PART 1

Getting the Information You Need

CHAPTER 1
Data Collection

Statistics is a process—a series of steps that leads to a goal. This text is divided into four parts to help the reader see the process of statistics.

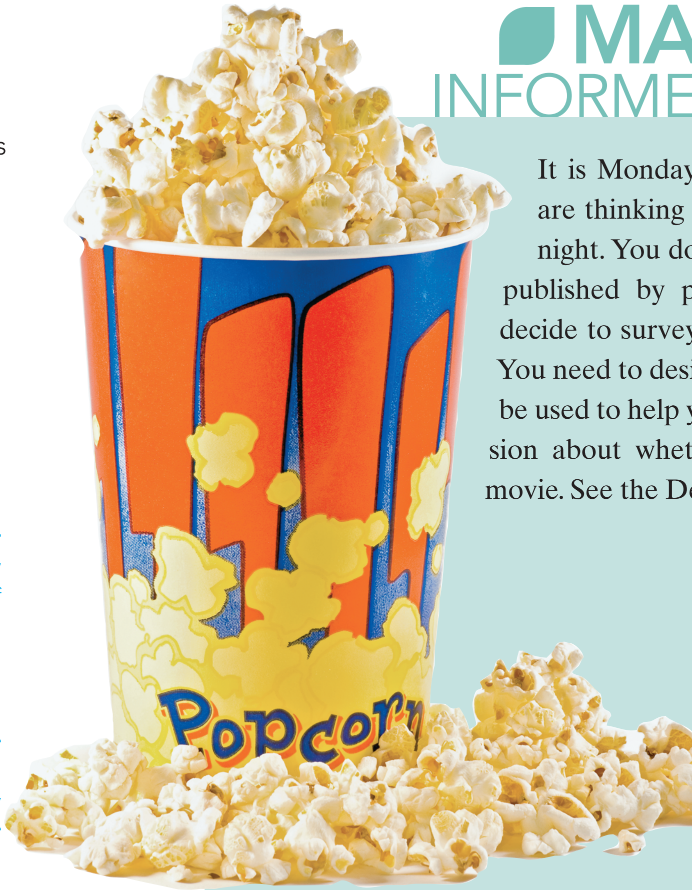
Part 1 focuses on the first step in the process, which is to determine the research objective or question to be answered. Then information is obtained to answer the questions stated in the research objective.

1 Data Collection

Outline

Note to Instructor

Due to the tremendous number of terms in this chapter, you may want students to create a list of terms, definitions, and their own examples for each term. This will allow students to make a personal connection to the concepts and will lead to a greater understanding of the vocabulary. Remind students that many terms presented in this chapter will be presented again in upcoming chapters, so they should not feel overwhelmed by the number of definitions.



MAKING AN INFORMED DECISION

It is Monday morning and already you are thinking about Friday night—movie night. You don't trust the movie reviews published by professional critics, so you decide to survey "regular" people yourself. You need to design a questionnaire that can be used to help you make an informed decision about whether to attend a particular movie. See the Decisions project on page 62.

PUTTING IT TOGETHER

Statistics plays a major role in many different areas of our lives. For example, it is used in sports to help a general manager decide which player might be the best fit for a team. It is used by politicians to help them understand how the public feels about various governmental policies. Statistics is used to help determine the effectiveness (efficacy) of experimental drugs.

Used appropriately, statistics can provide an understanding of the world around us. Used inappropriately, it can lend support to inaccurate beliefs. Understanding the methodologies of statistics will provide you with the ability to analyze and critique studies. With this ability, you will be an informed consumer of information, which will enable you to distinguish solid analysis from the bogus presentation of numerical "facts."

To help you understand the features of this text, and for hints to help you study, read the Pathway to Success on the front inside cover of the text.

1.1 INTRODUCTION TO THE PRACTICE OF STATISTICS

Objectives

1. Define statistics and statistical thinking
2. Explain the process of statistics
3. Distinguish between qualitative and quantitative variables
4. Distinguish between discrete and continuous variables
5. Determine the level of measurement of a variable

1 Define Statistics and Statistical Thinking

Note to Instructor

This first objective defines what statistics is, but also lists some key elements to solid statistical thinking.

1. Anecdotal claims can be refuted with statistical analysis.
2. Poorly collected data are not useful.
3. Watch out for lurking variables in a study.
4. Results in statistics are not certain.

The Guidelines for Assessment and Instruction in Statistics Education (GAISE) report suggests these as major components of statistical thinking.

What is statistics? When asked this question, many people respond that statistics is numbers. After all, we are bombarded by numbers that supposedly represent how we feel and who we are. For example, we hear on the radio that 50% of first marriages, 67% of second marriages, and 74% of third marriages end in divorce (Forest Institute of Professional Psychology, Springfield, MO).

Another interesting consideration about the "facts" we hear or read is that two different sources can report two different results. For example, a July 12–15, 2007, Gallup poll indicated that 66% of Americans disapproved of the job that Congress was doing. However, a July 25–29, 2007, poll conducted by the Pew Research Center indicated that 54% of Americans disapproved of the job that Congress was doing. Is it possible that Congress's disapproval rating could decrease by 12% in less than 2 weeks or is something else going on? Statistics helps to provide the answer.

Certainly, statistics has a lot to do with numbers, but this definition is only partially correct. Statistics is also about where the numbers come from (that is, how they were obtained) and how closely the numbers reflect reality.

Definition

Statistics is the science of collecting, organizing, summarizing, and analyzing information to draw conclusions or answer questions. In addition, statistics is about providing a measure of confidence in any conclusions.

It is helpful to consider this definition in four parts. The first part of the definition states that statistics involves the collection of information. The second refers to the organization and summarization of information. The third states that the information is analyzed to draw conclusions or answer specific questions. The fourth part states that results should be reported with some measure that represents how convinced we are that our conclusions reflect reality.

What is the information referred to in the definition? The information is *data*. The *American Heritage Dictionary* defines **data** as "a fact or proposition used to draw a conclusion or make a decision." Data can be numerical, as in height, or nonnumerical, as in gender. In either case, data describe characteristics of an individual. The reason that data are important in statistics can be seen in this definition: data are used to draw a conclusion or make a decision.

In Other Words

Anecdotal means that the information being conveyed is based on casual observation, not scientific research.

Analysis of data can lead to powerful results. Data can be used to offset anecdotal claims, such as the suggestion that cellular telephones cause brain cancer. After carefully collecting, summarizing, and analyzing data regarding this phenomenon, it was determined that there is no link between cell phone usage and brain cancer. See Examples 1 and 2 in Section 1.2.

Because data are powerful, they can be dangerous when misused. The misuse of data usually occurs when data are incorrectly obtained or analyzed. For example, radio or television talk shows regularly ask poll questions for which respondents must call in or use the Internet to supply their vote. Most likely, the individuals who are going to call in are those that have a strong opinion about the topic. This group is not likely to be representative of people in general, so the results of the poll are not meaningful. Whenever we look at data, we should be mindful of where the data come from.

Even when data tell us that a relation exists, we need to investigate. For example, a study showed that breast-fed children have higher IQs than those who were not breast-fed. Does this study mean that mothers should breast-feed their children? Not necessarily. It may be that some other factor contributes to the IQ of the children. In this case, it turns out that mothers who breast-feed generally have higher IQs than those who do not. Therefore, it may be genetics that leads to the higher IQ, not breast-feeding. This illustrates an idea in statistics known as the *lurking variable*. In statistics, we must consider lurking variables, because two variables are often influenced by a third variable. A good statistical study will have a way of dealing with lurking variables.

A key aspect of data is that they vary. To help understand this variability, consider the students in your classroom. Is everyone the same height? No. Does everyone have the same color hair? No. So, among a group of individuals there is variation. Now consider yourself. Do you eat the same amount of food each day? No. Do you sleep the same number of hours each day? No. So, even looking at an individual there is variation. Data vary. One goal of statistics is to describe and understand the sources of variation. Variability in data may help to explain the different results obtained by Gallup and Pew mentioned at the beginning of this section.

Because of this variability, the results that we obtain using data can vary. This is a very different idea than what you may be used to from your mathematics classes. In mathematics, if Bob and Jane are asked to solve $3x + 5 = 11$, they will both obtain $x = 2$ as the solution, if they use the correct procedures. In statistics, if Bob and Jane are asked to estimate the average commute time for workers in Dallas, Texas, they will likely get different answers, even though they both use the correct procedure. The different answers occur because they likely surveyed different individuals, and these individuals have different commute times. Note: The only way Bob and Jane would get the same result is if they both asked *all* commuters or the same commuters how long it takes to get to work, but how likely is this?

So, in mathematics when a problem is solved correctly, the results can be reported with 100% certainty. In statistics, when a problem is solved, the results do not have 100% certainty. In statistics, we might say that we are 95% confident that the average commute time in Dallas, Texas, is between 20 and 23 minutes. While uncertain results may sound disturbing now, it will become more apparent what this means as we proceed through the course.

Without certainty, how can statistics be useful? Statistics can provide an understanding of the world around us because recognizing where variability in data comes from can help us to control it. Understanding the techniques presented in this text will provide you with powerful tools that will give you the ability to analyze and critique media reports, make investment decisions (such as what mutual fund to invest in), or conduct research on major purchases (such as what type of car you should buy). This will help to make you an informed consumer of information and guide you in becoming a critical and statistical thinker.

2 Explain the Process of Statistics

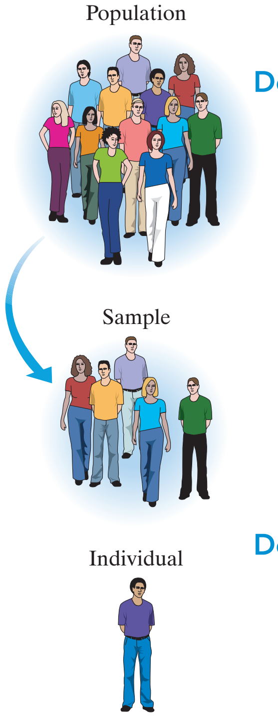
Consider the following scenario.

> You are walking down the street and notice that a person walking in front of you drops \$100. Nobody seems to notice the \$100 except you. Since you could keep the money without anyone knowing, would you keep the money or return it to the owner?

Note: Certainly, obtaining a truthful response to a question such as this is challenging. In Section 1.5, we present some techniques for obtaining truthful responses to sensitive questions. ◀

Suppose you wanted to use this scenario as a gauge of the morality of students at your school by determining the percent of students who would return the money. How might you go about doing this? Well, you could attempt to present the scenario to every student at the school, but this is likely to be difficult or impossible since the number of enrolled students is likely large. A second possibility is to present the scenario to 50 students and use the results from these 50 students to make a statement about all the students at the school.

Figure 1



Definitions

The entire group of individuals to be studied is called the **population**. An **individual** is a person or object that is a member of the population being studied. A **sample** is a subset of the population that is being studied. See Figure 1.

In the scenario presented, the population is all the students at the school. Each student is an individual. The sample is the 50 students selected to participate in the study.

Suppose 39 of the 50 students stated that they would return the money to the owner. We could present this result by saying the percent of students in the survey that stated they would return the money to the owner is 78%. This is an example of a *descriptive statistic* because it describes the results of the sample without making any general conclusions about the population.

Definitions

A **statistic** is a numerical summary of a sample. **Descriptive statistics** consist of organizing and summarizing data. Descriptive statistics describe data through numerical summaries, tables, and graphs.

So 78% is a statistic because it is based on a sample. Descriptive statistics make it easier to get an overview of what the data are telling us.

If we extend the results of our sample to the population and say that the proportion of all students at the school who would return the money is 78%, we are performing *inferential statistics.*

Definition

Inferential statistics uses methods that take a result from a sample, extend it to the population, and measure the reliability of the result.

When generalizing to a population from a sample, there is always uncertainty as to the accuracy of the generalization, because we cannot learn everything about a population by looking at a sample. Therefore, when performing inferential statistics, we always report a measure that quantifies how confident we are in our results. So, rather than saying that 78% of all students would return the money, we might say that we are 95% confident that between 76% and 80% of all students would return the money. Notice how this inferential statement includes a level of confidence (measure of reliability) in our results. It also provides a range of values to account for the variability in our results.

One goal of inferential statistics is to use statistics to estimate *parameters*.

Definition

A **parameter** is a numerical summary of a population.

EXAMPLE 1 Parameter versus Statistic

Suppose the percentage of all students on your campus that own a car is 48.2%. This value represents a parameter because it is a numerical summary of a population. Suppose a sample of 100 students is obtained, and from this sample we find that 46% own a car. This value represents a statistic because it is a numerical summary of a sample.

Now Work Problem 13

The methods of statistics follow a process.

CAUTION Many nonscientific studies are based on *convenience samples*, such as Internet surveys or phone-in polls. The results of any study performed using this type of sampling method are not reliable.

Note to Instructor

Notice how the parts of the text are divided according to the process of statistics. As you enter a new part of the course, be sure to remind students where they are in the process of statistics.

The Process of Statistics

1. *Identify the research objective.* A researcher must determine the question(s) he or she wants answered. The question(s) must be detailed so that it identifies the population that is to be studied.

2. *Collect the data needed to answer the question(s) posed in (1).* Gaining access to an entire population is often difficult and expensive. When conducting research, we typically look at a sample. The collection-of-data step is vital to the statistical process, because if the data are not collected correctly, the conclusions drawn are meaningless. Do not overlook the importance of appropriate data-collection processes. We discuss this step in detail in Sections 1.2 through 1.6.

3. *Describe the data.* Obtaining descriptive statistics allows the researcher to obtain an overview of the data and can provide insight as to the type of statistical methods the researcher should use. We discuss this step in detail in Chapters 2 through 4.

4. *Perform inference.* Apply the appropriate techniques to extend the results obtained from the sample to the population and report a level of reliability of the results. We discuss techniques for measuring reliability in Chapters 5 through 8 and inferential techniques in Chapters 9 through 15.

EXAMPLE 2 The Process of Statistics: Do You Favor Stricter Gun Laws?

A poll was conducted by the Gallup Organization on October 4–7, 2007, to learn how Americans feel about existing gun-control laws. The following statistical process allowed the researchers at Gallup to conduct their study.

1. *Identify the research objective.* The researchers wished to determine the percentage of Americans aged 18 years or older who were in favor of more strict gun-control laws. Therefore, the population being studied was Americans aged 18 years or older.

2. *Collect the information needed to answer the question posed in (1).* It is unreasonable to expect to survey the more than 200 million Americans aged 18 years or older to determine how they feel about gun-control laws. So the researchers surveyed a sample of 1,010 Americans aged 18 years or older. Of those surveyed, 515 stated they were in favor of more strict laws covering the sale of firearms.

3. *Describe the data.* Of the 1,010 individuals in the survey, 51% (= 515/1,010) are in favor of more strict laws covering the sale of firearms. This is a descriptive statistic because its value is determined from a sample.

4. *Perform inference.* The researchers at Gallup wanted to extend the results of the survey to all Americans aged 18 years or older. Remember, when generalizing results from a sample to a population, the results are uncertain. To account for this uncertainty, Gallup reported a 3% *margin of error.* This means that Gallup feels fairly certain (in fact, Gallup is 95% certain) that the percentage of *all* Americans aged 18 years or older in favor of more strict laws covering the sale of firearms is somewhere between 48% (51% − 3%) and 54% (51% + 3%).

Now Work Problem 57

3 Distinguish between Qualitative and Quantitative Variables

Note to Instructor
Emphasize the point that variables vary. This is key to helping students to develop statistical thinking.

Once a research objective is stated, a list of the information the researcher desires about the individuals must be created. **Variables** are the characteristics of the individuals within the population. For example, this past spring my son and I planted a tomato plant in our backyard. We decided to collect some information about the tomatoes harvested from the plant. The individuals we studied were the tomatoes. The variable that interested us was the weight of the tomatoes. My son noted that the tomatoes had different weights even though they all came from the same plant. He discovered that variables such as weight vary.

If variables did not vary, they would be constants, and statistical inference would not be necessary. Think about it this way: If all the tomatoes had the same weight, then knowing the weight of one tomato would be sufficient to determine the weights of all tomatoes. However, the weights vary from one tomato to the next. One goal of research is to learn the causes of the variability so that we can learn to grow plants that yield the best tomatoes.

Variables can be classified into two groups: *qualitative* or *quantitative*.

Definitions

Qualitative, or categorical, variables allow for classification of individuals based on some attribute or characteristic.

Quantitative variables provide numerical measures of individuals. Arithmetic operations such as addition and subtraction can be performed on the values of a quantitative variable and will provide meaningful results.

Many examples in this text will include a suggested **approach**, or a way to look at and organize a problem so that it can be solved. The approach will be a suggested method of *attack* toward solving the problem. This does not mean that the approach given is the only way to solve the problem, because many problems have more than one approach leading to a correct solution. For example, if you turn the key in your car's ignition and it doesn't start, one approach would be to look under the hood to try to determine what is wrong. (Of course, this approach will work only if you know how to fix cars.) A second, equally valid approach would be to call an automobile mechanic to service the car.

In Other Words
Typically, there is more than one correct approach to solving a problem.

EXAMPLE 3 Distinguishing between Qualitative and Quantitative Variables

Problem: Determine whether the following variables are qualitative or quantitative.

(a) Gender

(b) Temperature

(c) Number of days during the past week that a college student aged 21 years or older has had at least one drink

(d) Zip code

Approach: Quantitative variables are numerical measures such that meaningful arithmetic operations can be performed on the values of the variable. Qualitative variables describe an attribute or characteristic of the individual that allows researchers to categorize the individual.

Solution

(a) Gender is a qualitative variable because it allows a researcher to categorize the individual as male or female. Notice that arithmetic operations cannot be performed on these attributes.

(b) Temperature is a quantitative variable because it is numeric, and operations such as addition and subtraction provide meaningful results. For example, 70°F is 10°F warmer than 60°F.

(c) Number of days during the past week that a college student aged 21 years or older had at least one drink is a quantitative variable because it is numeric, and operations such as addition and subtraction provide meaningful results.

(d) Zip code is a qualitative variable because it categorizes a location. Notice that, even though they are numeric, the addition or subtraction of zip codes does not provide meaningful results.

Now Work Problem 21

On the basis of the result of Example 3(d), we conclude that a variable may be qualitative while having values that are numeric. Just because the value of a variable is numeric does not mean that the variable is quantitative.

4 Distinguish between Discrete and Continuous Variables

We can further classify quantitative variables into two types: *discrete* or *continuous*.

Definitions

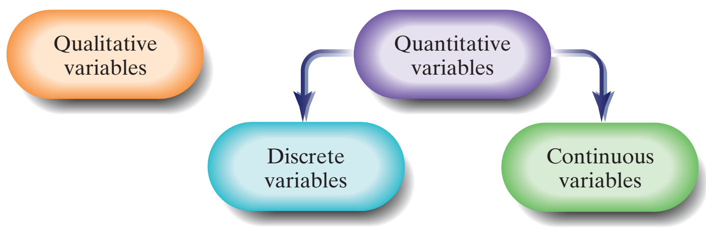
A **discrete variable** is a quantitative variable that has either a finite number of possible values or a countable number of possible values. The term *countable* means that the values result from counting, such as 0, 1, 2, 3, and so on.

A **continuous variable** is a quantitative variable that has an infinite number of possible values that are not countable.

In Other Words
If you count to get the value of a quantitative variable, it is discrete. If you measure to get the value of a quantitative variable, it is continuous. When deciding whether a variable is discrete or continuous, ask yourself if it is counted or measured.

Figure 2 illustrates the relationship among qualitative, quantitative, discrete, and continuous variables.

Figure 2



An example should help to clarify the definitions.

EXAMPLE 4 Distinguishing between Discrete and Continuous Variables

Problem: Determine whether the following quantitative variables are discrete or continuous.

(a) The number of heads obtained after flipping a coin five times.

(b) The number of cars that arrive at a McDonald's drive-through between 12:00 P.M. and 1:00 P.M.

(c) The distance a 2007 Toyota Prius can travel in city driving conditions with a full tank of gas.

Approach: A variable is discrete if its value results from counting. A variable is continuous if its value is measured.

Solution

(a) The number of heads obtained by flipping a coin five times would be a discrete variable because we would count the number of heads obtained. The possible values of the discrete variable are 0, 1, 2, 3, 4, 5.

(b) The number of cars that arrive at a McDonald's drive-through between 12:00 P.M. and 1:00 P.M. is a discrete variable because its value would result from counting the cars. The possible values of the discrete variable are 0, 1, 2, 3, 4, and so on. Notice that there is no predetermined upper limit to the number of cars that may arrive.

(c) The distance traveled is a continuous variable because we measure the distance.

Now Work Problem 29

Continuous variables are often rounded. For example, when the miles per gallon (mpg) of gasoline for a certain make of car is given as 24 mpg, it means that the miles per gallon is greater than or equal to 23.5 and less than 24.5, or $23.5 \le \text{mpg} < 24.5$.

The type of variable (qualitative, discrete, or continuous) dictates the methods that can be used to analyze the data.

The list of observed values for a variable is **data**. Gender is a variable; the observations male or female are data. **Qualitative data** are observations corresponding to a qualitative variable. **Quantitative data** are observations corresponding to a quantitative variable. **Discrete data** are observations corresponding to a discrete variable, and **continuous data** are observations corresponding to a continuous variable.

EXAMPLE 5 Distinguishing between Variables and Data

Problem: Table 1 presents a group of selected countries and information regarding these countries as of October, 2007. Identify the individuals, variables, and data in Table 1.

Table 1

Country	Government Type	Life Expectancy (years)	Population (in millions)
Australia	Federal parliamentary democracy	80.62	20.4
Canada	Constitutional monarchy	80.34	33.4
France	Republic	80.59	63.7
Morocco	Constitutional monarchy	71.22	33.8
Poland	Republic	75.19	38.52
Sri Lanka	Republic	74.80	20.93
United States	Federal republic	78.00	301.14

Source: CIA *World Factbook*

Approach: An individual is an object or person for whom we wish to obtain data. The variables are the characteristics of the individuals, and the data are the specific values of the variables.

Solution: The individuals in the study are the countries: Australia, Canada, and so on (in red ink). The variables measured for each country are *government type*, *life expectancy*, and *population* (in blue ink). The variable *government type* is qualitative because it categorizes the individual. The variables *life expectancy* and *population* are quantitative.

The quantitative variable *life expectancy* is continuous because it is measured. The quantitative variable *population* is discrete because we count people. The observations are the data (in green ink). For example, the data corresponding to the variable *life expectancy* are 80.62, 80.34, 80.59, 71.22, 75.19, 74.80, and 78.00. The following data correspond to the individual Poland: a *republic* government with residents whose *life expectancy* is 75.19 years and where *population* is 38.52 million people. Republic is an instance of qualitative data that results from observing the value of the qualitative variable *government type*. The life expectancy of 75.19 years is an instance of quantitative data that results from observing the value of the quantitative variable *life expectancy*.

Now Work Problem 51

5 Determine the Level of Measurement of a Variable

Rather than classify a variable as qualitative or quantitative, we can assign a level of measurement to the variable.

Definitions

A variable is at the **nominal level of measurement** if the values of the variable name, label, or categorize. In addition, the naming scheme does not allow for the values of the variable to be arranged in a ranked or specific order.

A variable is at the **ordinal level of measurement** if it has the properties of the nominal level of measurement and the naming scheme allows for the values of the variable to be arranged in a ranked or specific order.

A variable is at the **interval level of measurement** if it has the properties of the ordinal level of measurement and the differences in the values of the variable have meaning. A value of zero in the interval level of measurement does not mean the absence of the quantity. Arithmetic operations such as addition and subtraction can be performed on values of the variable.

A variable is at the **ratio level of measurement** if it has the properties of the interval level of measurement and the ratios of the values of the variable have meaning. A value of zero in the ratio level of measurement means the absence of the quantity. Arithmetic operations such as multiplication and division can be performed on the values of the variable.

In Other Words

The word *nominal* comes from the Latin word *nomen*, which means to name. When you see the word *ordinal*, think order.

Variables that are nominal or ordinal are qualitative variables, while variables that are interval or ratio are quantitative variables.

EXAMPLE 6 Determining the Level of Measurement of a Variable

Problem: For each of the following variables, determine the level of measurement.

(a) Gender

(b) Temperature

(c) Number of days during the past week that a college student aged 21 years or older has had at least one drink

(d) Letter grade earned in your statistics class

Approach: For each variable, we ask the following: Does the variable simply categorize each individual? If so, the variable is nominal. Does the variable categorize *and* allow ranking of each value of the variable? If so, the variable is ordinal. Do differences in values of the variable have meaning, but a value of zero does not mean the absence of the quantity? If so, the variable is interval. Do ratios of values of the variable have meaning *and* there is a natural zero starting point? If so, the variable is ratio.

Solution

(a) Gender is a variable measured at the nominal level because it only allows for categorization of male or female. Plus, it is not possible to rank gender classifications.

(b) Temperature is a variable measured at the interval level because differences in the value of the variable make sense. For example, 70°F is 10°F warmer than 60°F. Notice that the ratio of temperatures does not represent a meaningful result. For example, 60°F is not twice as warm as 30°F. In addition, 0°F does not represent the absence of heat.

(c) Number of days during the past week that a college student aged 21 years or older has had at least one drink is measured at the ratio level, because the ratio of two values makes sense and a value of zero has meaning. For example, a student who had four drinks had twice as many drinks as a student who had two drinks.

(d) Letter grade is a variable measured at the ordinal level because the values of the variable can be ranked, but differences in values have no meaning. For example, an A is better than a B, but A – B has no meaning.

Now Work Problem 37

When classifying variables according to their level of measurement, it is extremely important to be careful to recognize what the variable is intended to measure. For example, suppose we want to know whether cars with 4-cylinder engines get better gas mileage than cars with 6-cylinder engines. Here, engine size represents a category of data and so the variable is nominal. On the other hand, if we want to know the average number of cylinders in cars in the United States, the variable is classified as ratio (an 8-cylinder engine has twice as many cylinders as a 4-cylinder engine).

Note to Instructor

The purpose of this activity is to expose students to variability in measuring variables. It also allows students to see the importance of defining a variable. It should take about 20 to 30 minutes.

IN CLASS ACTIVITY

Validity, Reliability, and Variability

Divide the class into groups of four to six students.

(a) Select one student to be the group leader. Each student in the group measures the length of the right arm of the group leader. As the group leader is being measured, the other students in the group do not look on. Do not share the measurements obtained with others in the group until everyone has obtained a measurement! Record the results.

(b) The group leader measures the length of the right arm of each of the other students in the group. Record the results.

(c) **Validity** of a variable or measurement represents how close to the true value the measurement is. In other words, a variable is valid if it measures what it is supposed to measure. For example, if a student measured arm length from the shoulder to the wrist and another student measured arm length from the shoulder to the tip of the middle finger, the variable is not valid. How valid are the results obtained from part (a)? What could have been done by the group to increase the validity of the variable?

(d) **Reliability** of a variable or measurement represents the ability of different measurements of the same individual to yield the same results. How reliable are the measurements obtained in part (b)? Why is it likely that the results from part (b) are valid, but may not be reliable?

(e) Which set of data appears to have more variability, the data from part (a) or the data from part (b)? Why?

(f) Compare the results of all the groups. Which group do you think has the most valid results? Which group has the most reliable results?

1.1 ASSESS YOUR UNDERSTANDING

Concepts and Vocabulary

1. Define statistics.
2. Explain the difference between a population and a sample.
3. A(n) <u>individual</u> is a person or object that is a member of the population being studied.
4. <u>Descriptive</u> statistics consists of organizing and summarizing information collected, while <u>inferential</u> statistics uses methods that generalize results obtained from a sample to the population and measure the reliability of the results.
5. A(n) <u>statistic</u> is a numerical summary of sample. A(n) <u>parameter</u> is a numerical summary of a population.
6. <u>Variables</u> are the characteristics of the individuals of the population being studied.

7. Contrast the differences between qualitative and quantitative variables.
8. Discuss the differences between discrete and continuous variables.
9. In your own words, define the four levels of measurement of a variable. Give an example of each.
10. Explain what is meant when we say "data vary." How does this variability affect the results of statistical analysis?
11. Explain the process of statistics.
12. The age of a person is commonly considered to be a continuous random variable. Could it be considered a discrete random variable instead? Explain.

Skill Building

In Problems 13–20, determine whether the underlined value is a parameter or a statistic.

13. **State Government** Following the 2006 national midterm election, 18% of the governors of the 50 United States were female. Parameter
NW
Source: National Governors Association
14. **Calculus Exam** The average score for a class of 28 students taking a calculus midterm exam was 72%. Parameter
15. **Illegal Drugs** In a national survey of high school students (grades 9 to 12), 25% of respondents reported that someone had offered, sold, or given them an illegal drug on school property. Statistic
Source: Bureau of Justice Statistics jointly with the U.S. Department of Education, *Indicators of School Crime and Safety, 2006*, December 2006
16. **Alcohol Use** In a national survey on substance abuse, 66.4% of respondents who were full-time college students aged 18 to 22 reported using alcohol within the past month. Statistic
Source: Substance Abuse and Mental Health Services Administration, *Results from the 2006 National Survey on Drug Use and Health: National Findings*, September 2007
17. **Batting Average** Ty Cobb is one of Major League Baseball's greatest hitters of all time, with a career batting average of 0.366. Parameter
Source: baseball-almanac.com
18. **Moonwalkers** Only 12 men have walked on the moon. The average age of these men at the time of their moonwalks was 39 years, 11 months, 15 days. Parameter
Source: Wikipedia.org
19. **Hygiene Habits** A study of 6,076 adults in public rest rooms (in Atlanta, Chicago, New York City, and San Francisco) found that 23% did not wash their hands before exiting. Statistic
Source: American Society for Microbiology and the Soap and Detergent Association, *Press Release: Hygiene Habits Stall: Public Handwashing Down*. September 17, 2007
20. **Public Knowledge** Telephone interviews of 1,502 adults 18 years of age or older, conducted nationwide February 1–13, 2007, found that only 69% could identify the current vice-president. Statistic
Source: The Pew Research Center, *Public Knowledge of Current Affairs Little Changed by News and Information Revolutions: What Americans Know: 1989–2007*, April 15, 2007

In Problems 21–28, classify the variable as qualitative or quantitative.

21. Nation of origin Qualitative
NW
22. Number of siblings Quantitative
23. Grams of carbohydrates in a doughnut Quantitative
24. Number on a football player's jersey Qualitative
25. Number of unpopped kernels in a bag of ACT microwave popcorn Quantitative
26. Assessed value of a house Quantitative
27. Phone number Qualitative
28. Student ID number Qualitative

In Problems 29–36, determine whether the quantitative variable is discrete or continuous. 30. Continuous

29. Runs scored in a season by Albert Pujols Discrete
NW
30. Volume of water lost each day through a leaky faucet
31. Length (in minutes) of a country song Continuous
32. Number of sequoia trees in a randomly selected acre of Yosemite National Park Discrete
33. Temperature on a randomly selected day in Memphis, Tennessee Continuous
34. Internet connection speed in kilobytes per second Continuous
35. Points scored in an NCAA basketball game Discrete
36. Air pressure in pounds per square inch in an automobile tire Continuous

In Problems 37–44, determine the level of measurement of each variable. 41. Ordinal

37. Nation of origin Nominal
NW
38. Movie ratings of one star through five stars Ordinal
39. Volume of water used by a household in a day Ratio
40. Year of birth of college students Interval
41. Highest degree conferred (high school, bachelor's, and so on)
42. Eye color Nominal
43. Assessed value of a house Ratio
44. Time of day measured in military time Interval

In Problems 45–50, a research objective is presented. For each research objective, identify the population and sample in the study.

45. The Gallup Organization contacts 1,028 teenagers who are 13 to 17 years of age and live in the United States and asks whether or not they had been prescribed medications for any mental disorders, such as depression or anxiety.
46. A quality-control manager randomly selects 50 bottles of Coca-Cola that were filled on October 15 to assess the calibration of the filling machine.
47. A farmer wanted to learn about the weight of his soybean crop. He randomly sampled 100 plants and weighed the soybeans on each plant.
48. Every year the U.S. Census Bureau releases the *Current Population Report* based on a survey of 50,000 households. The

goal of this report is to learn the demographic characteristics of all households within the United States, such as income.

49. Folate and Hypertension Researcher John P. Forman and co-workers wanted to determine whether or not higher folate intake is associated with a lower risk of hypertension (high blood pressure) in younger women (27 to 44 years of age). To make this determination, they looked at 7,373 cases of hypertension in younger women and found that younger women who consumed at least 1,000 micrograms per day (μg/d) of total folate (dietary plus supplemental) had a decreased risk of hypertension compared with those who consumed less than 200 μg/d.

Source: John P. Forman, MD; Eric B. Rimm, ScD; Meir J. Stampfer, MD; Gary C. Curhan, MD, ScD, "Folate Intake and the Risk of Incident Hypertension among US Women," *Journal of the American Medical Association* 293:320–329, 2005

50. A large community college has noticed that an increasing number of full-time students are working while attending the school. The administration randomly selects 128 students and asks this question: How many hours per week do you work?

In Problems 51–54, identify the individuals, variables, and data corresponding to the variables. Determine whether each variable is qualitative, continuous, or discrete.

51. Widescreen TVs The following data relate to widescreen high-definition televisions.
NW

Model	Size (in.)	Screen Type	Price ($)
Hitachi #P50X901	50	Plasma	4,000
Mitsubishi #WD-73833	73	Projection	4,300
Sony #KDF-50E3000	50	Projection	1,500
Panasonic #TH-65PZ750U	65	Plasma	9,000
Phillips #60PP9200D37	60	Projection	1,600
Samsung #FP-T5884	58	Plasma	4,200
LG #52LB5D	52	Plasma	3,500

Source: bestbuy.com

52. BMW Cars The following information relates to the 2008 model year product line of BMW automobiles.

Model	Body Style	Weight (lb)	Number of Seats
3 Series	Coupe	3,351	4
5 Series	Sedan	3,505	5
6 Series	Convertible	4,277	4
7 Series	Sedan	4,486	5
X3	Sport utility	4,012	5
Z4 Roadster	Coupe	3,087	2

Source: www.motortrend.com

53. Driver's License Laws The following data represent driver's license laws for various states.

State	Minimum Age for Driver's License (unrestricted)	Mandatory Belt Use Seating Positions	Maximum Allowable Speed Limit (cars on rural interstate), mph, 2007
Alabama	17	Front	70
Colorado	17	Front	75
Indiana	18	All	70
North Carolina	16	All	70
Wisconsin	18	All	65

Source: Governors Highway Safety Association

54. Media Players The following information concerns various digital media players that can be purchased online at circuitcity.com.

Product	Memory Size (GB)	Weight (oz)	Price ($)
Samsung YP-U3	2	0.8	79.99
SanDisk Sansa c200	2	10.4	74.99
Microsoft Zune	4	8.3	149.99
SanDisk Sansa Connect	4	1.7	129.99
Apple iPod nano	4	1.7	149.99
Apple iPod touch	8	4.2	299.99
Archos 605	30	6.7	299.99

Applying the Concepts

55. A Cure for the Common Wart A study conducted by researchers was designed "to determine if application of duct tape is as effective as cryotherapy in the treatment of common warts." The researchers randomly divided 51 patients into two groups. The 26 patients in group 1 had their warts treated by applying duct tape to the wart for 6.5 days and then removing the tape for 12 hours, at which point the cycle was repeated for a maximum of 2 months. The 25 patients in group 2 had their warts treated by cryotherapy (liquid nitrogen applied to the wart for 10 seconds every 2 to 3 weeks) for a maximum of six treatments. Once the treatments were complete, it was determined that 85% of the patients in group 1 and 60% of the patients in group 2 had complete resolution of their warts. The researchers concluded that duct tape is significantly more effective in treating warts than cryotherapy.

Source: Dean R. Focht III, Carole Spicer, Mary P. Fairchok. "The Efficacy of Duct Tape vs. Cryotherapy in the Treatment of Verruca Vulgaris (The Common Wart)," *Archives of Pediatrics and Adolescent Medicine*, 156(10), 2002

(a) What is the research objective?
(b) What is the population being studied? What is the sample?
(c) What are the descriptive statistics?
(d) What are the conclusions of the study?

56. Early Epidurals A study was conducted at Northwestern University in Chicago to determine if pregnant women in first-time labor could receive early low-dose epidurals, an anesthesia to control pain during childbirth, without raising

55. (a) To determine if duct tape is as effective as cryotherapy in the treatment of warts. 55. (b) People with warts; 51 patients with warts
55. (c) 85% in group 1, 60% in group 2 had resolution of their warts 55. (d) Duct tape is more effective in treating warts.

their chances of a Cesarean section. In the study, reported in the *New England Journal of Medicine*, "728 women in first-time labor were divided into two groups. One group received the spinal shot and then got epidurals when the cervix dilated to about 2 centimeters. The other group initially received pain-relieving medicine directly into their bloodstreams, and put off epidurals until 4 centimeters if they could tolerate the pain." In the end, the C-section rate was 18% in the early epidural group and 21% in the delayed group. The researchers concluded that pregnant women in first-time labor can be given a low-dose epidural early without raising their chances of a C-section.

Source: Associated Press, February, 22, 2005

(a) What is the research objective?
(b) What is the population being studied? What is the sample?
(c) What are the descriptive statistics?
(d) What are the conclusions of the study?

57. When Are You Best? NW Gallup News Service conducted a survey of 1,019 American adults aged 18 years or older, August 13–16, 2007. The respondents were asked, "Generally speaking, at what hour of the day or night are you personally at your best?" Of the 1,019 adults surveyed, 55% said they were personally at their best in the morning (5 A.M. to 11:59 A.M.). Gallup reported that 55% of all adult Americans felt they were personally at their best in the morning, with a 3% margin of error with 95% confidence.

(a) What is the research objective?
(b) What is the population?
(c) What is the sample?
(d) List the descriptive statistics.
(e) What can be inferred from this survey?

58. Financial Worries? Gallup News Service conducted a survey of 1,006 American adults aged 18 years or older, September 24–27, 2007. The respondents were asked, "What, if anything, worries you most about your personal financial situation in the long term?" Of the 1,006 adults surveyed, 18% said they were most worried about having enough money for retirement. (Ironically, not having enough money for retirement was not a short-term concern.) Gallup reported that 18% of all adult Americans were most worried about not having enough money for retirement, with a 4% margin of error with 95% confidence.

(a) What is the research objective?
(b) What is the population?
(c) What is the sample?
(d) List the descriptive statistics.
(e) What can be inferred from this survey?

59. What Level of Measurement? It is extremely important for a researcher to clearly define the variables in a study because this helps to determine the type of analysis that can be performed on the data. For example, if a researcher wanted to describe baseball players based on jersey number, what level of measurement would the variable *jersey number* be? Now suppose the researcher felt that certain players who were of lower caliber received higher numbers. Does the level of measurement of the variable change? If so, how?

60. Interpreting the Variable Suppose a fundraiser holds a raffle for which each person that enters the room receives a ticket. The tickets are numbered 1 to N, where N is the number of people at the fundraiser. The first person to arrive receives ticket number 1, the second person receives ticket number 2, and so on. Determine the level of measurement for each of the following interpretations of the variable *ticket number*.

(a) The winning ticket number. Nominal
(b) The winning ticket number was announced as 329. An attendee noted his ticket number was 294 and stated, "I guess I arrived too early." Ordinal
(c) The winning ticket number was announced as 329. An attendee looked around the room and commented, "It doesn't look like there are 329 people in attendance." Ratio

61. Analyze the Article Read the newspaper article and identify (a) the research question the study addresses, (b) the population, (c) the sample, (d) the descriptive statistics, and (e) the inferences of the study.

Study: Educational TV for Toddlers OK

CHICAGO (AP)—*Arthur* and *Barney* are OK for toddler TV-watching, but not *Rugrats* and certainly not *Power Rangers*, reports a new study of early TV-watching and future attention problems.

The research involved children younger than 3, so TV is mostly a no–no anyway, according to the experts. But if TV is allowed, it should be of the educational variety, the researchers said.

Every hour per day that kids under 3 watched violent child-oriented entertainment their risk doubled for attention problems five years later, the study found. Even nonviolent kids' shows like *Rugrats* and *The Flintstones* carried a still substantial risk for attention problems, though slightly lower. On the other hand, educational shows, including *Arthur, Barney* and *Sesame Street* had no association with future attention problems.

Interestingly, the risks only occurred in children younger than age 3, perhaps because that is a particularly crucial period of brain development. Those results echo a different study last month that suggested TV-watching has less impact on older children's behavior than on toddlers.

The American Academy of Pediatrics recommends no television for children younger than 2 and limited TV for older children.

The current study by University of Washington researchers was prepared for release Monday in November's issue of the journal *Pediatrics*.

Previous research and news reports on TV's effects have tended to view television as a single entity, without regard to content. But "the reality is that it's not inherently good or bad. It really depends on what they watch," said Dr. Dimitri Christakis, who co-authored the study with researcher Frederick Zimmerman.

Their study was based on parent questionnaires. They acknowledge it's observational data that only suggests a link and isn't proof that TV habits cause attention problems. Still, they think the connection is plausible.

57. (a) To determine the hour of the day adults feel at their best
57. (b) American adults aged 18 years or older
57. (c) 1,019 American adults surveyed
57. (d) 55% felt they were best in the morning

58. (a) To determine what worries adults most about their personal financial situation in the long term
58. (b) American adults aged 18 years or older
58. (c) 1,006 American adults surveyed
58. (d) 18% were most worried about having enough money for retirement

The researchers called a show violent if it involved fighting, hitting people, threats or other violence that was central to the plot or a main character. Shows listed included *Power Rangers*, *Lion King* and *Scooby Doo*.

These shows, and other kids' shows without violence, also tend to be very fast-paced, which may hamper children's ability to focus attention, Christakis said.

Shows with violence also send a flawed message, namely that "if someone gets bonked on the head with a rolling pin, it just makes a funny sound and someone gets dizzy for a minute and then everything is back to normal," Christakis said.

Dennis Wharton of the National Association of Broadcasters, a trade association for stations and networks including those with entertainment and educational children's TV shows, said he had not had a chance to thoroughly review the research and declined to comment on specifics.

Wharton said his group believes "there are many superb television programs for children, and would acknowledge that it is important for parents to supervise the media consumption habits of young children."

The study involved a nationally representative sample of 967 children whose parents answered government-funded child development questionnaires in 1997 and 2002. Questions involved television viewing habits in 1997. Parents were asked in 2002 about their children's behavior, including inattentiveness, difficulty concentrating and restlessness.

The researchers took into account other factors that might have influenced the results—including cultural differences and parents' education levels—and still found a strong link between the non-educational shows and future attention problems.

Peggy O'Brien, senior vice president for educational programming and services at the Corporation for Public Broadcasting, said violence in ads accompanying shows on commercial TV might contribute to the study results.

She said lots of research about brain development goes into the production of educational TV programming for children, and that the slower pace is intentional.

"We want it to be kind of an extension of play" rather than fantasy, she said.

Source: Copyright @ 2008 The Associated Press.

1.2 OBSERVATIONAL STUDIES VERSUS DESIGNED EXPERIMENTS

Objectives

1. Distinguish between an observational study and an experiment
2. Explain the various types of observational studies

1 Distinguish between an Observational Study and an Experiment

Note to Instructor

If you like, you can wait to cover the material in Sections 1.2 through 1.6 until after Chapter 4. Provided you explain the concept of random sample, you can cover this material after Chapter 8.

Once our research question is developed, we must develop methods for obtaining the data that can be used to answer the questions posed in our research objective. There are two methods for collecting data, *observational studies* and *designed experiments*. To help see the difference between these two methods for obtaining data, read the following two studies.

EXAMPLE 1 Cellular Phones and Brain Tumors

Researchers Joachim Schüz and associates wanted "to investigate cancer risk among Danish cellular telephone users who were followed for up to 21 years." To do so, they kept track of 420,095 people whose first cellular telephone prescription was between 1982 and 1995. In 2002, they recorded the number of people out of the 420,095 people who had a brain tumor and compared the rate of brain tumors in this group to the rate of brain tumors in the general population. They found no significant difference in the rate of brain tumors between the two groups. The researchers concluded "cellular telephone use was not associated with increased risk for brain tumors." (*Source*: Joachim Schüz et al. "Cellular Telephone Use and Cancer Risk: Update of a Nationwide Danish Cohort," *Journal of the National Cancer Institute* 98(23): 1707–1713, 2006)

EXAMPLE 2 Cellular Phones and Brain Tumors

Researchers Joseph L. Roti Roti and associates examined “whether chronic exposure to radio frequency (RF) radiation at two common cell phone signals—835.62 megahertz, a frequency used by analogue cell phones, and 847.74 megahertz, a frequency used by digital cell phones—caused brain tumors in rats.” To do so, the researchers divided 480 rats into three groups. The rats in group 1 were exposed to the analogue cell phone frequency; the rats in group 2 were exposed to the digital frequency; the rats in group 3 served as controls and received no radiation. The exposure was done for 4 hours a day, 5 days a week for 2 years. The rats in all three groups were treated the same, except for the RF exposure.

After 505 days of exposure, the researchers reported the following after analyzing the data. “We found no statistically significant increases in any tumor type, including brain, liver, lung or kidney, compared to the control group.” (*Source*: M. La Regina, E. Moros, W. Pickard, W. Straube, J. L. Roti Roti, “The Effect of Chronic Exposure to 835.62 MHz FMCW or 847.7 MHz CDMA on the Incidence of Spontaneous Tumors in Rats,” Bioelectromagnetic Society Conference, June 25, 2002.)

In both studies, the goal was to determine if radio frequencies from cell phones increase the risk of contracting brain tumors. Whether or not brain cancer was contracted is the *response variable.* The level of cell phone usage is the *explanatory variable.* In research, we wish to determine how varying the amount of an **explanatory variable** affects the value of a **response variable**.

What are the differences between the study in Example 1 and the study in Example 2? Obviously, in Example 1 the study was conducted on humans, whereas the study in Example 2 was conducted on rats. However, there is a bigger difference. In Example 1, no attempt was made to influence the individuals in the study. The researchers simply let the 420,095 people go through their everyday lives and talk on the phone as much or as little as they wished. In other words, no attempt was made to influence the value of the explanatory variable, radio-frequency exposure (cell phone use). Because the researchers simply observed the behavior of the study participants, we call the study in Example 1 an *observational study.*

Definition

An **observational study** measures the value of the response variable without attempting to influence the value of either the response or explanatory variables. That is, in an observational study, the researcher observes the behavior of the individuals in the study without trying to influence the outcome of the study.

Now let’s consider the study in Example 2. In this study, the researchers obtained 480 rats and divided the rats into three groups. Each group was *intentionally* exposed to various levels of radiation. The researchers then compared the number of rats that had brain tumors. Clearly, there was an attempt to influence the individuals in this study because the value of the explanatory variable (exposure to radio frequency) was influenced. Because the researchers controlled the value of the explanatory variable, we call the study in Example 2 a *designed experiment.*

Definition

If a researcher assigns the individuals in a study to a certain group, intentionally changes the value of the explanatory variable, and then records the value of the response variable for each group, the researcher is conducting a **designed experiment**.

Now Work Problem 9

Which Is Better? A Designed Experiment or an Observational Study?

To answer this question, let’s consider another study.

EXAMPLE 3 Do Flu Shots Benefit Seniors?

Researchers wanted to determine the long-term benefits of the influenza vaccine on seniors aged 65 years and older. The researchers looked at records of over 36,000 seniors for 10 years. The seniors were divided into two groups. Group 1 were seniors who chose to get a flu vaccination shot, and group 2 were seniors who chose not to get a flu vaccination shot. After observing the seniors for 10 years, it was determined that seniors who get flu shots are 27% less likely to be hospitalized for pneumonia or influenza and 48% less likely to die from pneumonia or influenza. (*Source*: Kristin L. Nichol, MD, MPH, MBA, James D. Nordin, MD, MPH, David B. Nelson, PhD, John P. Mullooly, PhD, Eelko Hak, PhD. "Effectiveness of Influenza Vaccine in the Community-Dwelling Elderly," *New England Journal of Medicine* 357:1373–1381, 2007)

Wow! The results of this study sound great! All seniors should go out and get a flu shot. Right? Well, hold on a second. The authors of the study admitted that there may be some flaws in their results. They were concerned with *confounding*. That is, the authors were concerned that there might be a different explanation for lower hospitalization and death rates than the flu shot. Could it be that seniors that get flu shots are more health conscious in the first place? Could it be that seniors who get flu shots are able to get around more easily, so they can get to the clinic to get the flu shot? Does race, income, or gender play a role in whether one might contract (and possibly die from) influenza?

Definition

Confounding in a study occurs when the effects of two or more explanatory variables are not separated. Therefore, any relation that may exist between an explanatory variable and the response variable may be due to some other variable or variables not accounted for in the study.

Confounding is potentially a major problem with observational studies. Often, the cause of confounding is a *lurking variable.*

Definition

A **lurking variable** is an explanatory variable that was not considered in a study, but that affects the value of the response variable in the study. In addition, lurking variables are typically related to explanatory variables considered in the study.

In the influenza study, possible lurking variables might be age, health status, or mobility of the senior. How can we manage the effect of lurking variables? One possibility is to look at the individuals in the study to determine if they differ in any significant way. For example, it turns out in the influenza study that the seniors who elected to get a flu shot were actually *less* healthy than those who did not. The researchers also accounted for race and income. Another variable the authors identified as a potential lurking variable was *functional status*, meaning the ability of the seniors to conduct day-to-day activities on their own. The authors were able to adjust their results for this variable as well.

Even after accounting for all the potential lurking variables in the study, the authors were still careful to conclude that getting an influenza shot is *associated* with a lower risk of being hospitalized or dying from influenza. The reason the authors used the term *associated* instead of saying that influenza shots result in (or *cause*) a lower risk of death due to influenza is because the study was observational.

Observational studies do not allow a researcher to claim causation, only association.

Designed experiments, on the other hand, are used whenever control of certain variables is possible and desirable. This type of research allows the researcher to identify certain cause and effect relationships among the variables in the study.

So why ever conduct a study through an observational experiment? Often, it is unethical to conduct an experiment. Consider the link between smoking and lung cancer. Would you want to participate in a designed experiment to determine if smoking causes lung cancer in humans? To do so, a researcher would divide a group of volunteers into two groups. Group 1 would be told to smoke a pack of cigarettes every day for the next 10 years, while group 2 would not. In addition, the researcher would control eating habits, sleeping habits, and exercise so that the only difference between the two groups was smoking. After 10 years the researcher would compare the incidence rate of lung cancer (the response variable) in the smoking group to the nonsmoking group. If the two cancer rates differ significantly, we could say that smoking causes cancer. By approaching the study in this way, we are able to control many of the factors that might affect the incidence rate of lung cancer that were beyond our control in the observational study.

Other reasons exist for conducting observational studies over designed experiments. Kjell Benson and Arthur Hartz wrote an article in the *New England Journal of Medicine* in support of observational studies by stating, "observational studies have several advantages over designed experiments, including lower cost, greater timeliness, and a broader range of patients." (*Source*: Kjell Benson, BA, and Arthur J. Hartz, MD, PhD. "A Comparison of Observational Studies and Randomized, Controlled Trials," *New England Journal of Medicine* 342:1878–1886, 2000)

For the remainder of this section, and in Sections 1.3 through 1.5, we will look at obtaining data through various types of observational studies. We look at designed experiments in Section 1.6.

2 Explain the Various Types of Observational Studies

There are three major categories of observational studies: (1) cross-sectional studies, (2) case-control studies, and (3) cohort studies.

Cross-sectional Studies These are observational studies that collect information about individuals at a specific point in time or over a very short period of time.

For example, a researcher might want to assess the risk associated with smoking by looking at a group of people, determining how many are smokers and comparing the incidence rate of lung cancer of the smokers to the nonsmokers.

A clear advantage of cross-sectional studies is that they are cheap and quick to do. However, cross-sectional studies have limitations. For our lung cancer study, it could be that individuals develop cancer after the data are collected, so our study will not give the full picture.

Case-control Studies These studies are **retrospective**, meaning that they require individuals to look back in time or require the researcher to look at existing records. In case-control studies, individuals that have a certain characteristic are matched with those that do not.

For example, we might match individuals that have lung cancer with those that do not. When we say "match" individuals, we mean that we would like the individuals in the study to be as similar (homogeneous) as possible in terms of demographics and other variables that may affect the response variable. Once homogeneous groups are established, we would ask the individuals in each group how much they smoked. The incidence rate of lung cancer between the two groups would then be compared.

Certainly, a disadvantage to this type of study is that it requires individuals to recall information from the past. Plus, it requires the individuals to be truthful in their responses. An advantage of case-control studies is that they are relatively inexpensive to conduct and can be done relatively quickly.

Cohort Studies A cohort study first identifies a group of individuals to participate in the study (the cohort). The cohort is then observed over a period of time (sometimes a long period of time). Over this time period, characteristics about the individuals are recorded and some individuals in the study will be exposed to certain factors (not intentionally) and others will not. At the end of the study the value of the response variable is recorded for the individuals.

The observational study in Example 1 is a cohort study that took over 21 years to complete! The individuals were divided into groups depending on their cell phone usage. A cohort study was done to further advance the link between lung cancer and smoking. Typically, cohort studies require many individuals to participate over long periods of time. Because the data are collected over time, cohort studies are **prospective**. Another problem with cohort studies is that individuals tend to drop out due to the long time frame. This could lead to misleading results. Cohort studies definitely are the most powerful of the observational studies.

One of the largest cohort studies is the Framingham Heart Study. In this study, more than 10,000 individuals have been monitored since 1948. The study continues to this day, with the grandchildren of the original participants taking part in the study. This cohort study is responsible for many of the breakthroughs in understanding heart disease. The cost of this study is in excess of $10 million.

Some Concluding Remarks about Observational Studies versus Designed Experiments

Is a designed experiment superior to an observational study? Not necessarily. Plus, observational studies play a role in the research process. For example, because cross-sectional and case-control observational studies are relatively inexpensive, they provide an opportunity to explore possible associations prior to undertaking large cohort studies or designing experiments.

Also, it is not always possible to conduct an experiment. For example, we could not conduct an experiment to investigate the perceived link between high tension wires and leukemia. Do you see why?

Now Work Problem 19

Existing Sources of Data and Census Data

Have you ever heard this saying? *There is no point in reinventing the wheel.* Well, there is no point in spending energy obtaining data that already exist either. If a researcher wishes to conduct a study and a data set exists that can be used to answer the researcher's questions, then it would be silly to collect the data from scratch. For example, various federal agencies regularly collect data that are available to the public. Some of these agencies include the Centers for Disease Control and Prevention (www.cdc.gov), the Internal Revenue Service (www.irs.gov), and the Department of Justice (http://fjsrc.urban.org/index.cfm). In fact, a great website that lists virtually all the sources of federal data is www.fedstats.gov. Another great source of data is the General Social Survey (GSS) administered by the University of Chicago. This survey regularly asks "demographic and attitudinal questions" of individuals around the country. The website is www.gss.norc.org.

Another source of data is a *census*.

Definition

A **census** is a list of all individuals in a population along with certain characteristics of each individual.

The United States attempts to conduct a census every 10 years to learn the demographic makeup of the United States. Everyone whose usual residence is within the borders of the United States must fill out a questionnaire packet. There are two different census forms: a short form and a long form. The short form goes to every household in the United States and includes questions on name, gender, age, relationship of individuals living in the household, Hispanic origin, race, and housing tenure (whether the home is owned or rented). About 83% of all households received the short form in 2000, with the remaining households receiving the long form. The cost of obtaining the census in 2000 was approximately $6 billion; about 860,000 temporary workers were hired to assist in collecting the data.

Why is the U.S. Census so important? The results of the census are used to determine the number of representatives in the House of Representatives in each state, congressional districts, distribution of funds for government programs (such as Medicaid), and planning for the construction of schools and roads. The first census of the United States was obtained in 1790 under the direction of Thomas Jefferson. It is a constitutional mandate that a census be conducted every 10 years (Article 1, Section 2, of the U.S. Constitution).

Is the United States successful in obtaining a census? Not entirely. Inevitably, certain individuals in the United States go uncounted. Why? There are a number of reasons, but a few of the common reasons include illiteracy, language issues, and homelessness. Given what is at stake politically based on the results of the census, politicians often debate on how to count these individuals. In fact, statisticians have offered solutions to the counting problem. The interested reader can go to www.census.gov and in the search box type *count homeless*. You will find many articles related to the Census Bureau's attempt to count the homeless. The bottom line is that even census data can have flaws.

1.2 ASSESS YOUR UNDERSTANDING

Concepts and Vocabulary

1. In your own words, define explanatory variable and response variable.

2. What is an observational study? What is a designed experiment? Which allows the researcher to claim causation between an explanatory variable and a response variable?

3. Explain what is meant by confounding. What is a lurking variable?

4. Given a choice, would you conduct a study using an observational study or a designed experiment? Why?

5. What is a cross-sectional study? What is a case-control study? Which is the superior observational study? Why?

6. The data used in the influenza study presented in Example 3 were obtained from a cohort study. What does this mean? Why is a cohort study superior to a case-control study?

7. Explain why it would be unlikely to use a designed experiment to answer the research question posed in Example 3.

8. What does it mean when an observational study is retrospective? What does it mean when an observational study is prospective?

Skill Building

In Problems 9–16, determine whether the study depicts an observational study or an experiment.

9. NW Researchers wanted to know if there is a link between proximity to high-tension wires and the rate of leukemia in children. To conduct the study, researchers compared the incidence rate of leukemia for children who lived within $\frac{1}{2}$ mile of high-tension wires to the incidence rate of leukemia for children who did not live within $\frac{1}{2}$ mile of high-tension wires. Observational study

10. Rats with cancer are divided into two groups. One group receives 5 milligrams (mg) of a medication that is thought to fight cancer, and the other receives 10 mg. After 2 years, the spread of the cancer is measured. Experiment

11. Seventh-grade students are randomly divided into two groups. One group is taught math using traditional techniques; the other is taught math using a reform method. After 1 year, each group is given an achievement test to compare proficiency. Experiment

12. A poll is conducted in which 500 people are asked whom they plan to vote for in the upcoming election. Observational study

13. A survey is conducted asking 400 people, "Do you prefer Coke or Pepsi?" Observational study

14. While shopping, 200 people are asked to perform a taste test in which they drink from two randomly placed, unmarked cups. They are then asked which drink they prefer. Experiment

15. Sixty patients with carpal tunnel syndrome are randomly divided into two groups. One group is treated weekly with both acupuncture and an exercise regimen. The other is treated weekly with the exact same exercise regimen, but no acupuncture. After 1 year, both groups are questioned about their level of pain due to carpal tunnel syndrome. Experiment

16. Conservation agents netted 250 large-mouth bass in a lake and determined how many were carrying parasites. Observational study

Applying the Concepts

17. Daily Coffee Consumption Researchers wanted to determine if there was an association between daily coffee consumption and the occurrence of skin cancer. The researchers looked at 93,676 women enrolled in the Women's Health Initiative Observational Study and asked them to report their coffee-drinking habits. The researchers also determined which of the women had nonmelanoma skin cancer. After their analysis, the researchers concluded that consumption of six or more cups of caffeinated coffee per day was associated with a reduction in nonmelanoma skin cancer.

Source: *European Journal of Cancer Prevention*, 16(5): 446–452, October 2007

17. (a) Cross-sectional 17. (b) Whether the woman has nonmelanoma skin cancer or not; consumption of caffeinated coffee

(a) What type of observational study was this? Explain.
(b) What is the response variable in the study? What is the explanatory variable?
(c) In their report, the researchers stated that "After adjusting for various demographic and lifestyle variables, daily consumption of six or more cups was associated with a 30% reduced prevalence of nonmelanoma skin cancer." Why was it important to adjust for these variables?

18. Obesity and Artery Calcification Scientists were interested in determining if abdominal obesity is related to coronary artery calcification (CAC). The scientists studied 2,951 participants in the Coronary Artery Risk Development in Young Adults Study to investigate a possible link. Waist and hip girths were measured in 1985–1986, 1995–1996 (year 10), and in 2000–2001 (waist girth only). CAC measurements were taken in 2001–2002. The results of the study indicated that abdominal obesity measured by waist girth is associated with early atherosclerosclerosis as measured by the presence of CAC in participants.

Source: *American Journal of Clinical Nutrition*, 86(1): 48–54, 2007 18. (a) Cohort study

(a) What type of observational study was this? Explain.
(b) What is the response variable in the study? What is the explanatory variable? Presence of CAC; waist girth

19. Television in the Bedroom Researchers Christelle Delmas and associates wanted to determine if having a television (TV) in the bedroom is associated with obesity. The researchers administered a questionnaire to 379 twelve-year-old French adolescents. After analyzing the results, the researchers determined that the body mass index of the adolescents who had a TV in their bedroom was significantly higher than that of the adolescents who did not have a TV in their bedroom.

NW

Source: Christelle Delmas, Carine Platat, Brigette Schweitzer, Aline Wagner, Mohamed Oujaa, and Chantal Simon. "Association Between Television in Bedroom and Adiposity Throughout Adolescence," *Obesity*, 15:2495–2503, 2007

(a) Why is this an observational study? What type of observational study is this? Cross-sectional
(b) What is the response variable in the study? What is the explanatory variable?
(c) Can you think of any lurking variables that may affect the results of the study?
(d) In the report, the researchers stated, "These results remain significant after adjustment for socioeconomic status." What does this mean?
(e) Does a television in the bedroom cause a higher body mass index? Explain. No

20. Get Married, Gain Weight Researcher Penny Gordon-Larson and her associate wanted to determine whether young couples who marry or cohabitate are more likely to gain weight than those who stay single. The researchers followed 8,000 men and women from 1995 through 2002 as they matured from the teens to young adults. When the study began, none of the participants was married or living with a romantic partner. By 2002, 14% of the participants were married and 16% were living with a romantic partner. At the end of the study, married or cohabiting women gained, on average, 9 pounds more than single women, and married or cohabiting men gained, on average, 6 pounds more than single men.

(a) Why is this an observational study? What type of observational study is this? Cohort study
(b) What is the response variable in the study? What is the explanatory variable?
(c) Identify some potential lurking variables in this study.
(d) Does getting married or cohabiting cause one to gain weight? Explain.

21. Analyze the Article Write a summary of the following opinion. The opinion is posted at abcnews.com. Include the type of study conducted, possible lurking variables, and conclusions. What is the message of the author of the article?

Power Lines and Cancer—To Move or Not to Move

New Research May Cause More Fear Than Warranted, One Physician Explains

OPINION by JOSEPH MOORE, M.D.

May 30, 2007—

A recent study out of Switzerland indicates there might be an increased risk of certain blood cancers in people with prolonged exposure to electromagnetic fields, like those generated from high-voltage power lines.

If you live in a house near one of these high-voltage power lines, a study like this one might make you wonder whether you should move.

But based on what we know now, I don't think that's necessary. We can never say there is no risk, but we can say that the risk appears to be extremely small.

"Scare Science"

The results of studies like this add a bit more to our knowledge of potential harmful environmental exposures, but they should also be seen in conjunction with the results of hundreds of studies that have gone before. It cannot be seen as a definitive call to action in and of itself.

The current study followed more than 20,000 Swiss railway workers over a period of 30 years. True, that represents a lot of people over a long period of time.

However, the problem with many epidemiological studies, like this one, is that it is difficult to have an absolute control group of people to compare results with. The researchers compared the incidence of different cancers of workers with a high amount of electromagnetic field exposure to those workers with lower exposures.

These studies aren't like those that have identified definitive links between an exposure and a disease—like those involving smoking and lung cancer. In those studies, we can actually measure the damage done to lung tissue as a direct result of smoking. But usually it's very difficult for the conclusions of an epidemiological study to rise to the level of controlled studies in determining public policy.

Remember the recent scare about coffee and increased risk of pancreatic cancer? Or the always-simmering issue of cell phone use and brain tumors?

As far as I can tell, none of us have turned in our cell phones. In our own minds, we've decided that any links to cell phone use and brain cancer have not been proven definitively. While we can't say that there is absolutely no risk in using cell phones, individuals have determined on

19. (b) Body mass index; whether a TV is in the bedroom or not 20. (b) Weight gain; whether the individual is married/cohabitating or not
20. (c) Having children; eating out more often

their own that the potential risks appear to be quite small and are outweighed by the benefits.

Findings Shouldn't Lead to Fear

As a society, we should continue to investigate these and other related exposures to try to prove one way or another whether they are disease-causing. If we don't continue to study, we won't find out. It's that simple.

When findings like these come out, and I'm sure there will be more in the future, I would advise people not to lose their heads. Remain calm. You should take the results as we scientists do—as intriguing pieces of data about a problem we will eventually learn more about, either positively or negatively, in the future. It should not necessarily alter what we do right now.

What we can do is take actions that we know will reduce our chances of developing cancer.

Stop smoking and avoid passive smoke. It is the leading cause of cancer that individuals have control over.

Whenever you go outside, put on sunscreen or cover up.

Eat a healthy diet and stay physically active.

Make sure you get tested or screened. Procedures like colonoscopies, mammograms, pap smears and prostate exams can catch the early signs of cancer, when the chances of successfully treating them are the best.

Taking the actions above will go much farther in reducing your risks for cancer than moving away from power lines or throwing away your cell phone.

Dr. Joseph Moore is a medical oncologist at Duke University Comprehensive Cancer Center.

Source: Reprinted with the permission of the author.

22. Reread the article in Problem 61 from Section 1.1. What type of observational study does this appear to be? Name some lurking variables that the researchers accounted for.

23. Putting It Together: Passive Smoke The following abstract appears in *The New England Journal of Medicine:*

BACKGROUND. The relation between passive smoking and lung cancer is of great public health importance. Some previous studies have suggested that exposure to environmental tobacco smoke in the household can cause lung cancer, but others have found no effect. Smoking by the spouse has been the most commonly used measure of this exposure.

METHODS. In order to determine whether lung cancer is associated with exposure to tobacco smoke within the household, we conducted a case-control study of 191 patients with lung cancer who had never smoked and an equal number of persons without lung cancer who had never smoked. Lifetime residential histories including information on exposure to environmental tobacco smoke were compiled and analyzed. Exposure was measured in terms of "smoker-years," determined by multiplying the number of years in each residence by the number of smokers in the household.

RESULTS. Household exposure to 25 or more smoker-years during childhood and adolescence doubled the risk of lung cancer. Approximately 15 percent of the control subjects who had never smoked reported this level of exposure. Household exposure of less than 25 smoker-years during childhood and adolescence did not increase the risk of lung cancer. Exposure to a spouse's smoking, which constituted less than one third of total household exposure on average, was not associated with an increase in risk.

CONCLUSIONS. The possibility of recall bias and other methodologic problems may influence the results of case-control studies of environmental tobacco smoke. Nonetheless, our findings regarding exposure during early life suggest that approximately 17 percent of lung cancers among nonsmokers can be attributed to high levels of exposure to cigarette smoke during childhood and adolescence.

(a) What is the research objective?
(b) What makes this study a case-control study? Why is this a retrospective study?
(c) What is the response variable in the study? Is it qualitative or quantitative?
(d) What is the explanatory variable in the study? Is it qualitative or quantitative?
(e) Can you identify any lurking variables that may have affected this study?
(f) What is the conclusion of the study? Does exposure to smoke in the household cause lung cancer?
(g) Would it be possible to design an experiment to answer the research question in part (a)? Explain.

23. (c) Whether the individual has lung cancer; qualitative
23. (d) Number of smoker-years; quantitative
23. (e) Income; amount of exposure outside the home

1.3 SIMPLE RANDOM SAMPLING

Objective 1 Obtain a simple random sample

Sampling

Besides the observational studies that we looked at in Section 1.2, observational studies can also be conducted by administering a survey. Whenever administering a survey, the researcher must first identify the population that is to be targeted. For example, the Gallup Organization regularly surveys Americans about various pop-culture and political issues. Often, the population of interest in these surveys is adult Americans aged 18 years or older. Of course, it is unreasonable to expect the Gallup Organization to survey *all* adult Americans (there are over 200 million), so instead the Gallup Organization will typically survey a *random sample* of about 1,000 adult Americans.

Definition

Random sampling is the process of using chance to select individuals from a population to be included in the sample.

Note to Instructor

Spend some time in class talking about the following. Suppose a grocery store manager wants to get a feel for how customers view the store. Why would a sample of the first 100 shoppers leaving a grocery store not be representative of the population? Why is this sampling method not random sampling?

What allows a researcher to be confident the results of a survey accurately reflect the feelings of an entire population? For the results of a survey to be reliable, the characteristics of the individuals in the sample must be representative of the characteristics of the individuals in the population. How can this be accomplished? The key to obtaining a sample representative of a population is to let *chance* or *randomness* play a role in dictating which individuals are in the sample, rather than convenience. If convenience is used to obtain a sample, the results of the survey are meaningless.

For example, suppose that Gallup wants to know the proportion of adult Americans who consider themselves to be baseball fans. If Gallup obtained a sample by standing outside of Fenway Park (home of the Boston Red Sox professional baseball team), the results of the survey are not likely to be reliable. Why? Clearly, the individuals in the sample do not accurately reflect the makeup of the entire population. As another example, suppose you wanted to learn the proportion of students on your campus who work. It might be convenient to survey the students in your statistics class, but do the students in your class represent the overall student body? Is the proportion of freshmen, sophomores, juniors, and seniors in your class close to the proportion of freshmen, sophomores, juniors, and seniors on campus? Is the proportion of males and females in your class close to the proportion of males and females on campus? Probably not. For this reason, the convenient sample is not representative of the population, which means the results of your survey are misleading.

We will discuss four basic sampling techniques: *simple random sampling*, *stratified sampling*, *systematic sampling*, and *cluster sampling*. These sampling methods are designed so that any selection biases introduced (knowingly or unknowingly) by the surveyor during the selection process are eliminated. In other words, the surveyor does not have a choice as to which individuals are in the study. We will discuss simple random sampling now and the remaining three types of sampling in the next section.

1 Obtain a Simple Random Sample

The most basic sample survey design is *simple random sampling*.

Definition

A sample of size n from a population of size N is obtained through **simple random sampling** if every possible sample of size n has an equally likely chance of occurring. The sample is then called a **simple random sample**.

In Other Words

Simple random sampling is like selecting names from a hat.

The sample is always a subset of the population, meaning that the number of individuals in the sample is less than the number of individuals in the population.

Note to Instructor

The purpose of this activity is to allow students to visualize simple random sampling. It also helps to foreshadow the idea of a sampling distribution by showing that different samples lead to different individuals in the sample. It should take about 15 to 20 minutes.

IN CLASS ACTIVITY

Simple Random Sampling

This activity illustrates the idea of simple random sampling.

(a) Choose 5 students in the class to represent a population. Number the students 1 through 5.

(b) Form all possible samples of size $n = 2$ from the population of size $N = 5$. How many different simple random samples are possible?

(c) Write the numbers 1 through 5 on five pieces of paper and then place the paper in a hat. Select two of the numbers. The two individuals corresponding to these numbers are in the sample.

(d) Put the two numbers back in the hat. Select two of the numbers. The two individuals corresponding to these numbers are in the sample. Are the individuals in the second sample the same as the individuals in the first sample?

EXAMPLE 1 Illustrating Simple Random Sampling

Problem: Sophia has four tickets to a concert. Six of her friends, Yolanda, Michael, Kevin, Marissa, Annie, and Katie, have all expressed an interest in going to the concert. Sophia decides to randomly select three of her six friends to attend the concert.

(a) List all possible samples of size $n = 3$ (without replacement) from the population of size $N = 6$.

(b) Comment on the likelihood of the sample containing Michael, Kevin, and Marissa.

Approach: We list all possible combinations of three people chosen from the six. Remember, in simple random sampling, each sample of size 3 is equally likely to occur.

Solution

(a) The possible samples of size 3 are listed in Table 2.

Table 2

Yolanda, Michael, Kevin	Yolanda, Michael, Marissa	Yolanda, Michael, Annie	Yolanda, Michael, Katie
Yolanda, Kevin, Marissa	Yolanda, Kevin, Annie	Yolanda, Kevin, Katie	Yolanda, Marissa, Annie
Yolanda, Marissa, Katie	Yolanda, Annie, Katie	Michael, Kevin, Marissa	Michael, Kevin, Annie
Michael, Kevin, Katie	Michael, Marissa, Annie	Michael, Marissa, Katie	Michael, Annie, Katie
Kevin, Marissa, Annie	Kevin, Marissa, Katie	Kevin, Annie, Katie	Marissa, Annie, Katie

From Table 2, we see that there are 20 possible samples of size 3 from the population of size 6. We use the term *sample* to mean the individuals in the sample.

(b) There is 1 sample that contains Michael, Kevin, and Marissa and 20 possible samples, so there is a 1 in 20 chance that the simple random sample will contain Michael, Kevin, and Marissa. In fact, all the samples of size 3 have a 1 in 20 chance of occurring.

Now Work Problem 7

Obtaining a Simple Random Sample

The results of Example 1 leave one question unanswered: How do we select the individuals in a simple random sample? To obtain a simple random sample from a population, we could write the names of the individuals in the population on different sheets of paper and then select names from a hat.

Often, however, the size of the population is so large that performing simple random sampling in this fashion is not practical. Typically, random numbers are used by assigning each individual in the population a unique number between 1 and N, where N is the size of the population. Then n random numbers from this list are selected, where n represents the size of the sample. Because we must number the individuals in the population, we must have a list of all the individuals within the population, called a **frame**.

In Other Words

A frame lists all the individuals in a population. For example, a list of all registered voters in a particular precinct might be a frame.

EXAMPLE 2 Obtaining a Simple Random Sample

Problem: Senese and Associates has increased their accounting business. To make sure their clients are still satisfied with the services they are receiving, Senese and Associates decides to send a survey out to a simple random sample of 5 of its 30 clients.

Approach

Step 1: A list of the 30 clients must be obtained (the frame). Each client is then assigned a unique number from 01 to 30.

Step 2: Five unique numbers will be randomly selected. The clients corresponding to the numbers are sent a survey. This process is called *sampling without replacement*. When we **sample without replacement**, once an individual is selected, he or she is removed from the population and cannot be chosen again. Contrast this with **sampling with replacement**, which means the selected individual is placed back into

Note to Instructor

In-class discussion: What is meant by sampling "with replacement"? What is meant by sampling "without replacement"?

the population and so could be chosen a second time. We use sampling without replacement so that we don't select the same client twice.

Solution

Step 1: Table 3 shows the list of clients. We arrange the clients in alphabetic order (although this is not necessary). Because there are 30 clients, we number the clients from 01 to 30.

Table 3

01. ABC Electric	11. Fox Studios	21. R&Q Realty
02. Brassil Construction	12. Haynes Hauling	22. Ritter Engineering
03. Bridal Zone	13. House of Hair	23. Simplex Forms
04. Casey's Glass House	14. John's Bakery	24. Spruce Landscaping
05. Chicago Locksmith	15. Logistics Management, Inc.	25. Thors, Robert DDS
06. DeSoto Painting	16. Lucky Larry's Bistro	26. Travel Zone
07. Dino Jump	17. Moe's Exterminating	27. Ultimate Electric
08. Euro Car Care	18. Nick's Tavern	28. Venetian Gardens Restaurant
09. Farrell's Antiques	19. Orion Bowling	29. Walker Insurance
10. First Fifth Bank	20. Precise Plumbing	30. Worldwide Wireless

Step 2: A table of random numbers can be used to select the individuals to be in the sample. See Table 4.* We select a starting place in the table of random numbers. This

Column 4

Table 4

Row Number	Column Number 01–05	06–10	11–15	16–20	21–25	26–30	31–35	36–40	41–45	46–50
01	89392	**23**212	74483	36590	25956	36544	68518	40805	09980	00467
02	61458	17639	96252	95649	73727	33912	72896	66218	52341	97141
03	11452	74197	81962	48433	90360	26480	73231	37740	26628	44690
04	27575	04429	31308	02241	01698	19191	18948	78871	36030	23980
05	36829	59109	88976	46845	28329	47460	88944	08264	00843	84592
06	81902	93458	42161	26099	09419	89073	82849	09160	61845	40906
07	59761	55212	33360	68751	86737	79743	85262	31887	37879	17525
08	46827	25906	64708	20307	78423	15910	86548	08763	47050	18513
09	24040	66449	32353	83668	13874	86741	81312	54185	78824	00718
10	98144	96372	50277	15571	82261	66628	31457	00377	63423	55141
11	14228	17930	30118	00438	49666	65189	62869	31304	17117	71489
12	55366	51057	90065	14791	62426	02957	85518	28822	30588	32798
Row 13 — 13	961**01**	30646	35526	90389	73634	79304	96635	06626	94683	16696
14	381**52**	55474	30153	26525	83647	31988	82182	98377	33802	80471
15	850**07**	18416	24661	95581	45868	15662	28906	36392	07617	50248
16	855**44**	15890	80011	18160	33468	84106	40603	01315	74664	20553
17	104**46**	20699	98370	17684	16932	80449	92654	02084	19985	59321
18	672**37**	45509	17638	65115	29757	80705	82686	48565	72612	61760
19	230**26**	89817	05403	82209	30573	47501	00135	33955	50250	72592
20	674**11**	58542	18678	46491	13219	84084	27783	34508	55158	78742

We skip 52 because it is larger than 30.

*Each digit is in its own column. The digits are displayed in groups of five for ease of reading. The digits in row 1 are 893922321274483, and so on. The first digit, 8, is in column 1; the second digit, 9, is in column 2; the ninth digit, 1, is in column 9.

can be done by closing the eyes and placing a finger on the table. This may sound haphazard, but it accomplishes the goal of being random. Suppose we start in column 4, row 13. Because our data have two digits, we select two-digit numbers from the table using columns 4 and 5. We only select numbers greater than or equal to 01 and less than or equal to 30. Anytime we encounter 00, a number greater than 30, or a number already selected, we skip it and continue to the next number.

The first number in the list is 01, so the client corresponding to 01 will receive a survey. Moving down the list, the next number is 52. Because 52 is greater than 30, we skip it. Continuing down the list, the following numbers are selected from the list:

01, 07, 26, 11, 23

The clients corresponding to these numbers are

ABC Electric, Dino Jump, Travel Zone, Fox Studios, Simplex Forms

Each random number used to select the individuals in the sample is set in boldface type in Table 4 to help you to understand where the numbers come from.

EXAMPLE 3 Obtaining a Simple Random Sample Using Technology

Note to Instructor

Many types of statistical software packages and calculators can be used in statistics. Because of the number of statistical packages available, it is impossible for a text to present each of them. Therefore, the three most popular (MINITAB, Excel, and TI83/TI84), as learned from our surveys, are presented in this text. The technology used will vary from example to example. However, the purpose of the examples that use technology is to illustrate how it can obtain the results we need to solve a problem. Once we have the results, work still must be done to determine its meaning. So, just because the software or calculator used in a technology example is not the one used by you, do not feel obligated to skip it.

Problem: Find a simple random sample of five clients for the problem presented in Example 2.

Approach: The approach is similar to that given in Example 2.

Step 1: A list of the 30 clients must be obtained (the frame). The clients are then assigned a number from 01 to 30.

Step 2: Five numbers are randomly selected using a random number generator. The clients corresponding to the numbers are given a survey. We sample without replacement so that we don't select the same client twice. To use a random-number generator using technology, we must first set the *seed*. The **seed** in a random-number generator provides an initial point for the generator to start creating random numbers. It is just like selecting the initial point in the table of random numbers. The seed can be any nonzero number. Statistical software such as MINITAB or Excel can be used to generate random numbers, but we will use a TI-84 Plus graphing calculator. The steps for obtaining random numbers using MINITAB, Excel, and the TI-83/84 graphing calculator can be found in the Technology Step-by-Step on page 29.

Solution

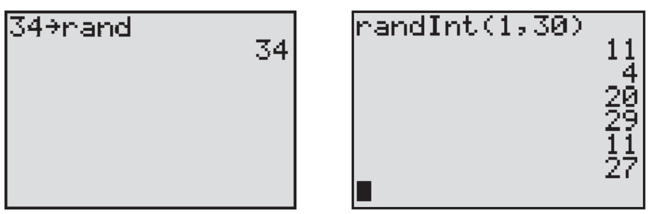
Step 1: Table 3 on page 25 shows the list of clients and numbers corresponding to the clients.

Step 2: See Figure 3(a) for an illustration of setting the seed using a TI-84 Plus graphing calculator, where the seed is set at 34. We are now ready to obtain the list of random numbers. Figure 3(b) shows the results obtained from a TI-84 Plus graphing calculator.

Using Technology

If you are using a different statistical package or type of calculator, the random numbers generated will likely be different. This does not mean you are wrong. There is no such thing as a wrong random sample as long as the correct procedures are followed.

Figure 3



(a) (b)

The following numbers are generated by the calculator:

$$11, 4, 20, 29, 11, 27$$

We ignore the second 11 because we are sampling without replacement. The clients corresponding to these numbers are the clients to be surveyed: Fox Studios, Casey's Glass House, Precise Plumbing, Walker Insurance, and Ultimate Electric.

Now Work Problem 11

CAUTION Random-number generators are not truly random, because they are programs, and programs do not act "randomly." The seed dictates the random numbers that are generated.

There is a very important consequence when comparing the by hand and technology solutions from Examples 2 and 3. Because both samples were obtained randomly, they resulted in different individuals in the sample! For this reason, each sample will likely result in different descriptive statistics. Any inference based on each sample *may* result in different conclusions regarding the population. This is the nature of statistics. Inferences based on samples will vary because the individuals in different samples vary.

1.3 ASSESS YOUR UNDERSTANDING

Concepts and Vocabulary

1. Explain why a frame is necessary to obtain a simple random sample.
2. Discuss why sampling is used in statistics.
3. What does it mean when sampling is done without replacement?
4. What is random sampling? Why is it used and how does it compare with convenience sampling?

Skill Building

5. **Literature** As part of a college literature course, students must select three classic works of literature from the provided list and complete critical book reviews for each selected work. Obtain a simple random sample of size 3 from this list. Write a short description of the process you used to generate your sample.

Pride and Prejudice	*The Sun Also Rises*	*The Jungle*
As I Lay Dying	*A Tale of Two Cities*	*Huckleberry Finn*
Death of a Salesman	*Scarlet Letter*	*Crime and Punishment*

6. **Team Captains** A coach must select two players to serve as captains at the beginning of a soccer match. He has 10 players on his team and, to be fair, wants to randomly select 2 players to be the captains. Obtain a simple random sample of size 2 from the following list. Write a short description of the process you used to generate your sample.

Mady	Breanne	Jory
Evin	Tori	Payton
Emily	Claire	Jordyn
Caty		

7. **Course Selection** A student entering a doctoral program in NW educational psychology is required to select two courses from the list of courses provided as part of his or her program.

EPR 616, Research in Child Development
EPR 630, Educational Research Planning and Interpretation
EPR 631, Nonparametric Statistics
EPR 632, Methods of Multivariate Analysis
EPR 645, Theory of Measurement
EPR 649, Fieldwork Methods in Educational Research
EPR 650, Interpretive Methods in Educational Research

(a) List all possible two-course selections.
(b) Comment on the likelihood that the pair of courses EPR 630 and EPR 645 will be selected.

8. **Merit Badge Requirements** To complete the Citizenship in the World merit badge, one must select TWO of the following organizations and describe their role in the world.
Source: Boy Scouts of America
   1. The United Nations
   2. The World Court
   3. World Organization of the Scout Movement
   4. The World Health Organization
   5. Amnesty International
   6. The International Committee of the Red Cross
   7. CARE

(a) List all possible pairs of organizations.
(b) Comment on the likelihood that the pair The United Nations and Amnesty International will be selected.

Applying the Concepts

9. **Sampling the Faculty** A small community college employs 87 full-time faculty members. To gain the faculty's opinions about an upcoming building project, the college president wishes to obtain a simple random sample that will consist of 9 faculty members. He numbers the faculty from 1 to 87.
(a) Using Table I from Appendix A, the president closes his eyes and drops his ink pen on the table. It points to the digit in row 5, column 22. Using this position as the starting point and proceeding downward, determine the numbers for the 9 faculty members who will be included in the sample.
(b) If the president uses technology, determine the numbers for the 9 faculty members who will be included in the sample.

10. **Sampling the Students** The same community college from Problem 9 has 7,656 students currently enrolled in classes.

9. (a) 83, 67, 84, 38, 22, 24, 36, 58, 34

To gain the students' opinions about an upcoming building project, the college president wishes to obtain a simple random sample of 20 students. He numbers the students from 1 to 7,656.

(a) Using Table I from Appendix A, the president closes his eyes and drops his ink pen on the table. It points to the digit in row 11, column 32. Using this position as the starting point and proceeding downward, determine the numbers for the 20 students who will be included in the sample.

(b) If the president uses technology, determine the numbers for the 20 students who will be included in the sample.

11. Obtaining a Simple Random Sample The following table lists the 50 states.

NW

(a) Obtain a simple random sample of size 10 using Table I in Appendix A, a graphing calculator, or computer software.

(b) Obtain a second simple random sample of size 10 using Table I in Appendix A, a graphing calculator, or computer software.

1. Alabama	11. Hawaii	21. Massachusetts	31. New Mexico	41. South Dakota
2. Alaska	12. Idaho	22. Michigan	32. New York	42. Tennessee
3. Arizona	13. Illinois	23. Minnesota	33. North Carolina	43. Texas
4. Arkansas	14. Indiana	24. Mississippi	34. North Dakota	44. Utah
5. California	15. Iowa	25. Missouri	35. Ohio	45. Vermont
6. Colorado	16. Kansas	26. Montana	36. Oklahoma	46. Virginia
7. Connecticut	17. Kentucky	27. Nebraska	37. Oregon	47. Washington
8. Delaware	18. Louisiana	28. Nevada	38. Pennsylvania	48. West Virginia
9. Florida	19. Maine	29. New Hampshire	39. Rhode Island	49. Wisconsin
10. Georgia	20. Maryland	30. New Jersey	40. South Carolina	50. Wyoming

12. Obtaining a Simple Random Sample The following table lists the 43 presidents of the United States.

(a) Obtain a simple random sample of size 8 using Table I in Appendix A, a graphing calculator, or computer software.

(b) Obtain a second simple random sample of size 8 using Table I in Appendix A, a graphing calculator, or computer software.

1. Washington	10. Tyler	19. Hayes	28. Wilson	37. Nixon
2. J. Adams	11. Polk	20. Garfield	29. Harding	38. Ford
3. Jefferson	12. Taylor	21. Arthur	30. Coolidge	39. Carter
4. Madison	13. Fillmore	22. Cleveland	31. Hoover	40. Reagan
5. Monroe	14. Pierce	23. B. Harrison	32. F. D. Roosevelt	41. George H. Bush
6. J. Q. Adams	15. Buchanan	24. Cleveland	33. Truman	42. Clinton
7. Jackson	16. Lincoln	25. McKinley	34. Eisenhower	43. George W. Bush
8. Van Buren	17. A. Johnson	26. T. Roosevelt	35. Kennedy	
9. W. H. Harrison	18. Grant	27. Taft	36. L. B. Johnson	

13. Obtaining a Simple Random Sample Suppose you are the president of the student government. You wish to conduct a survey to determine the student body's opinion regarding student services. The administration provides you with a list of the names and phone numbers of the 19,935 registered students.

(a) Discuss the procedure you would follow to obtain a simple random sample of 25 students.

(b) Obtain this sample.

14. Obtaining a Simple Random Sample Suppose the mayor of Justice, Illinois, asks you to poll the residents of the village. The mayor provides you with a list of the names and phone numbers of the 5,832 residents of the village.

(a) Discuss the procedure you would follow to obtain a simple random sample of 20 residents.

(b) Obtain this sample.

15. Future Government Club The Future Government Club wants to sponsor a panel discussion on the upcoming national election. The club wants four of its members to lead the panel discussion. Obtain a simple random sample of size 4 from the table. Write a short description of the process you used to generate your sample.

Blouin	Fallenbuchel	Niemeyer	Rice
Bolden	Grajewski	Nolan	Salihar
Bolt	Haydra	Ochs	Tate
Carter	Keating	Opacian	Thompson
Cooper	Khouri	Pawlak	Trudeau
Debold	Lukens	Pechtold	Washington
De Young	May	Ramirez	Wright
Engler	Motola	Redmond	Zenkel

16. Worker Morale The owner of a private food store is concerned about employee morale. She decides to survey the employees to see if she can learn about work environment and job satisfaction. Obtain a simple random sample of size 5 from the names in the given table. Write a short description of the process you used to generate your sample.

Archer	Foushi	Kemp	Oliver
Bolcerek	Gow	Lathus	Orsini
Bryant	Grove	Lindsey	Salazar
Carlisle	Hall	Massie	Ullrich
Cole	Hills	McGuffin	Vaneck
Dimas	Houston	Musa	Weber
Ellison	Kats	Nickas	Zavodny
Everhart			

TECHNOLOGY STEP-BY-STEP Obtaining a Simple Random Sample

TI-83/84 Plus

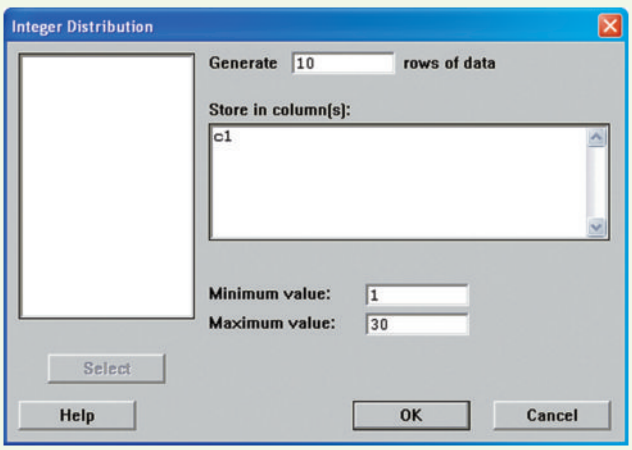
1. Enter any nonzero number (the seed) on the HOME screen.
2. Press the `STO ▶` button.
3. Press the `MATH` button.
4. Highlight the `PRB` menu and select `1: rand`.
5. From the HOME screen press `ENTER`.
6. Press the `MATH` button. Highlight `PRB` menu and select `5: randInt(`.
7. With `randInt(` on the HOME screen, enter `1`, *N*, where *N* is the population size. For example, if $N = 500$, enter the following:

`randInt(1,500)`

Press `ENTER` to obtain the first individual in the sample. Continue pressing `ENTER` until the desired sample size is obtained.

MINITAB

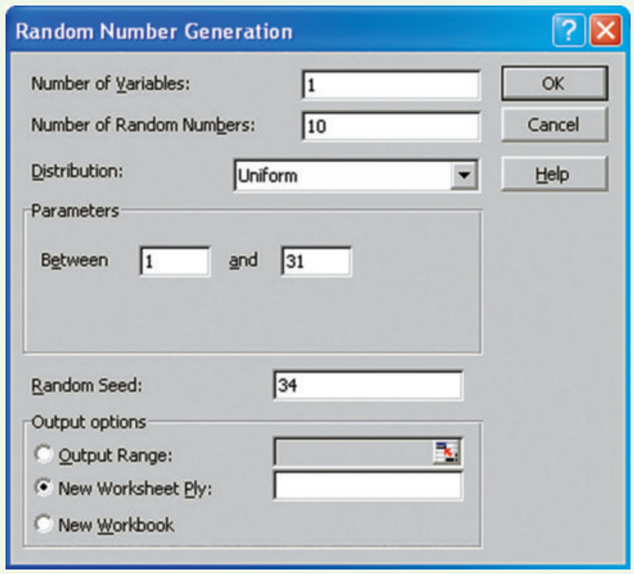
1. Select the **Calc** menu and highlight **Set Base** ....
2. Enter any seed number you desire. Note that it is not necessary to set the seed, because MINITAB uses the time of day in seconds to set the seed.
3. Select the **Calc** menu, highlight **Random Data**, and select **Integer** ....
4. Fill in the following window with the appropriate values. To obtain a simple random sample for the situation in Example 2, we would enter the following:



The reason we generate 10 rows of data (instead of 5) is in case any of the random numbers repeat. Select OK, and the random numbers will appear in column 1 (C1) in the spreadsheet.

Excel

1. Be sure the Data Analysis Tool Pak is activated. This is done by selecting the **Tools** menu and highlighting **Add – Ins** .... Check the box for the Analysis ToolPak and select OK.
2. Select **Tools** and highlight **Data Analysis** .... Highlight **Random Number Generation** and select OK.
3. Fill in the window with the appropriate values. To obtain a simple random sample for the situation in Example 2, we would fill in the following:



The reason we generate 10 rows of data (instead of 5) is in case any of the random numbers repeat. Notice also that the parameter is between 1 and 31, so any value greater than or equal to 1 and less than or equal to 31 is possible. In the unlikely event that 31 appears, simply ignore it. Select OK and the random numbers will appear in column 1 (A1) in the spreadsheet. Ignore any values to the right of the decimal place.

1.4 OTHER EFFECTIVE SAMPLING METHODS

Objectives
1. Obtain a stratified sample
2. Obtain a systematic sample
3. Obtain a cluster sample

Note to Instructor
The material in this section is optional and can be skipped without loss of continuity.

The goal of sampling is to obtain as much information as possible about the population at the least cost. Remember, we are using the word *cost* in a general sense. Cost includes monetary outlays, time, and other resources. With this goal in mind, we may find it advantageous to use sampling techniques other than simple random sampling.

1 Obtain a Stratified Sample

Under certain circumstances, *stratified sampling* provides more information about the population for less cost than simple random sampling.

Definition

A **stratified sample** is obtained by separating the population into nonoverlapping groups called *strata* and then obtaining a simple random sample from each stratum. The individuals within each stratum should be homogeneous (or similar) in some way.

For example, suppose Congress was considering a bill that abolishes estate taxes. In an effort to determine the opinion of her constituency, a senator asks a pollster to conduct a survey within her district. The pollster may divide the population of registered voters within the district into three strata: Republican, Democrat, and Independent. This grouping makes sense because the members within each of the three party affiliations may have the same opinion regarding estate taxes, but opinions between parties may differ. The main criterion in performing a stratified sample is that each group (stratum) must have a common attribute that results in the individuals being similar within the stratum.

In Other Words
Stratum is singular, while *strata* is plural. The word *strata* means divisions. So a stratified sample is a simple random sample of different divisions of the population.

An advantage of stratified sampling over simple random sampling is that it may allow fewer individuals to be surveyed while obtaining the same or more information. This result occurs because individuals within each subgroup have similar characteristics, so opinions within the group are not as likely vary much from one individual to the next. In addition, a stratified sample guarantees that each stratum is represented in the sample.

EXAMPLE 1 Obtaining a Stratified Sample

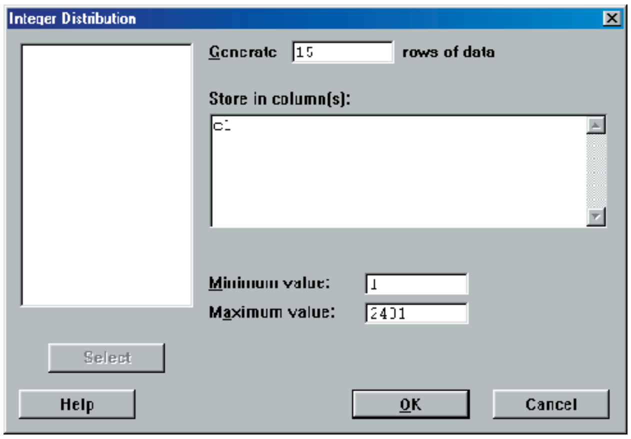
Problem: The president of DePaul University wants to conduct a survey to determine the community's opinion regarding campus safety. The president divides the DePaul community into three groups: resident students, nonresident (commuting) students, and staff (including faculty) so that he can obtain a stratified sample. Suppose there are 6,204 resident students, 13,304 nonresident students, and 2,401 staff, for a total of 21,909 individuals in the population. The president wants to obtain a sample of size 100, with the number of individuals selected from each stratum weighted by the population size. So resident students make up $6{,}204/21{,}909 = 28\%$ of the sample, nonresident students account for 61% of the sample, and staff constitute 11% of the sample. To obtain a sample of size 100, the president will obtain a stratified sample of $0.28(100) = 28$ resident students, $0.61(100) = 61$ nonresident students, and $0.11(100) = 11$ staff.

Approach: To obtain the stratified sample, conduct a simple random sample within each group. That is, obtain a simple random sample of 28 resident students (from the 6,204 resident students), a simple random sample of 61 nonresident students, and a simple random sample of 11 staff. Be sure to use a different seed for each stratum.

Solution: Using MINITAB, with the seed set to 4032 and the values shown in Figure 4, we obtain the following sample of staff:

240, 630, 847, 190, 2096, 705, 2320, 323, 701, 471, 744

Figure 4



CAUTION Do not use the same seed (or starting point in Table I) for all the groups in a stratified sample, because we want the simple random samples within each stratum to be independent of each other.

Repeat this procedure for the resident and nonresident students using a different seed.

An advantage of stratified sampling over simple random sampling is that the researcher is able to determine characteristics within each stratum. This allows an analysis to be performed on each subgroup to see if any significant differences between the groups exist. For example, we could analyze the data obtained in Example 1 to see if there is a difference in the opinions of students versus staff.

Now Work Problem 25

2 Obtain a Systematic Sample

In both simple random sampling and stratified sampling, it is necessary for a list of the individuals in the population being studied (the frame) to exist. Therefore, these sampling techniques require some preliminary work before the sample is obtained. A sampling technique that does not require a frame is *systematic sampling*.

Definition

A **systematic sample** is obtained by selecting every kth individual from the population. The first individual selected corresponds to a random number between 1 and k.

Because systematic sampling does not require a frame, it is a useful technique when you can't obtain a list of the individuals in the population that you wish to study.

The idea behind obtaining a systematic sample is relatively simple: Select a number k, randomly select a number between 1 and k and survey that individual, then survey every kth individual thereafter. For example, we might decide to survey every $k = 8$th individual. We randomly select a number between 1 and 8 such as 5. This means we survey the 5th, $5 + 8 = 13$th, $13 + 8 = 21$st, $21 + 8 = 29$th, and so on, individuals until we reach the desired sample size.

EXAMPLE 2 Obtaining a Systematic Sample without a Frame

Problem: The manager of Kroger Food Stores wants to measure the satisfaction of the store's customers. Design a sampling technique that can be used to obtain a sample of 40 customers.

Approach: A frame of Kroger customers would be difficult, if not impossible, to obtain. Therefore, it is reasonable to use systematic sampling by surveying every kth customer who leaves the store.

Solution: The manager decides to obtain a systematic sample by surveying every 7th customer. He randomly determines a number between 1 and 7, say 5. He then surveys the 5th customer exiting the store and every 7th customer thereafter, until a sample of 40 customers is reached. The survey will include customers 5, 12, 19, . . . , 278.*

But how do we select the value of k? If the size of the population is unknown, there is no mathematical way to determine k. It must be chosen by determining a value of k that is not so large that we are unable to achieve our desired sample size, but not so small that we do not obtain a sample size that is representative of the population.

To clarify this point, let's revisit Example 2. Suppose we chose a value of k that was too large, say 30. This means that we will survey every 30th shopper, starting with the 5th. To obtain a sample of size 40 would require that 1,175 shoppers visit Kroger on that day. If Kroger does not have 1,175 shoppers, the desired sample size will not be achieved. On the other hand, if k is too small, say 4, we would survey the 5th, 9th, . . . , 161st shopper. It may be that the 161st shopper exits the store at 3 P.M., which means our survey did not include any of the evening shoppers. Certainly, this sample is not representative of *all* Kroger patrons! An estimate of the size of the population would certainly help determine an appropriate value for k.

To determine the value of k when the size of the population, N, is known is relatively straightforward. Suppose we wish to survey a population whose size is known to be $N = 20{,}325$ and we desire a sample of size $n = 100$. To guarantee that individuals are selected evenly from both the beginning and the end of the population (such as early and late shoppers), we compute N/n and round down to the nearest integer. For example, $20{,}325/100 = 203.25$, so $k = 203$. Then we randomly select a number between 1 and 203 and select every 203rd individual thereafter. So, if we randomly selected 90 as our starting point, we would survey the 90th, 293rd, 496th, . . . , 20,187th individuals.

We summarize the procedure as follows:

Steps in Systematic Sampling

Step 1: If possible, approximate the population size, N.

Step 2: Determine the sample size desired, n.

Step 3: Compute $\frac{N}{n}$ and round down to the nearest integer. This value is k.

Step 4: Randomly select a number between 1 and k. Call this number p.

Step 5: The sample will consist of the following individuals:

$$p, p + k, p + 2k, \ldots, p + (n - 1)k$$

Because systematic sampling does not require a frame, it typically provides more information for a given cost than does simple random sampling. In addition, systematic sampling is easier to employ, so there is less likelihood of interviewer error occurring, such as selecting the wrong individual to be surveyed.

Now Work Problem 27

3 Obtain a Cluster Sample

A fourth sampling method is called *cluster sampling*. The previous three sampling methods discussed have benefits under certain circumstances. So does cluster sampling.

*Because we are surveying 40 customers, the first individual surveyed is the 5th, the second is the $5 + 7 = 12$th, the third is the $5 + (2)\,7 = 19$th, and so on, until we reach the 40th, which is the $5 + (39)\,7 = 278$th shopper.

Definition

A **cluster sample** is obtained by selecting all individuals within a randomly selected collection or group of individuals.

In Other Words
Imagine a mall parking lot. Each subsection of the lot could be a cluster (Section F-4, for example).

A school administrator wants to learn the characteristics of students enrolled in online classes. Rather than obtaining a simple random sample based on the frame of all students enrolled in online classes, the administrator could treat each online class as a cluster and then obtain a simple random sample of these clusters. The administrator would then survey *all* the students in the selected clusters.

EXAMPLE 3 Obtaining a Cluster Sample

Problem: A sociologist wants to gather data regarding household income within the city of Boston. Obtain a sample using cluster sampling.

Approach: The city of Boston can be set up so that each city block is a cluster. Once the city blocks have been identified, we obtain a simple random sample of the city blocks and survey all households on the blocks selected.

Solution: Suppose there are 10,493 city blocks in Boston. First, we must number the blocks from 1 to 10,493. Suppose the sociologist has enough time and money to survey 20 clusters (city blocks). Therefore, the sociologist should obtain a simple random sample of 20 numbers between 1 and 10,493 and survey all households from the clusters selected. Cluster sampling is a good choice in this example because it reduces the travel time to households that is likely to occur with both simple random sampling and stratified sampling. In addition, there is no need to obtain a detailed frame with cluster sampling. The only frame needed is one that provides information regarding city blocks.

Recall that in systematic sampling we had to determine an appropriate value for k, the number of individuals to skip between individuals selected to be in the sample. We have a similar problem in cluster sampling. The following are a few of the questions that arise:

- How do I cluster the population?
- How many clusters do I sample?
- How many individuals should be in each cluster?

CAUTION Stratified and cluster samples are different. In a stratified sample, we divide the population into two or more homogeneous groups. Then we obtain a simple random sample from each group. In a cluster sample, we divide the population into groups, obtain a simple random sample of some of the groups, and survey *all* individuals in the selected groups.

First, it must be determined whether the individuals within the proposed cluster are homogeneous (similar individuals) or heterogeneous (dissimilar individuals). Consider the results of Example 3. City blocks tend to have similar households. Surveying one house on a city block is likely to result in similar responses from another house on the same block. This results in duplicate information. We conclude the following: If the clusters have homogeneous individuals, it is better to have more clusters with fewer individuals in each cluster.

What if the cluster is heterogeneous? Under this circumstance, the heterogeneity of the cluster likely resembles the heterogeneity of the population. In other words, each cluster is a scaled-down representation of the overall population. For example, a quality-control manager might use shipping boxes that contain 100 light bulbs as a cluster, since the rate of defects within the cluster would closely mimic the rate of defects in the population, assuming the bulbs are randomly placed in the box. Thus, when each cluster is heterogeneous, fewer clusters with more individuals in each cluster are appropriate.

Now Work Problem 13

The four sampling techniques just presented are sampling techniques in which the individuals are selected randomly. Often, however, sampling methods are used in which the individuals are not randomly selected, such as *convenience sampling*.

Convenience Sampling

Have you ever been stopped in the mall by someone holding a clipboard? These folks are responsible for gathering information, but their methods of data collection are inappropriate, and the results of their analysis are suspect because they obtained their data using a *convenience sample*.

Definition

A **convenience sample** is a sample in which the individuals are easily obtained and not based on randomness.

CAUTION Studies that use convenience sampling generally have results that are suspect. The results should be looked on with extreme skepticism.

There are many types of convenience samples, but probably the most popular are those in which the individuals in the sample are **self-selected** (the individuals themselves decide to participate in a survey). These are also called **voluntary response** samples. Examples of self-selected sampling include phone-in polling; a radio personality will ask his or her listeners to phone the station to submit their opinions. Another example is the use of the Internet to conduct surveys. For example, *Dateline* will present a story regarding a certain topic and ask its viewers to "tell us what you think" by completing a questionnaire online or phoning in an opinion. Both of these samples are poor designs because the individuals who decide to be in the sample generally have strong opinions about the topic. A more typical individual in the population will not bother phoning or logging on to a computer to complete a survey. Any inference made regarding the population from this type of sample should be made with extreme caution.

Note to Instructor
Discuss some problems encountered in obtaining a sample from a nearby mall.

The reason convenience samples yield unreliable results is that the individuals chosen to participate in the survey are not chosen using random sampling. Instead, the interviewer or participant selects who is in the survey. Do you think an interviewer would select an ornery individual? Of course not! Therefore, the sample is likely not to be representative of the population.

Multistage Sampling

In practice, most large-scale surveys obtain samples using a combination of the techniques just presented.

As an example of multistage sampling, consider Nielsen Media Research. Nielsen randomly selects households and monitors the television programs these households are watching through a People Meter. The meter is an electronic box placed on each TV within the household. The People Meter measures what program is being watched and who is watching it. Nielsen selects the households with the use of a two-stage sampling process.

Stage 1: Using U.S. Census data, Nielsen divides the country into geographic areas (strata). The strata are typically city blocks in urban areas and geographic regions in rural areas. About 6,000 strata are randomly selected.

Stage 2: Nielsen sends representatives to the selected strata and lists the households within the strata. The households are then randomly selected through a simple random sample.

Nielsen sells the information obtained to television stations and companies. These results are used to help determine prices for commercials.

As another example of multistage sampling, consider the sample used by the Census Bureau for the Current Population Survey. This survey requires five stages of sampling:

Stage 1: Stratified sample
Stage 2: Cluster sample
Stage 3: Stratified sample
Stage 4: Cluster sample
Stage 5: Systematic sample

This survey is very important because it is used to obtain demographic estimates of the United States in noncensus years. A detailed presentation of the sampling method used by the Census Bureau can be found in *The Current Population Survey: Design and Methodology*, Technical Paper No. 40.

Sample Size Considerations

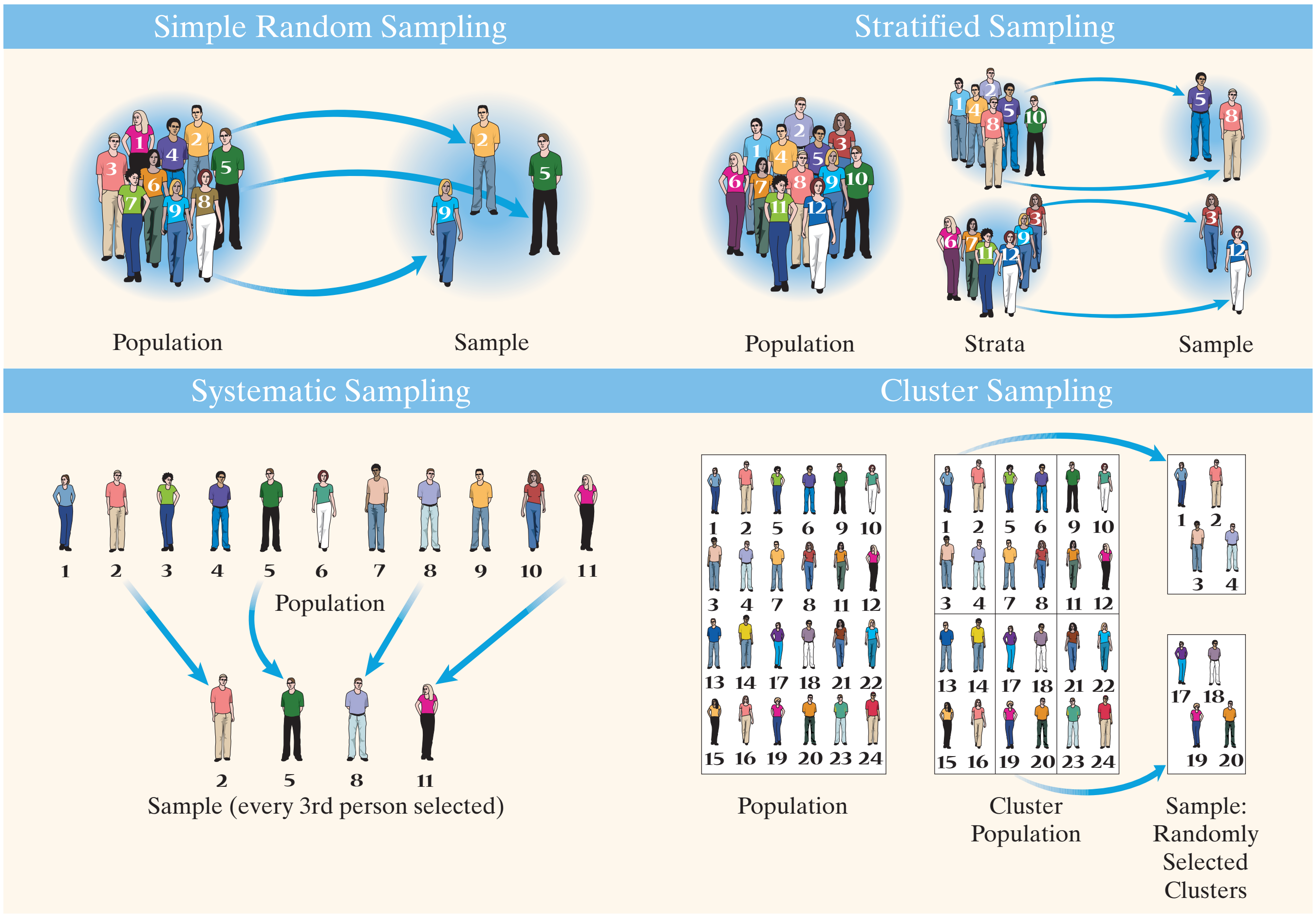
Throughout our discussion of sampling, we did not mention how to determine the sample size. Determining the sample size is key in the overall statistical process. In other words, the researcher must ask this question: "How many individuals must I survey to draw conclusions about the population within some predetermined margin of error?" The researcher must find the correct balance between the reliability of the results and the cost of obtaining these results. The bottom line is that time and money determine the level of confidence a researcher will place on the conclusions drawn from the sample data. The more time and money the researcher has available, the more accurate the results of the statistical inference will be.

Nonetheless, techniques do exist for determining the sample size required to estimate characteristics regarding the population within some margin of error. We will consider some of these techniques in Sections 9.1 and 9.3. (For a detailed discussion of sample size considerations, consult a text on sampling techniques such as *Elements of Sampling Theory and Methods* by Z. Govindarajulu, Prentice Hall, 1999.)

Summary

Figure 5 provides a summary of the four sampling techniques presented.

Figure 5



Note to Instructor
This activity should take about 20 to 30 minutes. You may also want students to discuss when each sampling technique might be appropriate.

IN CLASS ACTIVITY

Different Sampling Methods

The following question was recently asked by the Gallup Organization: In general, are you satisfied or dissatisfied with the way things are going in the country?

(a) Number the students in the class from 1 to *N*, where *N* is the number of students. Obtain a simple random sample and have them answer this question. Record the number of satisfied responses and the number of dissatisfied responses.

(b) Divide the students in the class by gender. Treat each gender as a stratum. Obtain a simple random sample from each stratum and have them answer this question. Record the number of satisfied responses and the number of dissatisfied responses.

(c) Treat each row of desks as a cluster. Obtain a simple random sample of clusters and have each student in the selected clusters answer this question. Record the number of satisfied responses and the number of dissatisfied responses.

(d) Number the students in the class from 1 to *N*, where *N* is the number of students. Obtain a systematic sample and have the selected students answer this question. Record the number of satisfied responses and the number of dissatisfied responses.

(e) Were there any differences in the results of the survey? State some reasons for any differences.

1.4 ASSESS YOUR UNDERSTANDING

Concepts and Vocabulary

1. Describe a circumstance in which stratified sampling would be an appropriate sampling method.

2. Which sampling method does not require a frame? Systematic

3. Why are convenience samples ill advised?

4. A(n) ______ is obtained by dividing the population into groups and selecting all individuals from within a random sample of the groups. cluster sample

5. A(n) ______ is obtained by dividing the population into homogeneous groups and randomly selecting individuals from each group. stratified sample

6. *True or False*: When taking a systematic random sample of size n, every group of size n from the population has the same chance of being selected. False

7. *True or False*: A simple random sample is always preferred because it obtains the same information as other sampling plans but requires a smaller sample size. False

8. *True or False*: When conducting a cluster sample, it is better to have fewer clusters with more individuals when the clusters are heterogeneous. True

9. *True or False*: Inferences based on voluntary response samples are generally not reliable. True

10. *True or False*: When obtaining a stratified sample, the number of individuals included within each stratum must be equal. False

Skill Building

In Problems 11–22, identify the type of sampling used.

11. To estimate the percentage of defects in a recent manufacturing batch, a quality-control manager at Intel selects every 8th chip that comes off the assembly line starting with the 3rd until she obtains a sample of 140 chips. Systematic

12. To determine the prevalence of human growth hormone (HGH) use among high school varsity baseball players, the State Athletic Commission randomly selects 50 high schools. All members of the selected high schools' varsity baseball teams are tested for HGH. Cluster

13. NW To determine customer opinion of its boarding policy, Southwest Airlines randomly selects 60 flights during a certain week and surveys all passengers on the flights. Cluster

14. A member of Congress wishes to determine her constituency's opinion regarding estate taxes. She divides her constituency into three income classes: low-income households, middle-income households, and upper-income households. She then takes a simple random sample of households from each income class. Stratified

15. In an effort to identify if an advertising campaign has been effective, a marketing firm conducts a nationwide poll by randomly selecting individuals from a list of known users of the product. Simple random

16. A radio station asks its listeners to call in their opinion regarding the use of U.S. forces in peacekeeping missions. Convenience

29. SRS: number from 1 to 1,280; randomly select 128 students to survey. Stratified: randomly select 4 students from each section to survey. Cluster: randomly select 4 sections, survey all students in these 4 sections. Answers will vary.

17. A farmer divides his orchard into 50 subsections, randomly selects 4, and samples all the trees within the 4 subsections to approximate the yield of his orchard. Cluster

18. A school official divides the student population into five classes: freshman, sophomore, junior, senior, and graduate student. The official takes a simple random sample from each class and asks the members' opinions regarding student services. Stratified

19. A survey regarding download time on a certain website is administered on the Internet by a market research firm to anyone who would like to take it. Convenience

20. The presider of a guest-lecture series at a university stands outside the auditorium before a lecture begins and hands every fifth person who arrives, beginning with the third, a speaker evaluation survey to be completed and returned at the end of the program. Systematic

21. To determine his DSL Internet connection speed, Shawn divides up the day into four parts: morning, midday, evening, and late night. He then measures his Internet connection speed at 5 randomly selected times during each part of the day. Stratified

22. 24 Hour Fitness wants to administer a satisfaction survey to its current members. Using its membership roster, the club randomly selects 40 club members and asks them about their level of satisfaction with the club. Simple random

23. A salesperson obtained a systematic sample of size 20 from a list of 500 clients. To do so, he randomly selected a number from 1 of 25, obtaining the number 16. He included in the sample the 16th client on the list and every 25th client thereafter. List the numbers that correspond to the 20 clients selected.

24. A quality-control expert wishes to obtain a cluster sample by selecting 10 of 795 clusters. She numbers the clusters from 1 to 795. Using Table I from Appendix A, she closes her eyes and drops a pencil on the table. It points to the digit in row 8, column 38. Using this position as the starting point and proceeding downward, determine the numbers for the 10 clusters selected. 763, 185, 377, 304, 626, 392, 315, 084, 565, 508

23. 16, 41, 66, 91, 116, 141, 166, 191, 216, 241, 266, 291, 316, 341, 366, 391, 416, 441, 466, 491

Applying the Concepts

25. Stratified Sampling NW The Future Government Club wants to sponsor a panel discussion on the upcoming national election. The club wants to have four of its members lead the panel discussion. To be fair, however, the panel should consist of two Democrats and two Republicans. From the list of current members of the club, obtain a stratified sample of two Democrats and two Republicans to serve on the panel.

Democrats		Republicans	
Bolden	Motola	Blouin	Ochs
Bolt	Nolan	Cooper	Pechtold
Carter	Opacian	De Young	Redmond
Debold	Pawlak	Engler	Rice
Fallenbuchel	Ramirez	Grajewski	Salihar
Haydra	Tate	Keating	Thompson
Khouri	Washington	May	Trudeau
Lukens	Wright	Niemeyer	Zenkel

26. Stratified Sampling The owner of a private food store is concerned about employee morale. She decides to survey the managers and hourly employees to see if she can learn about work environment and job satisfaction. From the list of workers at the store, obtain a stratified sample of two managers and four hourly employees to survey.

Managers		Hourly Employees		
Carlisle	Oliver	Archer	Foushi	Massie
Hills	Orsini	Bolcerek	Gow	Musa
Kats	Ullrich	Bryant	Grove	Nickas
Lindsey	McGuffin	Cole	Hall	Salazar
		Dimas	Houston	Vaneck
		Ellison	Kemp	Weber
		Everhart	Lathus	Zavodny

27. Systematic Sample NW The human resource department at a certain company wants to conduct a survey regarding worker morale. The department has an alphabetical list of all 4,502 employees at the company and wants to conduct a systematic sample.

(a) Determine k if the sample size is 50. 90

(b) Determine the individuals who will be administered the survey. More than one answer is possible.

28. Systematic Sample To predict the outcome of a county election, a newspaper obtains a list of all 945,035 registered voters in the county and wants to conduct a systematic sample.

(a) Determine k if the sample size is 130. 7269

(b) Determine the individuals who will be administered the survey. More than one answer is possible.

29. Which Method? The mathematics department at a university wishes to administer a survey to a sample of students taking college algebra. The department is offering 32 sections of college algebra, similar in class size and makeup, with a total of 1,280 students. They would like the sample size to be roughly 10% of the population of college algebra students this semester. How might the department obtain a simple random sample? A stratified sample? A cluster sample? Which method do you think is best in this situation?

30. Good Sampling Method? To obtain students' opinions about proposed changes to course registration procedures, the administration of a small college asked for faculty volunteers who were willing to administer a survey in one of their classes. Twenty-three faculty members volunteered. Each of these faculty members gave the survey to all the students in one course of their choosing. Would this sampling method be considered a cluster sample? Why or why not? No. The "clusters" were not randomly selected.

31. Sample Design The city of Naperville is considering the construction of a new commuter rail station. The city wishes to survey the residents of the city to obtain their opinion regarding the use of tax dollars for this purpose. Design a sampling method to obtain the individuals in the sample. Be sure to support your choice.

32. Sample Design A school board at a local community college is considering raising the student services fees. The board wants to obtain the opinion of the student body before proceeding. Design a sampling method to obtain the individuals in the sample. Be sure to support your choice.

33. Sample Design Target wants to open a new store in the village of Lockport. Before construction, Target's marketers want to obtain some demographic information regarding the area under consideration. Design a sampling method to obtain the individuals in the sample. Be sure to support your choice.

34. Sample Design The county sheriff wishes to determine if a certain highway has a high proportion of speeders traveling on it. Design a sampling method to obtain the individuals in the sample. Be sure to support your choice.

35. Sample Design A pharmaceutical company wants to conduct a survey of 30 individuals who have high cholesterol. The company has obtained a list from doctors throughout the country of 6,600 individuals who are known to have high cholesterol. Design a sampling method to obtain the individuals in the sample. Be sure to support your choice.

36. Sample Design A marketing executive for Coca-Cola, Inc., wants to identify television shows that people in the Boston area who typically drink Coke are watching. The executive has a list of all households in the Boston area. Design a sampling method to obtain the individuals in the sample. Be sure to support your choice.

37. Putting It Together: Comparing Sampling Methods Suppose a political strategist wants to get a sense of how American adults aged 18 years or older feel about health care and health insurance.

(a) In a political poll, what would be a good frame to use for obtaining a sample?

(b) Explain why simple random sampling may not guarantee that the sample has an accurate representation of registered Democrats, registered Republicans, and registered Independents.

(c) How can stratified sampling guarantee this representation?

37. (a) Registered voters who have voted in the past few elections

38. Putting It Together: Thinking about Randomness What is random sampling? Why is it necessary for a sample to be obtained randomly rather than conveniently? Will randomness guarantee that a sample will provide accurate information about the population? Explain.

39. Research the origins of the Gallup Poll and the current sampling method the organization uses. Report your findings to the class.

40. Research the sampling methods used by a market research firm in your neighborhood. Report your findings to the class. The report should include the types of sampling methods used, number of stages, and sample size.

1.5 BIAS IN SAMPLING

Objective 1 Explain the sources of bias in sampling

1 Explain the Sources of Bias in Sampling

Note to Instructor

If you are pressed for time, this section can be given to the students as a reading assignment with light in-class coverage.

So far we have looked at *how* to obtain samples, but not at some of the problems that inevitably arise in sampling. Remember, the goal of sampling is to obtain information about a population through a sample.

Definition

If the results of the sample are not representative of the population, then the sample has **bias**.

In Other Words

The word *bias* could mean to give preference to selecting some individuals over others. It could also mean that certain responses are more likely to occur in the sample than in the population.

There are three sources of bias in sampling:

1. Sampling bias
2. Nonresponse bias
3. Response bias

Sampling Bias

Sampling bias means that the technique used to obtain the individuals to be in the sample tends to favor one part of the population over another. Any convenience sample has sampling bias because the individuals are not chosen through a random sample. For example, a voluntary response sample will have sampling bias because the opinions of individuals who decide to be in the sample are probably not representative of the population as a whole.

Sampling bias also results due to *undercoverage*. **Undercoverage** occurs when the proportion of one segment of the population is lower in a sample than it is in the population. Undercoverage can result because the frame used to obtain the sample is incomplete or not representative of the population. Recall that the frame is the list of all individuals in the population under study. Sometimes, obtaining the frame would seem to be a relatively easy task, such as obtaining the list of all registered voters for a study regarding voter preference in an upcoming election. Even under this circumstance, however, the frame may be incomplete since people who recently registered to vote may not be on the published list of registered voters.

Sampling bias can result in incorrect predictions. For example, the magazine *Literary Digest* predicted that Alfred M. Landon would defeat Franklin D. Roosevelt in the 1936 presidential election. The *Literary Digest* conducted a poll by mailing questionnaires based on a list of its subscribers, telephone directories, and automobile owners. On the basis of the results, the *Literary Digest* predicted that Landon would win the election with 57% of the popular vote. However, Roosevelt won the election with about 62% of the popular vote. Bear in mind that this election was taking place during the height of the Great Depression. The incorrect prediction by the *Literary Digest* was the result of sampling bias. In 1936, most subscribers to the magazine, households with telephones, and automobile owners were Republican, the party of Landon. Therefore, the choice of the frame used to conduct the survey led to an incorrect prediction. Essentially, there was undercoverage of Democrats.

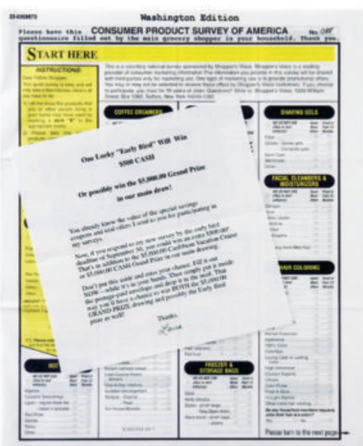
Often, it is difficult to gain access to a *complete* list of individuals in a population. For example, in public-opinion polls, random telephone surveys are frequently conducted, which implies that the frame is all households with telephones. This method of sampling will exclude any household that does not have a telephone, as well as homeless people. If the individuals without a telephone or homeless people differ in some way from people with a telephone or with homes, then the results of the sample may not be valid.

Nonresponse Bias

Nonresponse bias exists when individuals selected to be in the sample who do not respond to the survey have different opinions from those who do. Nonresponse can occur because individuals selected for the sample do not wish to respond or the interviewer was unable to contact them.

All surveys will suffer from nonresponse. The federal government uses a complex random sample to select individuals to participate in its Current Population Survey. Overall, the response rate is about 92%, but it varies depending on the age of the individual. For example, the response rate for 20- to 29-year-olds is 85%, while the response rate for individuals at least 70 years of age is 99%. Response rates in random digit dialing (RDD) telephone surveys are typically around 70%. Response rates for e-mail surveys typically hover around 40%, and mail surveys can have response rates as high as 60%.

Nonresponse bias can be controlled using callbacks. For example, if nonresponse occurs because a mailed questionnaire was not returned, a callback might mean phoning the individual to conduct the survey. If nonresponse occurs because an individual was not at home, a callback might mean returning to the home at other times in the day or on other days of the week.



Another method to improve nonresponse is using rewards and incentives. Rewards may include cash payments for completing a questionnaire. Incentives might include a cover letter that states that the responses to the questionnaire will determine future policy. For example, I received $1 with a survey regarding my satisfaction with a recent purchase. The $1 "payment" was meant to make me feel guilty enough to fill out the questionnaire. As another example, a city may send out questionnaires to households and state in a cover letter that the responses to the questionnaire will be used to decide pending issues within the city.

Let's consider the *Literary Digest* poll again. The *Literary Digest* mailed out more than 10 million questionnaires and 2.3 million people responded. The rather low response rate (23%) contributed to the *Literary Digest* making an incorrect prediction. After all, Roosevelt was the incumbent president and only those who were unhappy with his administration were likely to respond. By the way, in the same election, the 35-year-old George Gallup predicted that Roosevelt would win the election. He surveyed only 50,000 people to come to his conclusion.

Response Bias

Response bias exists when the answers on a survey do not reflect the true feelings of the respondent. Response bias can find its way into survey results in a number of ways.

Interviewer Error A trained interviewer is essential to obtain accurate information from a survey. A good interviewer will have the skill necessary to elicit responses from individuals within a sample and be able to make the interviewee feel comfortable enough to give truthful responses. For example, a good interviewer should be able to obtain truthful answers to questions as sensitive as "Have you ever cheated on your taxes?" Do not be quick to trust surveys that are conducted by poorly trained interviewers. Do not trust survey results if the sponsor has a vested interest in the results of the survey. Would you trust a survey conducted by a car dealer that reports 90% of customers say they would buy another car from the dealer?

Misrepresented Answers Some survey questions result in responses that misrepresent facts or are flat-out lies. For example, a survey of recent college graduates may find that self-reported salaries are somewhat inflated. Also, people may overestimate their abilities. For example, ask people how many push-ups they can do in 1 minute, and then ask them to do the push-ups. How accurate were they?

CAUTION The wording of questions can significantly affect the responses and, therefore, the validity of a study.

Wording of Questions The wording of a question plays a large role in the type of response given to the question. The way a question is worded can lead to response bias in a survey, so questions must always be asked in balanced form. For example, the "yes/no" question

Do you oppose the reduction of estate taxes?

should be written

Do you favor or oppose the reduction of estate taxes?

The second question is balanced. Do you see the difference? Consider the following report based on studies from Schuman and Presser (*Questions and Answers in Attitude Surveys*, 1981, p. 277), who asked the following two questions:

(A) Do you think the United States should forbid public speeches against democracy?

(B) Do you think the United States should allow public speeches against democracy?

For those respondents presented with question A, 21.4% gave "yes" responses, while for those given question B, 47.8% gave "no" responses. The conclusion you may arrive at is that most people are not necessarily willing to forbid something, but more people are willing not to allow something. These results imply that the wording of the question can alter the outcome of a survey.

Another consideration in wording a question is not to be vague. For example, the question "How much do you study?" is too vague. Does the researcher mean how much do I study for all my classes or just for statistics? Does the researcher mean per day or per week? The question should be written "How many hours do you study statistics each week?"

Ordering of Questions or Words Many surveys will rearrange the order of the questions within a questionnaire so that responses are not affected by prior questions. Consider the following example from Schuman and Presser in which the following two questions were asked:

(A) Do you think the United States should let Communist newspaper reporters from other countries come in here and send back to their papers the news as they see it?

(B) Do you think a Communist country such as Russia should let American newspaper reporters come in and send back to America the news as they see it?

For surveys conducted in 1980 in which the questions appeared in the order (A, B), 54.7% of respondents answered "yes" to A and 63.7% answered "yes" to B. If the questions were ordered (B, A), then 74.6% answered "yes" to A and 81.9% answered "yes" to B. When Americans are first asked if U.S. reporters should be allowed to report Communist news, they are more likely to agree that Communists should be allowed to report American news. Questions should be rearranged as much as possible to help reduce effects of this type.

Pollsters will also rearrange words within a question. For example, the Gallup Organization asked the following question of 1,017 adults aged 18 years or older:

> **Do you [rotated: approve (or) disapprove] of the job George W. Bush is doing as president?**

Notice how the words *approve* and *disapprove* were rotated. The purpose of this is to remove the effect that may occur by writing the word *approve* first in the question.

Type of Question One of the first considerations in designing a question is determining whether the question should be *open* or *closed*.

An **open question** is one for which the respondent is free to choose his or her response. For example:

What is the most important problem facing America's youth today?

A **closed question** is one for which the respondent must choose from a list of predetermined responses.

What is the most important problem facing America's youth today?

(a) Drugs
(b) Violence
(c) Single-parent homes
(d) Promiscuity
(e) Peer pressure

Not only should the order of the questions or certain words within the question be rearranged, but in closed questions the possible responses should also be rearranged. The reason is that respondents are likely to choose early choices in a list rather than later choices.

When designing an open question, be sure to phrase the question so that the responses are similar. (You don't want a wide variety of responses.) This allows for easy analysis of the responses. The benefit of closed questions is that they limit the number of respondent choices and, therefore, the results are much easier to analyze. However, this limits the choices and does not always allow the respondent to respond the way he or she might want to. If the desired answer is not provided as a choice, the respondent will be forced to choose a secondary answer or skip the question.

Survey designers recommend conducting pretest surveys with open questions and then using the most popular answers as the choices on closed-question surveys. Another issue to consider in the closed-question design is the number of responses the respondent may choose from. It is recommended that the option "no opinion" be omitted, because this option does not allow for meaningful analysis. The bottom line is to try to limit the number of choices in a closed-question format without forcing respondents to choose an option they otherwise would not. If the respondents choose an option they otherwise would not choose, the survey will have response bias.

Data-entry Error Although not technically a result of response bias, data-entry error will lead to results that are not representative of the population. Once data are collected, the results typically must be entered into a computer, which could result in input errors. For example, 39 may be entered as 93. It is imperative that data be checked for accuracy. In this text, we present some suggestions for checking for data error.

Can a Census Have Bias?

The discussion thus far has focused on bias in samples. This is not to imply that bias cannot occur when conducting a census, however. For example, it is entirely possible that a question on a census form is misunderstood, thereby leading to response bias in the results. We also mentioned that it is often difficult to contact each individual in a population. For example, the U.S. Census Bureau is challenged to count each homeless person in the country, so the census data published by the U.S. government likely suffers from nonresponse bias.

Sampling Error versus Nonsampling Error

Nonresponse bias, response bias, and data-entry errors are types of *nonsampling error*. However, whenever a sample is used to learn information about a population, there will inevitably also be *sampling error.*

Definitions

Nonsampling errors are errors that result from undercoverage, nonresponse bias, response bias, or data-entry error. Such errors could also be present in a complete census of the population. **Sampling error** is the error that results from using a sample to estimate information about a population. This type of error occurs because a sample gives incomplete information about a population.

By incomplete information, we mean that the individuals in the sample cannot reveal all the information about the population. Consider the following: Suppose that we wanted to determine the average age of the students enrolled in an introductory statistics course. To do this, we obtain a simple random sample of four students and ask them to write their age on a sheet of paper and turn it in. The average age of these four students is found to be 23.25 years. Assume that no students lied about their age, nobody misunderstood the question, and the sampling was done appropriately. If the actual average age of all 30 students in the class (the population) is 22.91 years, then the sampling error is $23.25 - 22.91 = 0.34$ year. Now suppose that the same survey is conducted, but this time one individual lies about his age. Then the results of the survey will also have nonsampling error.

In Other Words

We can think of sampling error as error that results from using a subset of the population to describe characteristics of the population. Nonsampling error is error that results from obtaining and recording the information collected.

IN CLASS ACTIVITY

A Classroom Survey

As a class, answer the following questions. Throughout the semester, the results of the survey can be used to illustrate various statistical concepts.

1. What is your gender?
2. What is your age?
3. How many semester hours are you enrolled in this semester?
4. How many minutes did you watch television last night?
5. What is your major? If you don't know, state undeclared.
6. How many hours did you work last week? If you don't work, write 0.
7. How many siblings do you have (include half- and step-siblings)?
8. Do you own your own car? If so, what make (Chevrolet, Honda, etc.)?
9. Do you speak more than one language fluently? If so, what language(s)?
10. How many hours do you study each week?
11. How many hours did you study last night?
12. How long does it take you (in minutes) to get to campus?
13. What is your eye color?

Note to Instructor

The purpose of this activity is to collect data about your class. You can then use the data to illustrate various concepts throughout the semester. To do some of the analysis, you will need to make the assumption that the data are the result of a simple random sample. This activity should only take 5 to 10 minutes.

1.5 ASSESS YOUR UNDERSTANDING

Concepts and Vocabulary

1. Why is it rare for frames to be completely accurate?
2. What are some solutions to nonresponse?
3. What is a closed question? What is an open question? Discuss the advantages and disadvantages of each type of question.
4. What does it mean when a part of the population is underrepresented?
5. Discuss the benefits of having trained interviewers.
6. What are the advantages of having a presurvey when constructing a questionnaire that has closed questions?
7. Discuss the pros and cons of telephone interviews that take place during dinner time in the early evening.
8. Why is a high response rate desired? How would a low response rate affect survey results?
9. Discuss why the order of questions or choices within a questionnaire are important in sample surveys.
10. Suppose a survey asks, "Do you own any CDs?" Explain how this could be interpreted in more than one way. Suggest a way in which the question could be improved.
11. What is bias? Name the three sources of bias and provide an example of each. How can a census have bias?
12. Distinguish between nonsampling error and sampling error.

Skill Building

In Problems 13–24, the survey has bias. (a) Determine the type of bias. (b) Suggest a remedy.

13. A retail store manager wants to conduct a study regarding the shopping habits of his customers. He selects the first 60 customers who enter his store on a Saturday morning. Sampling bias
14. The village of Oak Lawn wishes to conduct a study regarding the income level of households within the village. The village manager selects 10 homes in the southwest corner of the village and sends an interviewer to the homes to determine household income. Sampling bias
15. An antigun advocate wants to estimate the percentage of people who favor stricter gun laws. He conducts a nationwide survey of 1,203 randomly selected adults 18 years old and older. The interviewer asks the respondents, "Do you favor harsher penalties for individuals who sell guns illegally?" Response bias; poorly worded question
16. Suppose you are conducting a survey regarding the sleeping habits of students. From a list of registered students, you obtain a simple random sample of 150 students. One survey question is "How much sleep do you get?"
17. A polling organization conducts a study to estimate the percentage of households that speaks a foreign language as the primary language. It mails a questionnaire to 1,023 randomly selected households throughout the United States and asks the head of household if a foreign language is the primary language spoken in the home. Of the 1,023 households selected, 12 responded. Nonresponse
18. Cold Stone Creamery is considering opening a new store in O'Fallon. Before opening the store, the company would like to know the percentage of households in O'Fallon that regularly visit an ice cream shop. The market researcher obtains a list of households in O'Fallon and randomly selects 150 of them. He mails a questionnaire to the 150 households that asks about ice cream eating habits and flavor preferences. Of the 150 questionnaires mailed, 4 are returned. Nonresponse
19. A newspaper article reported, "The *Cosmopolitan* magazine survey of more than 5,000 Australian women aged 18–34 found about 42 percent considered themselves overweight or obese."

    Source: *Herald Sun*, September 9, 2007
20. A health teacher wishes to do research on the weight of college students. She obtains the weights for all the students in her 9 A.M. class by looking at their driver's licenses or state IDs. Sampling bias and possible response bias
21. A magazine is conducting a study on the effects of infidelity in a marriage. The editors randomly select 400 women whose husbands were unfaithful and ask, "Do you believe a marriage can survive when the husband destroys the trust that must exist between husband and wife?"
22. A textbook publisher wants to determine what percentage of college professors either require or recommend that their students purchase textbook packages with supplemental materials, such as study guides, digital media, and online tools. The publisher sends out surveys by e-mail to a random sample of 320 faculty members who have registered with its website and have agreed to receive solicitations. The publisher reports that 80% of college professors require or recommend that their students purchase some type of textbook package. Sampling bias; poor choice of a frame
23. Suppose you are conducting a survey regarding illicit drug use among teenagers in the Baltimore school district. You obtain a cluster sample of 12 schools within the district and sample all sophomore students in the randomly selected schools. The survey is administered by the teachers. Response bias
24. To determine the public's opinion of the police department, the police chief obtains a cluster sample of 15 census tracts within his jurisdiction and samples all households in the randomly selected tracts. Uniformed police officers go door to door to conduct the survey. Response bias

Applying the Concepts

25. **Response Rates** Surveys tend to suffer from low response rates. Based on past experience, a researcher determines that the typical response rate for an e-mail survey is 40%. She wishes to obtain a sample of 300 respondents, so she e-mails the survey to 1500 randomly selected e-mail addresses. Assuming the response rate for her survey is 40%, will the respondents form an unbiased sample? Explain.
26. **Delivery Format** The General Social Survey asked, "About how often did you have sex in the past 12 months?" About 47% of respondents indicated they had sex at least once a week. In a Web survey for a marriage and family wellness center, respondents were asked, "How often do you and your partner have sex (on average)?" About 31% of respondents indicated they had sex with their partner at least once a week. Explain how the delivery method for such a question could result in biased responses.

16. Response bias; poorly worded question 19. Sampling bias and possible interviewer error 21. Response bias; poorly worded question
25. No; undercoverage (sampling bias), nonresponse bias, and response bias

27. Order of the Questions Consider the following two questions:

(a) Suppose that a rape is committed in which the woman becomes pregnant. Do you think the criminal should or should not face additional charges if the woman becomes pregnant?

(b) Do you think abortions should be legal under any circumstances, legal under certain circumstances, or illegal in all circumstances?

Do you think the order in which the questions are asked will affect the survey results? If so, what can the pollster do to alleviate this response bias?

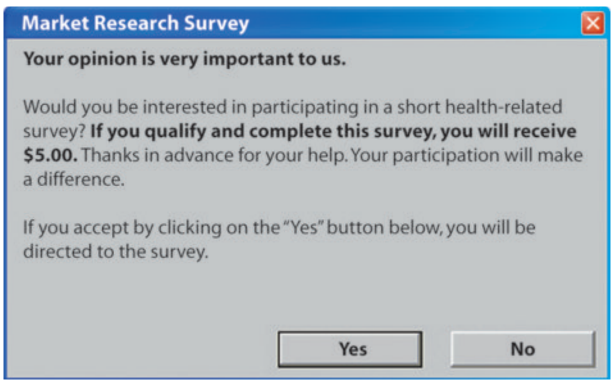
28. Order of the Questions Consider the following two questions:

(a) Do you believe that the government should or should not be allowed to prohibit individuals from expressing their religious beliefs at their place of employment?

(b) Do you believe that the government should or should not be allowed to prohibit teachers from expressing their religious beliefs in public school classrooms?

Do you think the order in which the questions are asked will affect the survey results? If so, what can the pollster do to alleviate this response bias? Discuss the choice of the word *prohibit* in the survey questions.

29. Improving Response Rates Suppose you are reading an article at psychcentral.com and the following text appears in a pop-up window:



What tactic is the company using to increase the response rate for its survey?

30. Rotating Choices Consider this question from a recent Gallup poll:

Which of the following approaches to solving the nation's energy problems do you think the U.S. should follow right now—[ROTATED: emphasize production of more oil, gas and coal supplies (or) emphasize more conservation by consumers of existing energy supplies]?

Why is it important to rotate the two choices presented in the question?

31. Random Digit Dialing Many polls use random digit dialing (RDD) to obtain a sample, which means a computer randomly generates phone numbers. What is the frame for this type of sampling? Who would be excluded from the survey and how might this affect the results of the survey?

32. Caller ID How do you think caller ID has affected phone surveys?

33. Don't Call Me! The Telephone Consumer Protection Act (TCPA) allows consumers to put themselves on a do-not-call registry. If a number is on the registry, commercial telemarketers are not allowed to call you. Do you believe this has affected the ability of surveyors to obtain accurate polling results? If so, how?

34. Current Population Survey In the federal government's Current Population Survey, the response rate for 20- to 29-year-olds is 85%, while response rates for individuals at least 70 years of age is 99%. Why do you think this is?

35. Analyze an Article Read the following article from the January 20, 2005 *USA Today*. What types of nonsampling errors led to incorrect exit polls?

Firms Report Flaws That Threw Off Exit Polls

Kerry backers' willingness, pollsters' inexperience cited

By Mark Memmott, *USA Today*

The exit polls of voters on Election Day so overstated Sen. John Kerry's support that, going back to 1988, they rank as the most inaccurate in a presidential election, the firms that did the work concede.

One reason the surveys were skewed, they say, was because Kerry's supporters were more willing to participate than Bush's. Also, the people they hired to quiz voters were on average too young and too inexperienced and needed more training.

The exit polls, which are supposed to help the TV networks shape their coverage on election night, were sharply criticized. Leaks of preliminary data showed up on the Internet in the early afternoon of Election Day, fueling talk that Kerry was beating President Bush. After the election, some political scientists, pollsters and journalists questioned their value.

In a report to the six media companies that paid them to conduct the voter surveys, pollsters Warren Mitofsky and Joseph Lenski said Wednesday that "on average, the results from each precinct overstated the Kerry-Bush difference by 6.5 (percentage) points. This is the largest (overstatement) we have observed . . . in the last five presidential elections."

Lenski said Wednesday that issuing the report was like "hanging out your dirty underwear. You hope it's cleaner than people expected."

Among the findings:

- They hired too many relatively young adults to conduct the interviews. Half of the 1,400 interviewers were younger than 35. That may explain in part why Kerry voters were more inclined to participate, since he drew more of the youth vote than did Bush. But Mitofsky and Lenski also found younger interviewers were more likely to make mistakes.
- Early results were skewed by a "programming error" that led to including too many female voters. Kerry outpolled Bush among women.
- Some local officials prevented interviewers from getting close to voters.

For future exit polls, Lenski and Mitofsky recommended hiring more experienced polltakers and giving them better training, and working with election officials to ensure access to polling places.

Lenski and Mitofsky noted that none of the media outlets they worked for—ABC, CBS, CNN, Fox News, NBC and the Associated Press—made any wrong "calls" on election night. Representatives of those six are reviewing the report. Many other news media, including *USA Today*, also paid to get some of the data.

Source: USA TODAY. January 20, 2005. Reprinted with Permission.

36. **Increasing Response Rates** Offering rewards or incentives is one way of attempting to increase response rates. Discuss a possible disadvantage of such a practice.

37. **Wording Survey Questions** Write a survey question that contains strong wording and a survey question that contains tempered wording. Present the strongly worded question to 10 randomly selected people and the tempered question to 10 different randomly selected people. How does the wording affect the response?

38. **Order in Survey Questions** Write two questions that could have different responses, depending on the order in which the questions are presented. Randomly select 20 people and present the questions in one order to 10 of the people and in the opposite order to the other 10 people. Did the results differ?

39. Research a survey method used by a company or government branch. Determine the sampling method used, the sample size, the method of collection, and the frame used.

40. **Informed Opinions** People often respond to survey questions without any knowledge of the subject matter. A common example of this is the discussion on banning dihydrogen monoxide. The Centers for Disease Control (CDC) reports that there were 1,493 deaths due to asbestos in 2002, but over 3,200 deaths were attributed to dihydrogen monoxide in 2000. Articles and Web sites, such as www.dhmo.org tell how this substance is widely used despite the dangers associated with it. Many people have joined the cause to ban this substance without realizing that dihydrogen monoxide is simply water (H_2O). Their eagerness to protect the environment or their fear of seeming uninformed may be part of the problem. Put together a survey that asks individuals whether dihydrogen monoxide should or should not be banned. Give the survey to 20 randomly selected students around campus and report your results to the class. An example survey might look like the following:

Dihydrogen monoxide is colorless, odorless, and kills thousands of people every year. Most of these deaths are caused by accidental inhalation, but the dangers of dihydrogen monoxide do not stop there. Prolonged exposure to its solid form can severely damage skin tissue. Symptoms of ingestion can include excessive sweating and urination and possibly a bloated feeling, nausea, vomiting, and body electrolyte imbalance. Dihydrogen monoxide is a major component of acid rain and can cause corrosion after coming in contact with certain metals.

Do you believe that the government should or should not ban the use of dihydrogen monoxide?

41. Name two biases that led to the *Literary Digest* making an incorrect prediction in the presidential election of 1936.

42. **Research on George Gallup** Research the polling done by George Gallup in the 1936 presidential election. Write a report on your findings. Be sure to include information about the sampling technique and sample size. Now research the polling done by Gallup for the 1948 presidential election. Did Gallup accurately predict the outcome of the election? What lessons were learned by Gallup?

43. **Putting It Together: Speed Limit** In the state of California, speed limits are established through traffic engineering surveys. One aspect of the survey is for city officials to measure the speed of vehicles on a particular road.

Source: www.ci.eureka.ca.gov, www.nctimes.com

(a) What is the population of interest for this portion of the engineering survey?
(b) What is the variable of interest for this portion of the engineering survey? Speed
(c) Is the variable qualitative or quantitative? Quantitative
(d) What is the level of measurement for the variable? Ratio
(e) Is a census feasible in this situation? Explain why or why not. No
(f) Is a sample feasible in this situation? If so, explain what type of sampling plan could be used? If not, explain why not. Systematic
(g) In July 2007, the Temecula City Council refused a request to increase the speed limit on Pechanga Parkway from 40 to 45 mph despite survey results indicating that the prevailing speed on the parkway favored the increase. Opponents were concerned that it was visitors to a nearby casino who were driving at the increased speeds and that city residents actually favored the lower speed limit. Explain how bias might be playing a role in the city council's decision.

1.6 THE DESIGN OF EXPERIMENTS

Objectives

1. Describe the characteristics of an experiment
2. Explain the steps in designing an experiment
3. Explain the completely randomized design
4. Explain the matched-pairs design
5. Explain the randomized block design

The major theme of this chapter has been data collection. Section 1.2 briefly discussed the idea of an experiment, but the main focus was on observational studies. Sections 1.3 through 1.5 focused on sampling and surveys. In this section, we further develop the idea of collecting data through an experiment.

1 Describe the Characteristics of an Experiment

Remember, in an observational study, if an association exists between an explanatory variable and response variable, the researcher cannot claim causality. If a researcher is interested in demonstrating how changes in the explanatory variable *cause* changes in the response variable, the researcher needs to conduct an *experiment*.

Definition

An **experiment** is a controlled study conducted to determine the effect varying one or more explanatory variables or **factors** has on a response variable. Any combination of the values of the factors is called a **treatment**.

Historical Note

Sir Ronald Fisher, often called the Father of Modern Statistics, was born in England on February 17, 1890. He received a BA in astronomy from Cambridge University in 1912. In 1914, he took a position teaching mathematics and physics at a high school. He did this to help serve his country during World War I. (He was rejected by the army because of his poor eyesight.) In 1919, Fisher took a job as a statistician at Rothamsted Experimental Station, where he was involved in agricultural research. In 1933, Fisher became Galton Professor of Eugenics at Cambridge University, where he studied Rh blood groups. In 1943 he was appointed to the Balfour Chair of Genetics at Cambridge. He was knighted by Queen Elizabeth in 1952. Fisher retired in 1957 and died in Adelaide, Australia, on July 29, 1962. One of his famous quotations is "To call in the statistician after the experiment is done may be no more than asking him to perform a postmortem examination: he may be able to say what the experiment died of."

In an experiment, the **experimental unit** is a person, object, or some other well-defined item upon which a treatment is applied. We often refer to the experimental unit as a **subject** when he or she is a person. The subject is analogous to the individual in a survey.

The overriding goal in an experiment is to determine the effect various treatments have on the response variable. For example, we might want to determine whether a new treatment is superior to an existing treatment (or no treatment at all). To make this determination, experiments require a *control group*. A **control group** serves as a baseline treatment that can be used to compare to other treatments. For example, a researcher in education might want to determine if students who do their homework using an online homework system do better on an exam than those who do their homework from the text. The students doing the text homework might serve as the control group (since this is the currently accepted practice). The factor is the type of homework. There are two treatments: online homework and text homework. A second method for defining the control group is through the use of a *placebo*. A **placebo** is an innocuous medication, such as a sugar tablet, that looks, tastes, and smells like the experimental medication.

In an experiment, it is important that each group be treated the same way. It is also important that individuals do not adjust their behavior in some way due to the treatment they are receiving. For this reason, many experiments use a technique called *blinding*. **Blinding** refers to nondisclosure of the treatment an experimental unit is receiving. There are two types of blinding: *single blinding* and *double blinding*.

Definitions

A **single-blind** experiment is one in which the experimental unit (or subject) does not know which treatment he or she is receiving. A **double-blind** experiment is one in which neither the experimental unit nor the researcher in contact with the experimental unit knows which treatment the experimental unit is receiving.

EXAMPLE 1 The Characteristics of an Experiment

Problem: Lipitor is a cholesterol-lowering drug by Pfizer. In the Collaborative Atorvastatin Diabetes Study (CARDS), the effect of Lipitor on cardiovascular disease was assessed in 2,838 subjects, ages 40 to 75, with type 2 diabetes, without prior history of cardiovascular disease. In this placebo-controlled, double-blind experiment, subjects were randomly allocated to either Lipitor 10 mg daily (1,429) or placebo (1,411) and were followed for 4 years. The response variable was the occurrence of any major cardiovascular event.

Lipitor significantly reduced the rate of major cardiovascular events (83 events in the Lipitor group versus 127 events in the placebo group). There were 61 deaths in the Lipitor group versus 82 deaths in the placebo group.

(a) What does it mean for the experiment to be placebo-controlled?
(b) What does it mean for the experiment to be double-blind?
(c) What is the population for which this study applies? What is the sample?
(d) What are the treatments?
(e) What is the response variable?

Approach: We will apply the definitions just presented.

Solution

(a) The placebo is a medication that looks, smells, and tastes like Lipitor. The purpose of the placebo control group is to serve as a baseline against which to compare the results from the group receiving Lipitor. Another reason for the placebo is to account for the fact that people tend to behave differently when they are in a study. By having a placebo control group, the effect of this is neutralized.

(b) Since the experiment is double-blind, the subjects do not know whether they are receiving Lipitor or the placebo. Plus, the individual monitoring the subjects does not know whether the subject is receiving Lipitor or the placebo. The reason we double-blind is so that the subjects receiving the medication do not behave differently from those receiving the placebo and the individual monitoring the subjects does not treat the folks in the Lipitor group differently from those in the placebo group.

(c) The population is individuals from 40 to 75 years of age with type 2 diabetes without a prior history of cardiovascular disease. The sample is the 2,838 subjects in the study.

(d) The treatments are 10 mg of Lipitor or a placebo daily.

(e) The response variable is whether the subject had any major cardiovascular event, such as a stroke, or not.

Now Work Problem 11

2 Explain the Steps in Designing an Experiment

To **design** an experiment means to describe the overall plan in conducting the experiment. The process of conducting an experiment requires a series of steps.

Step 1: *Identify the Problem to Be Solved.* The statement of the problem should be as explicit as possible. The statement should provide the experimenter with direction. In addition, the statement must identify the response variable and the population to be studied. Often, the statement is referred to as the *claim*.

Step 2: *Determine the Factors That Affect the Response Variable.* The factors are usually identified by an expert in the field of study. In identifying the factors, we must ask, "What things affect the value of the response variable?" Once the factors are identified, it must be determined which factors will be fixed at some predetermined level, which will be manipulated, and which will be uncontrolled.

Step 3: *Determine the Number of Experimental Units.* As a general rule, choose as many experimental units as time and money will allow. Techniques do exist for determining sample size, provided certain information is available. Some of these techniques are discussed later in the text.

Step 4: *Determine the Level of Each Factor.* There are two ways to deal with the factors:

1. **Control:** There are two ways to control the factors.

   (a) Fix their level at one predetermined value throughout the experiment. These are factors whose effect on the response variable is not of interest.

(b) Set them at predetermined levels. These are the factors whose effect on the response variable interests us. The combinations of the levels of these factors constitute the treatments in the experiment.

2. Randomize: Randomize the experimental units to various treatment groups so that the effect of factors whose levels cannot be controlled is minimized. The idea is that randomization averages out the effects of uncontrolled factors (explanatory variables). It is difficult, if not impossible, to identify all factors in an experiment. This is why randomization is so important. It mutes the effect of variation attributable to factors not controlled.

Step 5: *Conduct the Experiment.*

(a) The experimental units are randomly assigned to the treatments. **Replication** occurs when each treatment is applied to more than one experimental unit. By using more than one experimental unit for each treatment, we can be assured that the effect of a treatment is not due to some characteristic of a single experimental unit. It is a good idea to assign an equal number of experimental units to each treatment.

(b) Collect and process the data. Measure the value of the response variable for each replication. Then organize the results. The idea is that the value of the response variable for each treatment group is the same before the experiment because of randomization. Then any difference in the value of the response variable among the different treatment groups can be attributed to differences in the level of the treatment.

Step 6: *Test the Claim.* This is the subject of inferential statistics. **Inferential statistics** is a process in which generalizations about a population are made on the basis of results obtained from a sample. In addition, a statement regarding our level of confidence in our generalization is provided. We study methods of inferential statistics in Chapters 9 through 15.

3 Explain the Completely Randomized Design

The steps just given apply to any type of designed experiment. We now concentrate on the simplest type of experiment.

Definition

A **completely randomized design** is one in which each experimental unit is randomly assigned to a treatment.

We illustrate this type of experimental design using the steps just given.

EXAMPLE 2 A Completely Randomized Design

Problem: A farmer wishes to determine the optimal level of a new fertilizer on his soybean crop. Design an experiment that will assist him.

Approach: We follow the steps for designing an experiment.

Solution

Step 1: The farmer wants to identify the optimal level of fertilizer for growing soybeans. We define *optimal* as the level that maximizes yield. So the response variable will be crop yield.

Step 2: Some factors that affect crop yield are fertilizer, precipitation, sunlight, method of tilling the soil, type of soil, plant, and temperature.

Step 3: In this experiment, we will plant 60 soybean plants (experimental units).

In Other Words

The various levels of the factor are the treatments in a completely randomized design.

Step 4: We list the factors and their levels.

- **Fertilizer.** This factor will be set at three levels. We wish to measure the effect of varying the level of this variable on the response variable, yield. We will set the treatments (level of fertilizer) as follows:

  Treatment A: 20 soybean plants receive no fertilizer.

  Treatment B: 20 soybean plants receive 2 teaspoons of fertilizer per gallon of water every 2 weeks.

  Treatment C: 20 soybean plants receive 4 teaspoons of fertilizer per gallon of water every 2 weeks.

  See Figure 6.

Figure 6

- **Precipitation.** Although we cannot control the amount of rainfall, we can control the amount of watering we do. This factor will be controlled so that each plant receives the same amount of precipitation.
- **Sunlight.** This is an uncontrollable factor, but it will be roughly the same for each plant.
- **Method of tilling.** We can control this factor. We agree to use the round-up ready method of tilling for each plant.
- **Type of soil.** We can control certain aspects of the soil such as level of acidity. In addition, each plant will be planted within a 1-acre area, so it is reasonable to assume that the soil conditions for each plant are equivalent.
- **Plant.** There may be variation from plant to plant. To account for this, we randomly assign the plants to a treatment.
- **Temperature.** This factor is not within our control, but will be the same for each plant.

Step 5

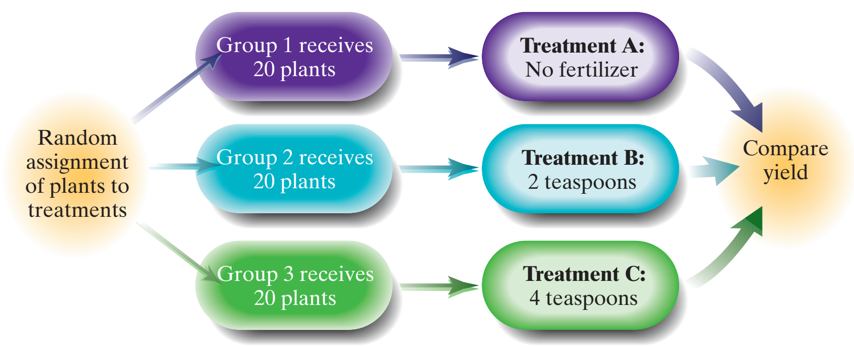
(a) We need to assign each plant to a treatment group. To do this, we will number the plants from 1 to 60. To determine which plants get treatment A, we randomly generate 20 numbers. The plants corresponding to these numbers get treatment A. Now number the remaining plants 1 to 40 and randomly generate 20 numbers. The plants corresponding to these numbers get treatment B. The remaining plants get treatment C. Now till the soil, plant the soybean plants, and fertilize according to the schedule prescribed.

(b) At the end of the growing season, determine the crop yield for each plant.

Step 6: Determine whether any differences in yield exist among the three treatment groups.

Figure 7 illustrates the experimental design.

Figure 7



Note to Instructor

It is useful to diagram the experimental designs in this chapter to target the understanding of the visual learner.

Example 2 is a completely randomized design because the experimental units (the plants) were randomly assigned to the treatments. It is the most popular experimental design because of its simplicity, but it is not always the best. We discuss inferential procedures for the completely randomized design in which there are two treatments in Section 11.2 and in which there are three or more treatments in Section 13.1.

Now Work Problem 13

4 Explain the Matched-Pairs Design

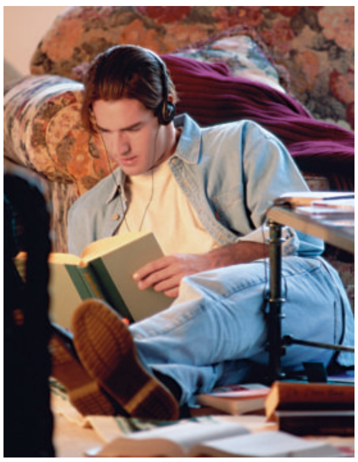
Another type of experimental design is called a *matched-pairs design*.

Definition

A **matched-pairs design** is an experimental design in which the experimental units are paired up. The pairs are matched up so that they are somehow related (that is, the same person before and after a treatment, twins, husband and wife, same geographical location, and so on). There are only two levels of treatment in a matched-pairs design.

In matched-pairs design, one matched individual will receive one treatment and the other matched individual receives a different treatment. The assignment of the matched pair to the treatment is done randomly using a coin flip or a random-number generator. We then look at the difference in the results of each matched pair. One common type of matched-pairs design is to measure a response variable on an experimental unit before a treatment is applied, and then to measure the response variable on the same experimental unit after the treatment is applied. In this way, the individual is matched against itself. These experiments are sometimes called before–after or pretest–posttest experiments.

EXAMPLE 3 A Matched-Pairs Design



Problem: An educational psychologist wanted to determine whether listening to music has an effect on a student's ability to learn. Design an experiment to help the psychologist answer the question.

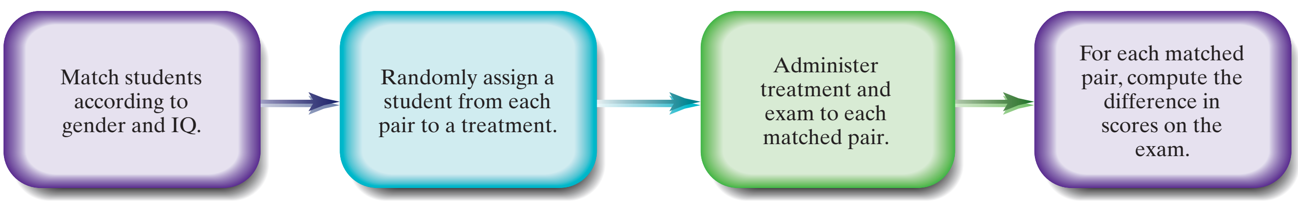
Approach: We will use a matched-pairs design by matching students according to IQ and gender (just in case gender plays a role in learning with music).

Solution: We match students according to IQ and gender. For example, a female with an IQ in the 110 to 115 range will be matched with a second female with an IQ in the 110 to 115 range.

For each pair of students, we will flip a coin to determine whether the first student in the pair is assigned the treatment of a quiet room or a room with music playing in the background.

Each student will be given a statistics textbook and asked to study Section 1.1. After 2 hours, the students will enter a testing center and take a short quiz on the material in the section. We compute the difference in the scores of each matched pair. Any differences in scores will be attributed to the treatment. Figure 8 illustrates the design.

Figure 8



Now Work Problem 15

We discuss statistical inference for the matched-pairs design in Section 11.1.

5 Explain the Randomized Block Design

The completely randomized design is the simplest experimental design. However, its simplicity can lead to flaws. Before we introduce a slightly more complicated experimental design, consider the following story.

I coach my son's soccer team, which is comprised of four 10-year-olds, six 9-year-olds, and four 8-year-olds. After each practice, I like to have a 15-minute scrimmage where I randomly assign seven players to each team. I quickly learned that randomly assigning players, without taking age into consideration, would sometimes result in very unequal teams (once I randomly created a team of four 10-year-olds and three 9-year-olds). So I learned that I should randomly assign two 10-year-olds to one team with the other two 10-year-olds going on the other team. I then assigned three 9-year-olds to each team and two 8-year-olds to each team. Using the language of statistics, I was *blocking* by age.

Definition

Grouping similar (homogeneous) experimental units together and then randomizing the experimental units within each group to a treatment is called **blocking**. Each group of homogeneous individuals is called a **block**.

With the soccer story in mind, let's revisit our crop-yield experiment from Example 2, where we assumed that the soybean plants were the same variety. Suppose we found out that we in fact had two varieties of soybeans: Chemgro and Pioneer. Using the completely randomized design, it is impossible to know whether any differences in yield were due to the differences in fertilizer or to the differences in variety.

It may be that the Chemgro variety naturally has higher yield potential than Pioneer. If a majority of the soybean plants that were randomly assigned to treatment C was Chemgro, we would not know whether to attribute the different crop yields to the level of fertilizer or to the plant variety. This effect is known as *confounding*. **Confounding** occurs when the effect of two factors (explanatory variables) on the response variable cannot be distinguished.

To resolve issues of this type, we introduce a third experimental design.

Definition

A **randomized block design** is used when the experimental units are divided into homogeneous groups called blocks. Within each block, the experimental units are randomly assigned to treatments.

In a randomized block design, we do not wish to determine whether the differences between blocks result in any difference in the value of the response variable. Our goal is to isolate any variability in the response variable that may be attributable to the block. We discuss this design in the next example.

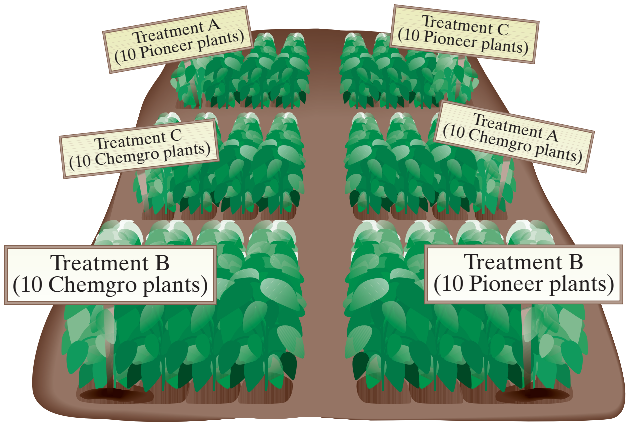
EXAMPLE 4 The Randomized Block Design

Problem: Suppose that the 60 soybean plants were actually two different varieties: 30 Chemgro soybean plants and 30 Pioneer soybean plants. In addition, the Chemgro variety may have a different yield potential than the Pioneer variety. Design an experiment that could be used to measure the effect of level of fertilizer while taking into account the variety of soybean.

Approach: Because the yield of Chemgro may differ from that of Pioneer, we will use a randomized block design to determine which level of fertilizer results in the highest yield. We will block by soybean variety.

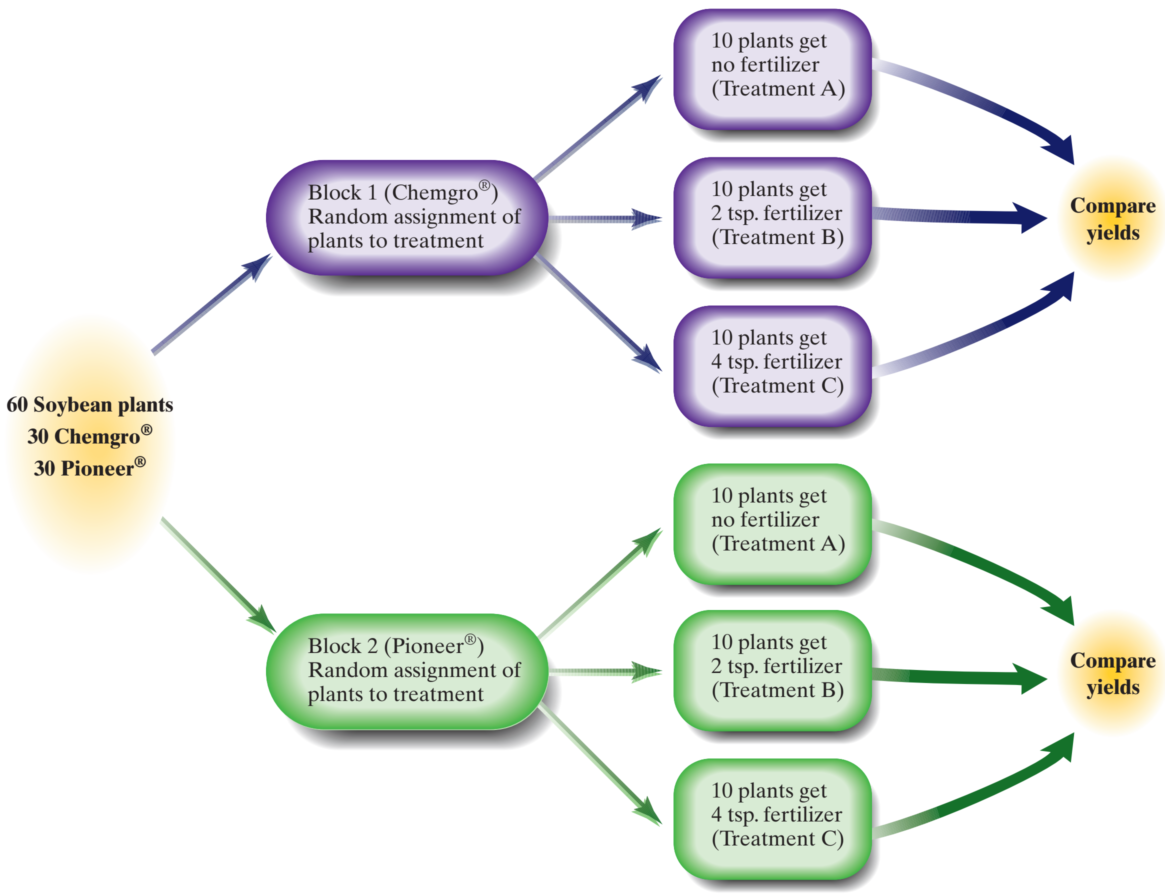
Solution: In a randomized block design, we divide 60 plants into two blocks. Block 1 will contain the 30 Chemgro plants and block 2 will contain the 30 Pioneer plants. Within each block, we randomly assign 10 plants to each treatment (A, B, or C). The other factors that may affect yield are controlled as they were in Example 2. See Figure 9.

Figure 9



The idea is that in each block the plants are all the same variety of plant, so plant variety does not affect the value of the response variable. Figure 10 illustrates the design. Notice that we do not compare yields across plant variety because this was not our goal—we already know which variety yields more. The reason for blocking is to reduce variability due to plant variety by comparing yields within each variety.

Figure 10



Now Work Problem 23

One note about the relation between a designed experiment and simple random sampling: It is often the case that the experimental units selected to participate in a study are not randomly selected. This is because we often need the experimental units to have some common trait, such as high blood pressure. For this reason, participants in experiments are recruited or volunteer to be in a study. However, once we have the experimental units, we use simple random sampling to assign them to treatment groups. With random assignment we assume that the participants are similar at the start of the experiment. Because the treatment is the only difference between the groups, we can say the treatment *caused* the difference observed in the response variable.

Note to Instructor

The purpose of this activity is to allow students to see why good experiments need control, randomization, and replication. In addition, students recognize how much thought needs to go into an experiment, even for something simple like paper frogs.

There are several possible designs for the frogs (e.g., www.frogsonice.com/froggy/origami/index.s/html or http://origami.ousaan.com/studio/traditional/frog/01/index.html), so choose whichever seems easiest for students to use.

Before the jumping, there should be discussions about why the order of the jumps is randomized, why the same person should jump all the frogs, and so on. In addition, students should decide how they will measure the distance jumped *before* they start jumping.

A good design might be a matched-pairs design so that the variability due to who is making the frog is eliminated and type of paper is not confounded with the person making the frog.

This activity may be done with paper airplanes as well.

IN CLASS ACTIVITY

Experimental Design (Hippity-Hop)

You are commissioned by the board of directors of Paper Toys, Inc. to design a new paper frog for their Christmas catalog. The design for the construction of the frog has already been completed and will be provided to you. However, the material with which to make the frogs has not yet been determined. The Materials Department has narrowed the choices down to either newspaper or brown paper (such as that used in grocery bags). You have decided to test both types of paper. Management decided to build the frogs from sheets of paper 9 inches square.

The goal of the experiment is to determine the material that results in frogs that jump farther.

(a) As a class, design an experiment that will answer the research question.

(b) Make the frogs.

(c) Conduct the experiment.

(d) As a class, discuss the strengths and weaknesses of the design. Would you change anything?

1.6 ASSESS YOUR UNDERSTANDING

Concepts and Vocabulary

1. Define the following:
(a) Experimental unit
(b) Treatment
(c) Response variable
(d) Factor
(e) Placebo
(f) Confounding

2. What is replication in an experiment?

3. Explain the difference between a single-blind and a double-blind experiment.

4. List the steps in designing an experiment.

5. A(n) ______ ______ design is one in which each experimental unit is randomly assigned to a treatment. A(n) ______ ______ design is one in which the experimental units are paired up. completely randomized; matched pairs

6. Grouping together similar experimental units and then randomly assigning the experimental units within each group to a treatment is called ______. blocking

7. *True or False*: Generally, the goal of an experiment is to determine the effect that treatments will have on the response variable. True

8. *True or False*: Observational studies can be used to determine causality between explanatory and response variables.

9. Discuss why control groups are needed in experiments.

10. Discuss how a randomized block design is similar to a stratified random sample. What is the purpose of blocking?

8. False

Applying the Concepts

11. Caffeinated Sports Drinks NW Researchers conducted a double-blind, placebo-controlled, repeated-measures experiment to compare the effectiveness of a commercial caffeinated carbohydrate–electrolyte sports drink with a commercial noncaffeinated carbohydrate–electrolyte sports drink and a flavored-water placebo. Sixteen highly trained cyclists each completed three trials of prolonged cycling in a warm environment: one while receiving the placebo, one while receiving the noncaffeinated sports drink, and one while receiving the caffeinated sports drink. For a given trial, one beverage treatment was administered throughout a 2-hour variable-intensity cycling bout followed by a 15-minute performance ride. Total work in kilojoules (kJ) performed during the final 15 minutes was used to measure performance. The beverage order for the individual subjects was randomly assigned. A period of at least 5 days separated the trials. All trials took place at approximately the same time of day in an environmental chamber at 28.5°C and 60% relative humidity with fan airflow of approximately 2.5 meters per second (m/s).

The researchers found that cycling performance, as assessed by the total work completed during the performance ride, was 23% greater for the caffeinated sports drink than for the placebo and 15% greater for the caffeinated sports drink than for the noncaffeinated sports drink. Cycling performances for the noncaffeinated sports drink and the placebo were not significantly different. The researchers

11. (e) Caffeinated sports drink, noncaffeinated sports drink, or flavored-water placebo

concluded that the caffeinated carbohydrate–electrolyte sports drink substantially enhanced physical performance during prolonged exercise compared with the noncaffeinated carbohydrate–electrolyte sports drink and the placebo.

Source: Kirk J. Cureton, Gordon L. Warren et al. "Caffeinated Sports Drink: Ergogenic Effects and Possible Mechanisms," *International Journal of Sport Nutrition and Exercise Metabolism*, 17(1):35–55, 2007

(a) What does it mean for the experiment to be placebo-controlled?
(b) What does it mean for the experiment to be double-blind? Why do you think it is necessary for the experiment to be double-blind?
(c) How is randomization used in this experiment?
(d) What is the population for which this study applies? What is the sample? All athletes; 16 cyclists
(e) What are the treatments?
(f) What is the response variable? Total work
(g) This experiment used a *repeated-measures design*, a design type that has not been directly discussed in this textbook. Using this experiment as a guide, determine what it means for the design of the experiment to be repeated-measures. How does this design relate to the matched-pairs design?

12. Alcohol Dependence To determine if topiramate is a safe and effective treatment for alcohol dependence, researchers conducted a 14-week trial of 371 men and women aged 18 to 65 years diagnosed with alcohol dependence. In this double-blind, randomized, placebo-controlled experiment, subjects were randomly given either 300 milligrams (mg) of topiramate (183 subjects) or a placebo (188 subjects) daily, along with a weekly compliance enhancement intervention. The variable used to determine the effectiveness of the treatment was self-reported percentage of heavy drinking days. Results indicated that topiramate was more effective than placebo at reducing the percentage of heavy drinking days. The researchers concluded that topiramate is a promising treatment for alcohol dependence.

Source: Bankole A. Johnson, Norman Rosenthal, et al. "Topiramate for Treating Alcohol Dependence: A Randomized Controlled Trial," *Journal of the American Medical Association*, 298(14):1641–1651, 2007

(a) What does it mean for the experiment to be placebo-controlled?
(b) What does it mean for the experiment to be double-blind? Why do you think it is necessary for the experiment to be double-blind?
(c) What does it mean for the experiment to be randomized?
(d) What is the population for which this study applies? What is the sample?
(e) What are the treatments?
(f) What is the response variable? Percentage of heavy drinking days

13. School Psychology NW A school psychologist wants to test the effectiveness of a new method for teaching reading. She recruits 500 first-grade students in District 203 and randomly divides them into two groups. Group 1 is taught by means of the new method, while group 2 is taught via traditional methods. The same teacher is assigned to teach both groups. At the end of the year, an achievement test is administered and the results of the two groups are compared.

(a) What is the response variable in this experiment?
(b) Think of some of the factors in the study. How are they controlled?
(c) What are the treatments? How many treatments are there? New teaching method and traditional method; 2
(d) How are the factors that are not controlled dealt with? Random assignment
(e) Which group serves as the control group? Group 2
(f) What type of experimental design is this?
(g) Identify the subjects. 500 students
(h) Draw a diagram similar to Figure 7, 8, or 10 to illustrate the design.

14. Pharmacy A pharmaceutical company has developed an experimental drug meant to relieve symptoms associated with the common cold. The company identifies 300 adult males 25 to 29 years old who have a common cold and randomly divides them into two groups. Group 1 is given the experimental drug, while group 2 is given a placebo. After 1 week of treatment, the proportions of each group that still have cold symptoms are compared. 14. (a) Proportion of subjects with cold

(a) What is the response variable in this experiment?
(b) Think of some of the factors in the study. How are they controlled?
(c) What are the treatments? How many treatments are there? Experimental drug and placebo; 2
(d) How are the factors that are not controlled dealt with? Random assignment
(e) What type of experimental design is this?
(f) Identify the subjects. 300 males
(g) Draw a diagram similar to Figure 7, 8, or 10 to illustrate the design.

15. Whiter Teeth NW An ad for Crest Whitestrips Premium claims that the strips will whiten teeth in 7 days and the results will last for 12 months. A researcher who wishes to test this claim studies 20 sets of identical twins. Within each set of twins, one is randomly selected to use Crest Whitestrips Premium in addition to regular brushing and flossing, while the other just brushes and flosses. Whiteness of teeth is measured at the beginning of the study, after 7 days, and every month thereafter for 12 months. 15. (b) Whiteness level

(a) What type of experimental design is this? Matched pairs
(b) What is the response variable in this experiment?
(c) What are the treatments?
(d) What are other factors (controlled or uncontrolled) that could affect the response variable?
(e) What might be an advantage of using identical twins as subjects in this experiment?

16. Assessment To help assess student learning in her developmental math courses, a mathematics professor at a community college implemented pre- and posttests for her developmental math students. A knowledge-gained score was obtained by taking the difference of the two test scores.

12. (d) All 18–65 year olds with alcohol dependence; 371 men and women aged 18 to 65 years of age with alcohol dependence
12. (e) 300 mg of topiramate or a placebo daily 13. (a) Score on achievement test 13. (b) Method of teaching, grade level, intelligence, school district, teacher; Fixed: grade level, school district, teacher; Set to a predetermined level: method of teaching
13. (f) Completely randomized design 14. (b) Gender, age, geographic location, overall health, drug intervention; Fixed: gender, age, location; Set to a predetermined level: drug intervention 14. (e) Completely randomized design 15. (c) Whitening method; Crest Whitestrips Premium with brushing and flossing versus brushing and flossing alone

16. (b) Difference in test scores 17. (b) Adults with insomnia
17. (c) Terminal wake time after sleep onset (WASO) 17. (d) CBT, progressive muscle relaxation training, placebo

(a) What type of experimental design is this? Matched pair
(b) What is the response variable in this experiment?
(c) What is the treatment? Math course

17. Insomnia Researchers Jack D. Edinger and associates wanted to test the effectiveness of a new cognitive behavioral therapy (CBT) compared with both an older behavioral treatment and a placebo therapy for treating insomnia. They identified 75 adults with chronic insomnia. Patients were randomly assigned to one of three treatment groups. Twenty-five patients were randomly assigned to receive CBT (sleep education, stimulus control, and time-in-bed restrictions), another 25 received muscle relaxation training (RT), and the final 25 received a placebo treatment. Treatment lasted 6 weeks, with follow-up conducted at 6 months. To measure the effectiveness of the treatment, researchers used wake time after sleep onset (WASO). Cognitive behavioral therapy produced larger improvements than did RT or placebo treatment. For example, the CBT-treated patients achieved an average 54% reduction in their WASO, whereas RT-treated and placebo-treated patients, respectively, achieved only 16% and 12% reductions in this measure. Results suggest that CBT treatment leads to significant sleep improvements within 6 weeks, and these improvements appear to endure through 6 months of follow-up. 17. (a) Completely randomized design

Source: Jack D. Edinger, PhD; William K. Wohlgemuth, PhD; Rodney A. Radtke, MD; Gail R. Marsh, PhD; Ruth E. Quillian, PhD. "Cognitive Behavioral Therapy for Treatment of Chronic Primary Insomnia," *Journal of the American Medical Association*, 285:1856–1864, 2001

(a) What type of experimental design is this?
(b) What is the population being studied?
(c) What is the response variable in this study?
(d) What are the treatments?
(e) Identify the experimental units. 75 adults with insomnia
(f) Draw a diagram similar to Figure 7, 8, or 10 to illustrate the design.

18. Depression Researchers wanted to compare the effectiveness and safety of an extract of St. John's wort with placebo in outpatients with major depression. To do this, they recruited 200 adult outpatients diagnosed as having major depression and having a baseline Hamilton Rating Scale for Depression (HAM-D) score of at least 20. Participants were randomly assigned to receive either St. John's wort extract, 900 milligrams per day (mg/d) for 4 weeks, increased to 1200 mg/d in the absence of an adequate response thereafter, or a placebo for 8 weeks. The response variable was the change on the HAM-D over the treatment period. After analysis of the data, it was concluded that St. John's wort was not effective for treatment of major depression.

Source: Richard C. Shelton, MD, et al. "Effectiveness of St. John's Wort in Major Depression," *Journal of the American Medical Association* 285:1978–1986, 2001

(a) What type of experimental design is this?
(b) What is the population that is being studied?
(c) What is the response variable in this study?
(d) What are the treatments?
(e) Identify the experimental units. 200 outpatients
(f) What is the control group in this study? Placebo group
(g) Draw a diagram similar to Figure 7, 8, or 10 to illustrate the design.

19. The Memory Drug? Researchers wanted to evaluate whether ginkgo, an over-the-counter herb marketed as enhancing memory, improves memory in elderly adults as measured by objective tests. To do this, they recruited 98 men and 132 women older than 60 years and in good health. Participants were randomly assigned to receive ginkgo, 40 milligrams (mg) 3 times per day, or a matching placebo. The measure of memory improvement was determined by a standardized test of learning and memory. After 6 weeks of treatment, the data indicated that ginkgo did not increase performance on standard tests of learning, memory, attention, and concentration. These data suggest that, when taken following the manufacturer's instructions, ginkgo provides no measurable increase in memory or related cognitive function to adults with healthy cognitive function.

Source: Paul R. Solomon et al. "Ginkgo for Memory Enhancement," *Journal of the American Medical Association* 288:835–840, 2002

(a) What type of experimental design is this?
(b) What is the population being studied?
(c) What is the response variable in this study?
(d) What is the factor that is set to predetermined levels? What are the treatments?
(e) Identify the experimental units.
(f) What is the control group in this study? Placebo group
(g) Draw a diagram similar to Figure 7, 8, or 10 to illustrate the design.

20. Treating Depression Researchers wanted to test whether a new drug therapy results in a more rapid response in patients with major depression. To do this, they recruited 63 inpatients with a diagnosis of major depression. Patients were randomly assigned to two treatment groups receiving either placebo (31 patients) or the new drug therapy (32 patients). The response variable was the Hamilton Rating Scale for Depression score. After collecting and analyzing the data, it was concluded that the new drug therapy is effective in the treatment of major depression.

Source: Jahn Holger, MD, et al. "Metyrapone as Additive Treatment in Major Depression," *Archives of General Psychiatry*, 61:1235–1244, 2004

(a) What type of experimental design is this?
(b) What is the population that is being studied?
(c) What is the response variable in this study?
(d) What are the treatments? Placebo or new drug therapy
(e) Identify the experimental units.
(f) Draw a diagram similar to Figure 7, 8, or 10 to illustrate the design.

21. Dominant Hand Professor Andy Neill wanted to determine if the reaction time of people differs in their dominant hand versus their nondominant hand. To do this, he recruited

18. (a) Completely randomized design 18. (b) Adult outpatients diagnosed as having major depression and having a baseline Hamilton Rating Scale for Depression (HAM-D) score of at least 20 18. (c) Change in the HAM-D over the treatment period 18. (d) Type of drug: St. John's wort extract or placebo 19. (a) Completely randomized design 19. (b) Adults older than 60 years in good health 19. (c) Standardized test of learning and memory 19. (d) Drug; 40 mg of ginkgo 3 times per day or a matching placebo 19. (e) 98 men and 132 women older than 60 in good health 20. (a) Completely randomized design 20. (b) Inpatients with a diagnosis of major depression 20. (c) Hamilton Rating Scale for Depression score 20. (e) 63 inpatients with a diagnosis of major depression

21. (a) Matched-pairs design

15 students. Each student was asked to hold a yardstick between the index finger and thumb. The student was asked to open the hand, release the yardstick, and then asked to catch the yardstick between the index finger and thumb. The distance that the yardstick fell served as a measure of reaction time. A coin flip was used to determine whether the student would use their dominant hand first or the nondominant hand. Results indicated that the reaction time in the dominant hand exceeded that of the nondominant hand.

(a) What type of experimental design is this?
(b) What is the response variable in this study?
(c) What is the treatment? Dominant versus nondominant hand
(d) Identify the experimental units. 15 students
(e) Why did Professor Neill use a coin flip to determine whether the student should begin with the dominant hand or the nondominant hand?
(f) Draw a diagram similar to Figure 7, 8, or 10 to illustrate the design.

22. Golf Anyone? A local golf pro wanted to compare two styles of golf club. One golf club had a graphite shaft and the other had the latest style of steel shaft. It is a common belief that graphite shafts allow a player to hit the ball farther, but the manufacturer of the new steel shaft said the ball travels just as far with its new technology. To test this belief, the pro recruited 10 golfers from the driving range. Each player was asked to hit one ball with the graphite-shafted club and one ball with the new steel-shafted club. The distance that the ball traveled was determined using a range finder. A coin flip was used to determine whether the player hit with the graphite club or the steel club first. Results indicated that the distance the ball was hit with the graphite club was no different than the distance when using the steel club.

(a) What type of experimental design is this?
(b) What is the response variable in this study?
(c) What is the factor that is set to predetermined levels? What is the treatment?
(d) Identify the experimental units. 10 golfers
(e) Why did the golf pro use a coin flip to determine whether the golfer should hit with the graphite first or the steel first?
(f) Draw a diagram similar to Figure 7, 8, or 10 to illustrate the design.

23. Marketing NW A marketing research firm wishes to determine the most effective method of advertising: print, radio, or television. They recruit 300 volunteers to participate in the study. The chief researcher believes that level of education plays a role in the effectiveness of advertising, so she segments the volunteers by level of education. Of the 300 volunteers, 120 have a high school education, 120 have a college diploma, and 60 have advanced degrees. The 120 volunteers with a high school diploma are randomly assigned to either the print advertising group, the radio group, or the television group. The same procedure is followed for the college graduate and advanced degree volunteers. Each group is exposed to the advertising. After 1 hour, a recall exam is given and the proportion of correct answers recorded.

(a) What type of experimental design is this?
(b) What is the response variable in this experiment?
(c) What are the treatments? Print, radio, television
(d) What variable serves as the block? Level of education
(e) Draw a diagram similar to Figure 7, 8, or 10 to illustrate the design.

24. Social Work A social worker wants to examine methods that can be used to deter truancy. Three hundred chronically truant students from District 103 volunteer for the study. Because the social worker believes that socioeconomic class plays a role in truancy, she divides the 300 volunteers according to household income. Of the 300 students, 120 fall in the low-income category, 132 fall in the middle-income category, and the remaining 48 fall in the upper-income category. The students within each income category are randomly divided into three groups. The students in group 1 receive no intervention. The students in group 2 are treated with positive reinforcement in which, for each day the student is not truant, he or she receives a star that can be traded in for rewards. The students in group 3 are treated with negative reinforcement such that each truancy results in a 1-hour detention. However, the hours of detention are cumulative, meaning that the first truancy results in 1 hour of detention, the second truancy results in 2 hours, and so on. After a full school year, the total number of truancies are compared.

(a) What type of experimental design is this?
(b) What is the response variable in this experiment?
(c) What are the treatments?
(d) What variable serves as the block? Income
(e) Draw a diagram similar to Figure 7, 8, or 10 to illustrate the design.

25. Drug Effectiveness A pharmaceutical company wants to test the effectiveness of an experimental drug meant to reduce high cholesterol. The researcher at the pharmaceutical company has decided to test the effectiveness of the drug through a completely randomized design. She has obtained 20 volunteers with high cholesterol: Ann, John, Michael, Kevin, Marissa, Christina, Eddie, Shannon, Julia, Randy, Sue, Tom, Wanda, Roger, Laurie, Rick, Kim, Joe, Colleen, and Bill. Number the volunteers from 1 to 20. Use a random-number generator to randomly assign 10 of the volunteers to the experimental group. The remaining volunteers will go into the control group. List the individuals in each group.

26. Effects of Alcohol A researcher has recruited 20 volunteers to participate in a study. The researcher wishes to measure the effect of alcohol on an individual's reaction time. The 20 volunteers are randomly divided into two groups. Group 1 will serve as a control group in which participants drink four 1-ounce cups of a liquid that looks, smells, and tastes like alcohol in 15-minute increments. Group 2 will serve as an experimental group in which participants drink four 1-ounce cups of 80-proof alcohol in 15-minute increments. After drinking the last 1-ounce cup, the participants sit for 20 minutes. After the 20-minute resting period, the reaction time to a stimulus is measured.

(a) What type of experimental design is this?
(b) Use Table I in Appendix A or a random-number generator to divide the 20 volunteers into groups 1 and 2 by assigning the volunteers a number between 1 and 20. Then randomly select 10 numbers between 1 and 20. The individuals corresponding to these numbers will go into group 1.

21. (b) Distance yardstick falls 21. (e) To eliminate bias due to starting on dominant or nondominant first for each trial 22. (a) Matched-pairs design 22. (b) Distance ball hit 22. (c) Shaft type: graphite shaft versus steel shaft 22. (e) To eliminate bias due to type of shaft used first 23. (a) Randomized block design 23. (b) Score on recall exam 24. (a) Randomized block design 24. (b) Total number of truancies 24. (c) No intervention, positive reinforcement, negative reinforcement 26. (a) Completely randomized design

29. Matched-pair design (match by type of exterior finish)
30. Matched-pair design (match by car model)

27. Tomatoes An oncologist wants to perform a long-term study on the benefits of eating tomatoes. In particular, she wishes to determine whether there is a significant difference in the rate of prostate cancer among adult males after eating one serving of tomatoes per week for 5 years, after eating three servings of tomatoes per week for 5 years, and after eating five servings of tomatoes per week for 5 years. Help the oncologist design the experiment. Include a diagram to illustrate your design. Completely randomized design

28. Batteries An engineer wants to determine the effect of temperature on battery voltage. In particular, he is interested in determining if there is a significant difference in the voltage of the batteries when exposed to temperatures of 90°F, 70°F, and 50°F. Help the engineer design the experiment. Include a diagram to illustrate your design. Completely randomized design

29. The Better Paint Suppose you are interested in comparing Benjamin Moore's MoorLife Latex house paint with Sherwin Williams' LowTemp 35 Exterior Latex paint. Design an experiment that will answer this question: Which paint is better for painting the exterior of a house? Include a diagram to illustrate your design.

30. Tire Design An engineer has just developed a new tire design. However, before going into production, the tire company wants to determine if the new tire reduces braking distance on a car traveling 60 miles per hour compared with radial tires. Design an experiment to help the engineer determine if the new tire reduces braking distance.

31. Octane Does the level of octane in gasoline affect gas mileage? To answer this question, an automotive engineer obtains 60 cars. Twenty of the cars are considered to be compact, 20 are full size, and 20 are sport utility vehicles (SUVs). Design an experiment for the engineer. Include a diagram to illustrate your design. Randomized block design

32. School Psychology The school psychologist first presented in Problem 13 worries that girls and boys may react differently to the two methods of instruction. Design an experiment that will eliminate the variability due to gender on the response variable, the score on the achievement test. Include a diagram to illustrate your design. Randomized block design

33. Designing an Experiment Researchers wish to know if there is a link between hypertension (high blood pressure) and consumption of salt. Past studies have indicated that the consumption of fruits and vegetables offsets the negative impact of salt consumption. It is also known that there is quite a bit of person-to-person variability as far as the ability of the body to process and eliminate salt. However, no method exists for identifying individuals who have a higher ability to process salt. The U.S. Department of Agriculture recommends that daily intake of salt should not exceed 2400 milligrams (mg). The researchers want to keep the design simple, so they choose to conduct their study using a completely randomized design.

(a) What is the response variable in the study? Blood pressure
(b) Name three factors that have been identified.
(c) For each factor identified, determine whether the variable can be controlled or cannot be controlled. If a factor cannot be controlled, what should be done to reduce variability in the response variable?
(d) How many treatments would you recommend? Why?

34. Search a newspaper, magazine, or other periodical that describes an experiment. Identify the population, experimental unit, response variable, treatment, factors, and their levels.

35. Research the *placebo effect* and the *Hawthorne effect*. Write a paragraph that describes how each affects the outcome of an experiment.

36. Coke or Pepsi Suppose you want to perform an experiment whose goal is to determine whether people prefer Coke or Pepsi. Design an experiment that utilizes the completely randomized design. Design an experiment that utilizes the matched-pairs design. In both designs, be sure to identify the response variable, the role of blinding, and randomization. Which design do you prefer? Why?

37. Putting It Together: Mosquito Control In an attempt to identify ecologically friendly methods for controlling mosquito populations, researchers conducted field experiments in India where aquatic nymphs of the dragonfly *Brachytron pretense* were used against the larvae of mosquitoes. For the experiment, the researchers selected ten 300-liter (L) outdoor, open, concrete water tanks, which were natural breeding places for mosquitoes. Each tank was manually sieved to ensure that it was free of any nonmosquito larvae, nymphs, or fish. Only larvae of mosquitoes were allowed to remain in the tanks. The larval density in each tank was assessed using a 250-milliliter (mL) dipper. For each tank, 30 dips were taken and the mean larval density per dip was calculated. Ten freshly collected nymphs of *Brachytron pretense* were introduced into each of five randomly selected tanks. No nymphs were released into the remaining five tanks, which served as controls. After 15 days, larval densities in all the tanks were assessed again and all the introduced nymphs were removed. After another 15 days, the larval densities in all the tanks were assessed a third time.

In the nymph-treated tanks, the density of larval mosquitoes dropped significantly from 7.34 to 0.83 larvae per dip 15 days after the *Brachytron pretense* nymphs were introduced. Further, the larval density increased significantly to 6.83 larvae per dip 15 days after the nymphs were removed. Over the same time period, the control tanks did not show a significant difference in larval density, with density measurements of 7.12, 6.83, and 6.79 larvae per dip. The researchers concluded that *Brachytron pretense* can be used effectively as a strong, ecologically friendly control of mosquitoes and mosquito borne diseases.

Source: S. N. Chatterjee, A. Ghosh, and G. Chandra. "Eco-Friendly Control of Mosquito Larvae by *Brachytron pretense* Nymph," *Journal of Environmental Health*, 69(8):44–48, 2007

(a) Identify the research objective.
(b) What type of experimental design is this?
(c) What is the response variable? It is quantitative or qualitative? If quantitative, is it discrete or continuous?
(d) What is the factor the researchers controlled and set to predetermined levels? What are the treatments?

33. (b) Daily consumption of salt, daily consumption of fruits and vegetables, body's ability to process salt 33. (c) Salt: controlled; Fruits and vegetables: controlled; Body: cannot be controlled. To deal with variability in body types, randomize experimental units to each treatment group. 33. (d) Answers will vary, but three might be a good choice: one below RDA, one equal to RDA, one above RDA.
37. (a) To determine if nymphs of *Brachytron pretense* will control mosquitoes 37. (b) Completely randomized design
37. (c) Mosquito larvae density; quantitative; discrete

37. (d) Introduction of nymphs of *Brachytron pretense* into water tanks containing mosquito larvae; nymphs added or no nymphs added
37. (e) Temperature, amount of rainfall, fish, other larvae, sunlight

(e) Can you think of other factors that may affect larvae of mosquitoes? How are they controlled or dealt with?
(f) What is the population for which this study applies? What is the sample?
(g) List the descriptive statistics.
(h) How did the researchers control this experiment?
(i) Draw a diagram similar to Figure 7, 8, or 10 to illustrate the design.
(j) State the conclusion made in the study.

37. (f) All mosquito larvae in all breeding places; the mosquito larvae in the ten 300-liter outdoor, open concrete water tanks
37. (g) During the study period, mosquito larvae density changed from 7.34 to 0.83 to 6.83 larvae per dip in the treatment tanks, while changing only from 7.12 to 6.83 to 6.79 larvae per dip in the control tanks.

Consumer Reports® Emotional "Aspirin"

Americans have a long history of altering their moods with chemicals, ranging from alcohol and illicit drugs to prescription medications, such as diazepam (Valium) for anxiety and fluoxetine (Prozac) for depression. Today, there's a new trend: the over-the-counter availability of apparently effective mood modifiers in the form of herbs and other dietary supplements.

One problem is that many people who are treating themselves with these remedies may be sufficiently anxious or depressed to require professional care and monitoring. Self-treatment can be dangerous, particularly with depression, which causes some 20,000 reported suicides a year in the United States. Another major pitfall is that dietary supplements are largely unregulated by the government, so consumers have almost no protection against substandard preparations.

To help consumers and doctors, *Consumer Reports* tested the amounts of key ingredients in representative brands of several major mood-changing pills. To avoid potential bias, we tested samples from different lots of the pills using a randomized statistical design. The table contains a subset of the data from this study.

Each of these pills has a label claim of 200 mg of SAM-E. The column labeled Random Code contains a set of 3-digit random codes that were used so that the laboratory did not know which manufacturer was being tested. The column labeled Mg SAM-E contains the amount of SAM-E measured by the laboratory.

(a) Why is it important to label the pills with random codes?

(b) Why is it important to randomize the order in which the pills are tested instead of testing all of brand A first, followed by all of brand B, and so on?

Run Order	Brand	Random Code	Mg SAM-E
1	B	461	238.9
2	D	992	219.2
3	C	962	227.1
4	A	305	231.2
5	B	835	263.7
6	D	717	251.1
7	A	206	232.9
8	D	649	192.8
9	C	132	213.4
10	B	923	224.6
11	A	823	261.1
12	C	515	207.8

(c) Sort the data by brand. Does it appear that each brand is meeting its label claims?

(d) Design an experiment that follows the steps presented to answer the following research question: "Is there a difference in the amount of SAM-E contained in brands A, B, C, and D?"

Note to Readers: *In many cases, our test protocol and analytical methods are more complicated than described in this example. The data and discussion have been modified to make the material more appropriate for the audience.*

37. (j) Nymphs of *Brachytron pretense* can be used effectively as a strong, ecologically friendly control of mosquitoes and mosquito-borne diseases.

CHAPTER 1 REVIEW

Summary

We defined statistics as a science in which data are collected, organized, summarized, and analyzed to infer characteristics regarding a population. Statistics also provides a measure of confidence in the conclusions that are drawn. Descriptive statistics consists of organizing and summarizing information, while inferential statistics consists of drawing conclusions about a population based on results obtained from a sample. The population is a collection of individuals about which information is desired and the sample is a subset of the population.

Data are the observations of a variable. Data can be either qualitative or quantitative. Quantitative data are either discrete or continuous.

Data can be obtained from four sources: a census, existing sources, observational studies, or a designed experiment. A census will list all the individuals in the population, along with certain characteristics. Due to the cost of obtaining a census, most researchers opt for obtaining a sample. In observational studies, the response variable is measured without attempting to influence its value. In addition, the explanatory variable is not manipulated. Designed experiments are used when control of the individuals in the study is desired to isolate the effect of a certain treatment on a response variable.

We introduced five sampling methods: simple random sampling, stratified sampling, systematic sampling, cluster sampling, and convenience sampling. All the sampling methods, except for convenience sampling, allow for unbiased statistical inference to be made. Convenience sampling typically leads to an unrepresentative sample and biased results.

Vocabulary

Be sure you can define the following:

Statistics (p. 3)
Data (pp. 3, 9)
Population (p. 5)
Individual (p. 5)
Sample (p. 5)
Descriptive statistics (p. 5)
Statistic (p. 5)
Inferential statistics (pp. 5, 48)
Parameter (p. 5)
Variable (p. 7)
Qualitative variable (p. 7)
Quantitative variable (p. 7)
Discrete variable (p. 8)
Continuous variable (p. 8)
Qualitative data (p. 9)
Quantitative data (p. 9)
Discrete data (p. 9)
Continuous data (p. 9)
Nominal level of measurement (p. 10)
Ordinal level of measurement (p. 10)
Interval level of measurement (p. 10)
Ratio level of measurement (p. 10)
Validity (p. 11)
Reliability (p. 11)
Explanatory variable (p. 16)
Response variable (p. 16)
Observational study (p. 16)
Designed experiment (p. 16)
Confounding (pp. 17, 51)
Lurking variable (p. 17)
Retrospective (p. 18)
Prospective (p. 19)
Census (p. 19)
Random sampling (p. 23)
Simple random sampling (p. 23)
Simple random sample (p. 23)
Frame (p. 24)
Sampling without replacement (p. 24)
Sampling with replacement (p. 24)
Seed (p. 26)
Stratified sample (p. 30)
Systematic sample (p. 31)
Cluster sample (p. 33)
Convenience sample (p. 34)
Self-selected (p. 34)
Voluntary response (p. 34)
Bias (p. 38)
Sampling bias (p. 38)
Undercoverage (p. 38)
Nonresponse bias (p. 39)
Response bias (p. 40)
Open question (p. 41)
Closed question (p. 41)
Nonsampling error (p. 42)
Sampling error (p. 42)
Experiment (p. 46)
Factors (p. 46)
Treatment (p. 46)
Experimental unit (p. 46)
Subject (p. 46)
Control group (p. 46)
Placebo (p. 46)
Blinding (p. 46)
Single-blind (p. 46)
Double-blind (p. 46)
Design (p. 47)
Replication (p. 48)
Completely randomized design (p. 48)
Matched-pairs design (p. 50)
Blocking (p. 51)
Block (p. 51)
Randomized block design (p. 51)

Objectives

Section	You should be able to . . .	Example(s)	Review Exercises
1.1	1 Define statistics and statistical thinking (p. 3)	pp. 3–4	1
	2 Explain the process of statistics (p. 4)	1, 2	7, 14, 15
	3 Distinguish between qualitative and quantitative variables (p. 7)	3	11–13
	4 Distinguish between discrete and continuous variables (p. 8)	4, 5	11–13
	5 Determine the level of measurement of a variable (p. 10)	6	16–19
1.2	1 Distinguish between an observational study and an experiment (p. 15)	1–3	20–21
	2 Explain the various types of observational studies (p. 18)	pp. 18–19	6, 22
1.3	1 Obtain a simple random sample (p. 23)	1–3	28, 30
1.4	1 Obtain a stratified sample (p. 30)	1	25
	2 Obtain a systematic sample (p. 31)	2	26, 29
	3 Obtain a cluster sample (p. 32)	3	24
1.5	1 Explain the sources of bias in sampling (p. 38)	pp. 38–42	8, 9, 27
1.6	1 Describe the characteristics of an experiment (p. 46)	1	5
	2 Explain the steps in designing an experiment (p. 47)	pp. 47–48	10
	3 Explain the completely randomized design (p. 48)	2	31, 34(a), 35
	4 Explain the matched-pairs design (p. 50)	3	35
	5 Explain the randomized block design (p. 51)	4	32, 34(b)

Review Exercises

In Problems 1–5, provide a definition using your own words.

1. Statistics
2. Population
3. Sample
4. Observational study
5. Designed experiment
6. List and describe the three major types of observational studies.
7. What is meant by the *process of statistics*?
8. List and explain the three sources of bias in sampling. Provide some methods that might be used to minimize bias in sampling.
9. Distinguish between sampling and nonsampling error.
10. Explain the steps in designing an experiment.

In Problems 11–13, classify the variable as qualitative or quantitative. If the variable is quantitative, state whether it is discrete or continuous.

11. Number of new automobiles sold at a dealership on a given day Quantitative; discrete

12. Weight in carats of an uncut diamond Quantitative; continuous

13. Brand name of a pair of running shoes Qualitative

In Problems 14 and 15, determine whether the underlined value is a parameter or a statistic.

14. In a survey of 1011 people age 50 or older, 73% agreed with the statement "I believe in life after death." Statistic

Source: Bill Newcott. "Life after Death," *AARP Magazine*, Sept./Oct. 2007

15. Completion Rate In the 2007 NCAA Football Championship Game, quarterback Chris Leak completed 69% of his passes for a total of 213 yards and 1 touchdown. Parameter

In Problems 16–19, determine the level of measurement of each variable.

16. Birth year Interval

17. Marital status Nominal

18. Stock rating (strong buy, buy, hold, sell, strong sell) Ordinal

19. Number of siblings Ratio

In Problems 20 and 21, determine whether the study depicts an observational study or a designed experiment.

20. A parent group examines 25 randomly selected PG-13 movies and 25 randomly selected PG movies and records the number of sexual innuendos and curse words that occur in each. They then compare the number of sexual innuendos and curse words between the two movie ratings. Observational study

21. A sample of 504 patients in early stages of Alzheimer's disease is divided into two groups. One group receives an experimental drug; the other receives a placebo. The advance of the disease in the patients from the two groups is tracked at 1-month intervals over the next year. Experiment

22. Read the following description of an observational study and determine whether it is a cross-sectional, a case-control, or a cohort study. Explain your choice.

The Cancer Prevention Study II (CPS-II) examines the relationship among environmental and lifestyle factors of cancer cases by tracking approximately 1.2 million men and women. Study participants completed an initial study questionnaire in 1982 providing information on a range of lifestyle factors, such as diet, alcohol and tobacco use, occupation, medical history, and family cancer history. These data have been examined extensively in relation to cancer mortality. The vital status of study participants is updated biennially. Cohort study

Source: American Cancer Society

In Problems 23–26, determine the type of sampling used.

23. On election day, a pollster for Fox News positions herself outside a polling place near her home. She then asks the first 50 voters leaving the facility to complete a survey.

24. An Internet service provider randomly selects 15 residential blocks from a large city. It then surveys every household in these 15 blocks to determine the number that would use a high-speed Internet service if it were made available.

25. Thirty-five sophomores, 22 juniors, and 35 seniors are randomly selected to participate in a study from 574 sophomores, 462 juniors, and 532 seniors at a certain high school.

26. Officers for the Department of Motor Vehicles pull aside every 40th tractor trailer passing through a weigh station, starting with the 12th, for an emissions test.

27. Each survey has bias. Determine the type of bias and suggest a remedy.

(a) A politician sends a survey about tax issues to a random sample of subscribers to a literary magazine.

(b) An interviewer with little foreign language knowledge is sent to an area where her language is not commonly spoken. Response bias; interviewer error

(c) A data-entry clerk mistypes survey results into his computer. Data-entry error

28. Obtaining a Simple Random Sample The mayor of a small town wants to conduct personal interviews with small business owners to determine if there is anything the mayor could to do to help improve business conditions. The following list gives the names of the companies in the town. Obtain a simple random sample of size 5 from the companies in the town.

Allied Tube and Conduit	Lighthouse Financial	Senese's Winery
Bechstien Construction Co.	Mill Creek Animal Clinic	Skyline Laboratory
Cizer Trucking Co.	Nancy's Flowers	Solus, Maria, DDS
D & M Welding	Norm's Jewelry	Trust Lock and Key
Grace Cleaning Service	Papoose Children's Center	Ultimate Carpet
Jiffy Lube	Plaza Inn Motel	Waterfront Tavern
Levin, Thomas, MD	Risky Business Security	WPA Pharmacy

29. Obtaining a Systematic Sample A quality-control engineer wants to be sure that bolts coming off an assembly line are within prescribed tolerances. He wants to conduct a systematic sample by selecting every 9th bolt to come off the assembly line. The machine produces 30,000 bolts per day, and the engineer wants a sample of 32 bolts. Which bolts will be sampled?

30. Obtaining a Simple Random Sample Based on the Military Standard 105E (ANS1/ASQC Z1.4, ISO 2859) Tables, a lot of 91 to 150 items with an acceptable quality level (AQL) of 1% and a normal inspection plan would require a sample of size 13 to be inspected for defects. If the sample contains no defects, the entire lot is accepted. Otherwise, the entire lot is rejected. A shipment of 100 night-vision goggles is received and must be inspected. Discuss the procedure you would follow to obtain a simple random sample of 13 goggles to inspect.

31. Ballasts An electronics company has just developd a new electric ballast to be used in fluorescent bulbs. To determine if the new ballast is more energy efficient than the older ballast, the company randomly divides 200 fluorescent bulbs

23. Convenience sample 24. Cluster sample 25. Stratified sample 26. Systematic sample
27. (a) Sampling bias; undercoverage or nonrepresentative sample due to poor sampling frame

31. (a) Completely randomized design 31. (b) Energy required to light the bulb 31. (d) Group 2 (old ballast)
32. (a) Randomized block design 32. (b) Exam grade 32. (c) notecard use; with notecard or without notecard
32. (d) Instructor's statistics students

into two groups. The group 1 bulbs are to be given the new ballast and the group 2 bulbs are to be given the old ballast. The amount of energy required to light each bulb is measured.

(a) What type of experimental design is this?
(b) What is the response variable in this experiment?
(c) What are the treatments? Old and new ballasts
(d) Which group serves as the control group?
(e) What are the experimental units? 200 fluorescent bulbs
(f) What role does randomization play in this experiment?
(g) Draw a diagram similar to Figure 7, 8, or 10 to illustrate the design.

32. Exam Grades A statistics instructor wants to see if allowing students to use a notecard on exams will affect their overall exam scores. The instructor thinks that the results could be affected by the type of class: traditional, Web-enhanced, or online. He randomly selects half of his students from each type of class to use notecards. At the end of the semester, the average exam grades for those using notecards is compared to the average for those without notecards.

(a) What type of experimental design is this?
(b) What is the response variable in this experiment?
(c) What is the factor that is set to predetermined levels? What are the treatments?
(d) Identify the experimental units.
(e) Draw a diagram similar to Figure 7, 8, or 10 to illustrate the design.

33. Multiple Choice A common tip for taking multiple-choice tests is to always pick (b) or (c) if you are unsure. The idea is that instructors tend to feel the answer is more hidden if it is surrounded by distractor answers. An astute statistics instructor is aware of this and decides to use a table of random digits to select which choice will be the correct answer. If each question has five choices, use Table I in Appendix A or a random-number generator to determine the correct answers for a 20-question multiple-choice exam.

34. Humor in Advertising A marketing research firm wants to know whether information presented in a commercial is better recalled when presented using humor or serious commentary by adults between 18 and 35 years of age. They will use an exam that asks questions of 50 subjects about information presented in the ad. The response variable will be percentage of information recalled.

(a) Create a completely randomized design to answer the question. Be sure to include a diagram to illustrate your design.
(b) The chief of advertising criticizes your design because she knows that women react differently from men to advertising. Redesign the experiment as a randomized block design assuming that 30 of the subjects in your study are female and the remaining 20 are male. Be sure to include a diagram to illustrate your design.

35. Describe what is meant by a matched-pairs design. Contrast this experimental design with a completely randomized design.

CHAPTER TEST

Full solutions for this chapter test are available on the CD that accompanies this text. 3. Ratio, quantitative, continuous

1. List the four components that comprise the definition of statistics.

2. What is meant by the *process of statistics*?

In Problems 3–5, determine the level of measurement for the variable and identify if the variable is qualitative or quantitative. If the variable is quantitative, determine if it is discrete or continuous.

3. Time to complete the 500-meter race in speed skating.

4. Video game rating system by the Entertainment Software Rating Board (EC, E, E10+, T, M, AO, RP) Ordinal, qualitative

5. The number of surface imperfections on a camera lens. Ratio, quantitative, discrete

In Problems 6 and 7, determine whether the study depicts an observational study or a designed experiment. Identify the response variable in each case.

6. A random sample of 30 digital cameras is selected and divided into two groups. One group uses a brand-name battery, while the other uses a generic plain-label battery. All variables besides battery type are controlled. Pictures are taken under identical conditions and the battery life of the two groups is compared. Experiment, battery life

7. A sports reporter asks 100 baseball fans if Barry Bonds's 756th homerun ball should be marked with an asterisk when sent to the Baseball Hall of Fame. Observational study, asterisk or not

8. Contrast the three major types of observational studies in terms of the time frame when the data are collected.

9. Compare and contrast observational studies and designed experiments. Which study allows a researcher to claim causality? Experiment

10. Explain why it is important to use a control group and blinding in an experiment.

11. List the steps required to conduct an experiment.

12. A tanning company is looking for ways to improve customer satisfaction. They want to select a simple random sample of four stores from their 15 franchises in which to conduct customer satisfaction surveys. Discuss the procedure you would use, and then use the procedure to select a simple random sample of size $n = 4$. The locations are as follows:

Afton	Ballwin	Chesterfield	Clayton	Deer Creek
Ellisville	Farmington	Fenton	Ladue	Lake St. Louis
O'Fallon	Pevely	Shrewsbury	Troy	Warrenton

13. A congresswoman wants to survey her constituency regarding public policy. She asks one of her staff members to obtain a sample of residents of the district. The frame she has available lists 9,012 Democrats, 8,302 Republicans, and 3,012 Independents. Obtain a stratified random sample of 8 Democrats, 7 Republicans, and 3 Independents. Be sure to discuss the procedure used.

19. (f) 225 patients with skin irritations
20. (b) Bone mineral density; amount of cola consumed per week 20. (c) Quantitative
20. (e) Smoking status; alcohol consumption; physical activity; calcium intake

14. A farmer has a 500-acre orchard in Florida. Each acre is subdivided into blocks of 5. Altogether, there are 2,500 blocks of trees on the farm. After a frost, he wants to get an idea of the extent of the damage. Obtain a sample of 10 blocks of trees using a cluster sample. Be sure to discuss the procedure used.

15. A casino manager wants to inspect a sample of 14 slot machines in his casino for quality-control purposes. There are 600 sequentially numbered slot machines operating in the casino. Obtain a systematic sample of 14 slot machines. Be sure to discuss how you obtained the sample.

16. Describe what is meant by an experiment that is a completely randomized design. Contrast this experimental design with a randomized block design.

17. Each of the following surveys has bias. Identify the type of bias. 17. (a) Sampling bias (voluntary response)

(a) A television survey that gives 900 phone numbers for viewers to call with their vote. Each call costs $2.00.
(b) An employer distributes a survey to her 450 employees asking them how many hours each week, on average, they surf the Internet during business hours. Three of the employees complete the survey. Nonresponse bias
(c) A question on a survey asks, "Do you favor or oppose a minor increase in property tax to ensure fair salaries for teachers and properly equipped school buildings?"
(d) A researcher conducting a poll about national politics sends a survey to a random sample of subscribers to *Time* magazine. Sampling bias (undercoverage)

18. The four members of Skylab had their lymphocyte count per cubic millimeter measured 1 day before lift-off and measured again on their return to Earth. 18. (a) Lymphocyte count

(a) What is the response variable in this experiment?
(b) What is the treatment? Space flight
(c) What type of experimental design is this? Matched pair
(d) Identify the experimental units. 4 members
(e) Draw a diagram similar to Figure 7, 8, or 10 to illustrate the design.

19. Nucryst Pharmaceuticals, Inc., announced the results of its first human trial of NPI 32101, a topical form of its skin ointment. A total of 225 patients diagnosed with skin irritations were randomly divided into three groups as part of a double-blind, placebo-controlled study to test the effectiveness of the new topical cream. The first group received a 0.5% cream, the second group received a 1.0% cream, and the third group received a placebo. Groups were treated twice daily for a 6-week period.

Source: www.nucryst.com

(a) What type of experimental design is this?
(b) What is the response variable in this experiment?
(c) What is the factor that is set to predetermined levels? What are the treatments? Topical cream; 0.5%, 1.0%, 0%
(d) What does it mean for this study to be double-blind?
(e) What is the control group for this study?
(f) Identify the experimental units.
(g) Draw a diagram similar to Figure 7, 8, or 10 to illustrate the design.

17. (c) Response bias (poorly worded question) 19. (a) Completely randomized design 19. (b) Condition (dermatitis) improvement
19. (e) Subjects who get the 0% cream treatment

20. Researchers Katherine Tucker and associates wanted to determine whether consumption of cola is associated with lower bone mineral density. They looked at 1,125 men and 1,413 women in the Framingham Osteoporosis Study, which is a cohort that began in 1971. The first examination in this study began between 1971 and 1975, with participants returning for an examination every 4 years. Based on results of questionnaires, the researchers were able to determine cola consumption on a weekly basis. Analysis of the results indicated that women who consumed at least one cola per day (on average) had a bone mineral density that was significantly lower at the femoral neck than those who consumed less than one cola per day. The researchers did not find this relation in men.

Source: "Colas, but not other carbonated beverages, are associated with low bone mineral density in older women: The Framingam Osteoporosis Study," *American Journal of Clinical Nutrition* 84:936–942, 2006

(a) Why is this a cohort study?
(b) What is the response variable in this study? What is the explanatory variable?
(c) Is the response variable qualitative or quantitative?
(d) The following appears in the article: "Variables that could potentially confound the relation between carbonated beverage consumption and bone mineral density were obtained from information collected (in the questionnaire)." What does this mean?
(e) Can you think of any lurking variables that should be accounted for?
(f) What are the conclusions of the study? Does increased cola consumption cause a lower bone mineral density? Women who consumed at least one cola per day had a bone mineral density that was significantly lower at femoral neck than those who consumed less than one cola per day. No, they are associated.

MAKING AN INFORMED DECISION

What Movie Should I Go To?

One of the most difficult tasks of surveying is phrasing questions so that they are not misunderstood. In addition, questions must be phrased so that the researcher obtains answers that allow for meaningful analysis. We wish to create a questionnaire that can be used to make an informed decision about whether to attend a certain movie. Select a movie that you wish to see. If the movie is still in theaters, make sure that it has been released for at least a couple of weeks so that it is likely that a number of people have seen it. Design a questionnaire to be filled out by individuals who have seen the movie. You may wish to include questions regarding the demographics of the respondents first (such as age, gender, level of education, and so on). Ask as many questions as you feel are necessary to obtain an opinion regarding the movie. The questions can be open or closed. Administer the survey to at least 20 randomly selected people who have seen the movie. While administering the survey, keep track of those individuals who have not seen the movie. In particular, keep track of their demographic information. After administering the survey, summarize your findings. On the basis of the survey results, do you think that you will enjoy the movie? Why? Now see the movie. Did you like it? Did the survey accurately predict whether you would enjoy the movie? Now answer the following questions:

(a) What sampling method did you use? Why? Did you have a frame for the population?

(b) Did you have any problems with respondents misinterpreting your questions? How could this issue have been resolved?

(c) What role did the demographics of the respondents have in forming your opinion? Why?

(d) Did the demographics of individuals who did not see the movie play a role while you were forming your opinion regarding the movie?

(e) Look up a review of the movie by a professional movie critic. Did the movie critic's opinion agree with yours? What might account for the similarities or differences in your opinions?

(f) Describe the problems that you had in administering the survey. If you had to do this survey over again, would you change anything? Why?

CASE STUDY

Chrysalises for Cash

The colorful butterfly symbolizes the notion of personal change. Increasingly, people are turning to butterflies to consecrate meaningful events, such as birthdays, weddings, and funerals. To fill this need, a new industry has hatched.

Suppliers of butterflies are closely monitored by governmental agencies to ensure that local environments are not subjected to the introduction of invasive species. In addition to following these regulations, butterfly suppliers need to ensure the quality and quantity of their product, while maintaining a profit. To this end, an individual supplier may hire a number of smaller independent contractors to hatch the varieties needed. These entrepreneurs are paid a small fee for each chrysalis delivered, with a 50% bonus added for each hatched healthy butterfly. This fee structure provides little room for profit. Therefore, it is important that these small contractors deliver a high proportion of healthy butterflies that emerge at a fairly predictable rate.

In Florida, one such entrepreneur specializes in harvesting the black swallowtail butterfly. In the southern United States, the black swallowtail has at least three broods a year. The female flutters through open fields seeking plants of the Apiaceae family, such as carrot and parsley, upon which to lay her eggs. The resulting caterpillars are dark brown with a small white saddle mark. As the caterpillars consume the leaves of their host plant, they increase in size, changing color to a vibrant green with intermittent stripes of yellow and black. Once a caterpillar has eaten enough, it secures itself and sheds its skin, revealing an emerald chrysalis. During this resting phase, environmental factors such as temperature and humidity may affect the transformation process. Typically, the black swallowtail takes about 1 week to complete its metamorphosis and emerge from its chrysalis. The transformation occasionally results in a deformed butterfly. Deformities range from wings that will not fully open to missing limbs.

The Florida contractor believes that there are differences in quality and emergence time among his broods. Not having taken a scientific approach to the problem, he relies on his memory of seasons past. It seems to him that late-season butterflies emerge sooner and with a greater number of deformities than their early-season counterparts. He also speculates that the type and nutritional value of the food consumed by the caterpillar might contribute to any observed differences. This year he is committed to a more formal approach to his butterfly harvest.

Since it takes 2 days to deliver the chrysalises from the contractor to the supplier, it is important that the butterflies do not emerge prematurely. It is equally important that the number of defective butterflies be minimized. With these two goals in mind, the contractor seeks the best combination of food source, fertilizer, and brood season to maximize his profits. To examine the effects of these variables on emergence time and number of deformed butterflies, the entrepreneur designed the following experiment.

Eight identical pots were filled with equal amounts of a particular soil mixture. The watering of all the pots and the plants contained within them was carefully monitored for consistency. Four pots were set outside during the early part of the brood season. Of these, two contained carrot plants, while the remaining pair grew parsley. For the carrot pair, one pot was fed a fixed amount of liquid fertilizer, while the other pot was fed a nutritionally similar amount of solid fertilizer. The two pots containing parsley were similarly fertilized. All four pots were placed next to each other to ensure similar exposures to environmental conditions such as temperature and solar radiation. Five black swallowtail caterpillars of similar age were placed into each container. The caterpillars were allowed to mature and form a chrysalis. The time from chrysalis formation until emergence was reported to the nearest day. The occurrence of any defects was also noted. The same procedure was followed with the four pots that were placed outdoors during the late brood season.

Write a report describing the experimental goals and design for the entrepreneur's experiment. Follow the procedure outlined in the box on the steps in designing and conducting an experiment (pp. 47–48). Step 5(b), *collect and process the data*, of this procedure is provided in the following table and should be included in your report.

Florida Black Swallowtail Chrysalis Experiment Data

Season	Food	Fertilizer	Number Deformed	Emergence Time (Days)
Early	Parsley	Solid	0	6, 6, 7, 7, 7
Early	Parsley	Liquid	0	6, 7, 7, 8, 8
Early	Carrot	Solid	1	3, 6, 6, 7, 8
Early	Carrot	Liquid	0	6, 6, 7, 8, 8
Late	Parsley	Solid	2	2, 3, 4, 4, 5
Late	Parsley	Liquid	1	2, 3, 4, 5, 5
Late	Carrot	Solid	2	3, 3, 3, 4, 5
Late	Carrot	Liquid	0	2, 4, 4, 4, 5

In your report, provide a general descriptive analysis of these data. Be sure to include recommendations for the combination of season, food source, and type of fertilizer that result in the fewest deformed butterflies while achieving a long emergence time.

Conclude your report with recommendations for further experiments. For each proposed experiment, be sure to do the following:

1. State the problem to be solved and define the response variables.
2. Define the factors that affect the response variables.
3. State the number of experimental units.
4. State the treatment.