Our assessment of the lie detector was based on baseline studies of 120 guilty people and 120 innocent people. How might the makeup of this baseline test group affect our interpretation of the results in our two applications? Discuss.

EXERCISE 6-17 A cancer researcher wants to test a new combination of chemotherapy and radiation on skin tumors in laboratory mice. The researcher administers the treatment to each of four laboratory mice having the type of skin tumor under study. After a week of treatment, the researcher records failure or success for each mouse, depending on whether or not skin tumor cells are observed on the animal.

a. Write down a sample space for this problem, where an outcome shows success or failure for each of the four animals in the experiment.

b. Suppose that the treatment combination has no effect on the skin tumors. However, there is a .1 probability that the tumor will disappear spontaneously over a week of observation. If the mice are independent of one another with regard to disappearance of the skin tumor, find the probability function for the sample space in this problem.

c. With the probability model above, what is the probability that three or four of the mice will be free of skin tumors at the end of the week of treatment? What is the probability that one or none of the mice will be free of skin tumors at the end of the week of treatment?

d. Define a random variable X to be the number of mice that are free of skin tumors at the end of the week of treatment. Find the values X takes on and the probability that X takes on each of these values. Find the expected value, variance, and standard deviation of X.

EXERCISE 6-18

The experiment is the same as in Exercise 6-17. Suppose now that the treatment does have an effect on the skin tumor. Assume there is a probability of .8 that the skin tumor will disappear by the end of the week of treatment.

a. If mice are independent of one another with respect to disappearance of skin tumors, find the probability function for the sample space in this problem.

b. With this probability model, what is the probability that three or four of the mice will be free of skin tumors at the end of the week of treatment? What is the probability that one or none of the mice will be free of skin tumors at the end of the week of treatment?

c. Define a random variable X to be the number of mice that are free of skin tumors at the end of the week of treatment. Find the values X takes on and the probability that X takes on each of these values. Find the expected value, variance, and standard deviation of X.

d. How do the answers to this exercise compare with the answers to Exercise 6-17?

EXERCISE 6-19

Ralph's history teacher gives a surprise quiz one day in class. The quiz consists of four multiple-choice questions. Each question has six answers to choose from, one being the correct answer and the other five incorrect. Suppose that Ralph has not studied history all semester and that he guesses on this quiz in such a way that he is equally likely to choose any one of the six answers to any question and his answer to one question does not in any way influence his answer to another question. (Maybe he rolls a fair die to decide on his answer.) Suppose an experiment consists of Ralph taking the quiz. An outcome shows whether Ralph's answer was correct or incorrect for each of the four questions.

a. Write down a sample space and find the probability of each outcome in the sample space.

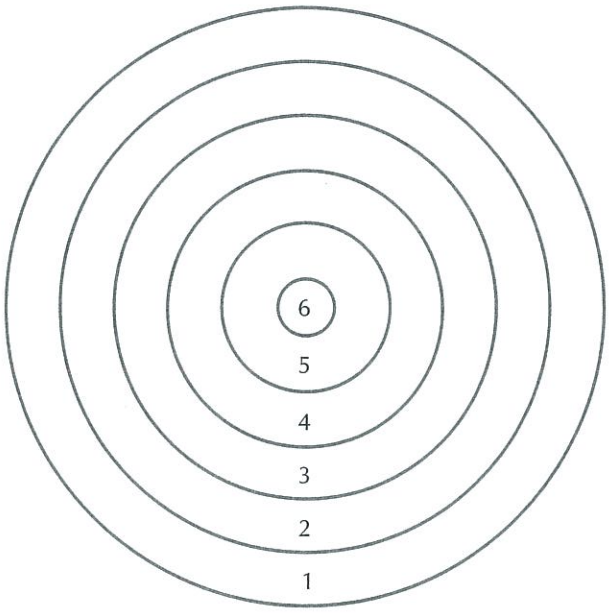
b. Define a random variable W that counts the number of correct answers Ralph guesses on the quiz. List the possible values for W and the probability that W takes on each of these values.

c. Find the expected value, variance, and standard deviation of W.

d. What is the probability that Ralph guesses at least two answers correctly? That is, find $P(W \geq 2)$.

EXERCISE 6-20

Vin throws a dart at a target, with each throw scored in the following way. A shot hitting the bull's eye in the center of the target is worth 6 points. Hits in concentric circles about the center are worth from 5 points down to 1 point, 5 points within the circle closest to the center and 1 point for the circle farthest from the center. A shot that misses the target is worth 0 points. Vin is a good dart player and experience has shown that under ideal conditions his likelihood of achieving any given score is the following:



Score	Probability
6	.15
5	.25
4	.20
3	.15
2	.10
1	.10
0	.05

Let the random variable Y represent the score Vin gets with one throw under ideal conditions.

a. Find $P(Y \geq 4)$, the probability that Vin gets 4 or more points.

b. Find $P(Y < 3)$, the probability that Vin gets less than 3 points.

c. Find $P(3 \leq Y \leq 5)$, the probability that Vin gets 3, 4, or 5 points.

d. Find the expected value, variance, and standard deviation of Y. How would you interpret each of these values?

EXERCISE 6-21

An experiment consists of rolling a six-sided die once and noting how many dots are face up at the end of the roll. This is a fair die, so that each of the six sides is equally likely to be face up at the end of the roll. [Part (b) uses optional material on odds ratios.]

a. Write down the sample space for this experiment and the probability function for this sample space.

b. What are the odds that 1, 2, or 3 dots will be face up at the end of the roll?

Suppose a gambler wins \$.50, \$1.50, or \$2.50 if the die comes up 4, 5, or 6, respectively. He loses \$.50, \$1.50, or \$2.50 if the die comes up 3, 2, or 1, respectively. Let the random variable W denote the gambler's winnings at the end of one die roll, where a loss is the same as a negative gain.

c. Find the values that W takes on and the probability that W takes on each of these values.

d. Find $P(W > 0)$, the probability that the gambler wins some money.

e. Calculate the gambler's expected gain, which is the same as the expected value of W.

f. Find the variance and standard deviation of W.

EXERCISE 6-22

An experiment consists of rolling a six-sided die once and noting how many dots are face up at the end of the roll. For this die, the chance of a particular side being face up at the end of a roll is proportional to the number of dots on that side. [Part (b) uses optional material on odds ratios.]

a. Write down the sample space S for this experiment. Find the probability function for S. (*Hint:* Let $P(i) = c \times i$, where c is a constant and i denotes the number of dots face up at the end of the roll. Show that the probabilities of the outcomes in S sum to 1 if $c = \frac{1}{21}$.)

b. What are the odds that 1, 2, or 3 dots will be face up at the end of the roll?

Suppose a gambler wins \$.50, \$1.50, or \$2.50 if the die comes up 4, 5, or 6, respectively. He loses \$.50, \$1.50, or \$2.50 if the die comes up 3, 2, or 1, respectively. Let the random variable W denote the gambler's winnings at the end of one die roll, where a loss is the same as a negative gain.

c. Find the values that W takes on and the probability that W takes on each of these values.

d. Find $P(W > 0)$, the probability that the gambler wins some money.

e. Calculate the gambler's expected gain, which is the same as the expected value of W.

f. Find the variance and standard deviation of W.

g. Compare your answers to this exercise with your answers to Exercise 6-21.

EXERCISE 6-23

A six-sided die is rolled once and each outcome is equally likely. A gambler wins \$1 for each dot that appears face up at the end of the roll. Let the random variable Y denote the gambler's winnings at the end of one die roll.

a. Find the values Y takes on and the probability that Y takes on each of these values.

b. Find $P(Y > 0)$, the probability that the gambler wins some money.

c. Find the expected value of Y, the gambler's expected winnings.

d. Find the variance and standard deviation of Y.

e. Compare these answers with the answers to Exercise 6-21.

EXERCISE 6-24

A six-sided die is rolled once and the probabilities of the outcomes are those you found in Exercise 6-22. A gambler wins \$1 for each dot that appears face up at the end of the roll. Let the random variable Y denote the gambler's winnings at the end of one die roll.

a. Find the values Y takes on and the probability that Y takes on each of these values.

b. Find $P(Y > 0)$, the probability that the gambler wins some money.

c. Find the expected value of Y, the gambler's expected winnings.

d. Find the variance and standard deviation of Y.

e. Compare these answers with the answers to Exercises 6-22 and 6-23.

EXERCISE 6-25 Suppose X is a finite random variable. Show that if $P(X = c) = 1$ for some constant c, then the variance of X is 0. Show that if X takes on at least two different values with positive probability, then the variance of X is greater than 0.

EXERCISE 6-26 If no treatment program is implemented, an outbreak of a rare disease is expected to kill 600 people. With one possible treatment plan, call it program B, there is a $\frac{1}{3}$ probability that all 600 of the disease victims will be saved and a $\frac{2}{3}$ probability that none of them will be saved. Find the expected number of disease victims saved under this treatment plan. (*Hint:* Let the random variable X denote the number of disease victims saved under the program B treatment plan. Find the expected value of X.)

EXERCISE 6-27 If no treatment program is implemented, an outbreak of a rare disease is expected to kill 600 people. With one possible treatment plan, call it program D, there is a $\frac{2}{3}$ probability that all 600 disease victims will die and a $\frac{1}{3}$ probability that none of them will die. Find the expected number of deaths under this treatment plan. (*Hint:* Let the random variable Y denote the number of disease victims who die under the program D treatment plan. Find the expected value of Y.)

EXERCISE 6-28 This problem illustrates why a special definition of independence for several events is necessary. Suppose a fair die is rolled twice, and the two rolls are independent. (A fair die is one with each of its six sides equally likely to be face up at the end of a roll.) Let an outcome be represented by the number of dots face up on each of the two rolls. For instance, (3, 5) indicates that 3 dots came up on the first roll and 5 dots came up on the second roll.

a. Write down a sample space and probability function for this experiment.

Let A be the event that the first roll results in 1, 2, or 3 dots face up. Let B be the event that the second roll results in 4, 5, or 6 dots face up. Let C be the event that the total number of dots face up on the two rolls is 3 or 11. Let D be the event that the second roll results in 2, 5, or 6 dots face up.

b. Find the probability of each of these events: A, B, C, A and B, A and C, B and C, A and B and C.

c. Use your answer to part (b) to show that A, B, and C are pairwise independent but the probability of their intersection does not equal the product of their separate probabilities.

d. Find the probability of each of these events: *A, C, D, A* and *C, A* and *D, C* and *D, A* and *C* and *D.*

e. Use your answer to part (d) to show that the probability of *A* and *C* and *D* equals the product of their separate probabilities. Show that *A, C,* and *D* are *not* pairwise independent.

(*Note:* To say several events are pairwise independent means that for any two of the events, the probability of their intersection equals the product of their separate probabilities.)

EXERCISE 6-29

Joe DiMaggio had a lifetime batting average of .325. Suppose .325 was the probability of his getting a hit in any single time at bat. Suppose also that his times at bat were independent of each other as far as his getting a hit was concerned. With these assumptions, calculate the probability that Joe DiMaggio would have at least one hit in four at-bats during a single game. How reasonable are the assumptions you made in order to calculate this probability? (This exercise is based on an article by Tom Short and Larry Wasserman, "Should We Be Surprised by the Streak of Streaks?" *Chance,* Volume 2, Spring 1989, page 13.)

EXERCISE 6-30

In a simple genetic model for a characteristic, an individual has two genes for the characteristic. A gene can be one of two types in this model: dominant or recessive. If the individual has at least one dominant-type gene, then the dominant form of the characteristic is expressed. If the individual has two recessive-type genes, then the recessive form of the characteristic is expressed. Phenotype refers to the form of the characteristic expressed. If we denote the dominant form by **B** and the recessive form by **b**, then the gene combinations (or genotypes) **BB**, **Bb**, and **bB** result in phenotype **B**. The genotype **bb** results in phenotype **b**.

a. In a hybrid cross in genetics, both parents have one dominant gene and one recessive gene for the characteristic. An offspring receives one gene for the characteristic from each parent. In the simplest Mendelian genetics model of independent selection, we assume that each of a parent's two genes is equally likely to be passed on to the offspring. We also assume that the parents are independent of one another with regard to the gene passed on to the offspring. With these assumptions, develop a probability model for the genotype of the offspring. Show that under this model we expect a 3:1 ratio of dominant to recessive phenotypes among the offspring of the hybrid cross.

b. In a dihybrid cross, both parents have one dominant and one recessive gene for each of two characteristics. Say, each parent has a dominant gene **B** and a recessive gene **b** for the first characteristic, and a dominant gene **C** and a recessive gene **c** for the second characteristic. In the simplest Mendelian genetics model of independent selection, we assume that for each

characteristic, each of a parent's two genes is equally likely to be passed on to the offspring. We also assume independence between characteristics and between parents with regard to the genes passed on to the offspring. With these assumptions, develop a probability model for the genotype of the offspring for these two characteristics. Show that under this model, we expect a 9:3:3:1 ratio of the phenotypes **BC**:**Bc**:**bC**:**bc** (dominant for both characteristics:dominant for the first and recessive for the second:recessive for the first and dominant for the second:recessive for both).