C H A P T E R

1

Introduction to Statistics

Preview

Before we begin our discussion of statistics, we ask you to read the following paragraph taken from the philosophy of Wrong Shui (Candappa, 2000).

> The Journey to Enlightenment
>
> In Wrong Shui, life is seen as a cosmic journey, a struggle to overcome unseen and unexpected obstacles at the end of which the traveler will find illumination and enlightenment. Replicate this quest in your home by moving light switches away from doors and over to the far side of each room.*

Why did we begin a statistics book with a bit of twisted philosophy? Actually, the paragraph is an excellent (and humorous) counterexample for the purpose of this book. Specifically, our goal is to help you avoid stumbling around in the dark by providing lots of easily available light switches and plenty of illumination as you journey through the world of statistics. To accomplish this, we try to present sufficient background and a clear statement of purpose as we introduce each new statistical procedure. Remember that all statistical procedures were developed to serve a purpose. If you understand why a new procedure is needed, you will find it much easier to learn.

The objectives for this first chapter are to provide an introduction to the topic of statistics and to give you some background for the rest of the book. We discuss the role of statistics within the general field of scientific inquiry, and we introduce some of the vocabulary and notation that are necessary for the statistical methods that follow.

As you read through the following chapters, keep in mind that the general topic of statistics follows a well-organized, logically developed progression that leads from basic concepts and definitions to increasingly sophisticated techniques. Thus, the material presented in the early chapters of this book serves as a foundation for the material that follows. The content of the first nine chapters, for example, provides an essential background and context for the statistical methods presented in Chapter 10. If you turn directly to Chapter 10 without reading the first nine chapters, you will find the material confusing and incomprehensible. However, if you learn and use the background material, you will have a good frame of reference for understanding and incorporating new concepts as they are presented.

*Candappa, R. (2000). *The little book of wrong shui.* Kansas City: Andrews McMeel Publishing. Reprinted by permission.

1.1 STATISTICS, SCIENCE, AND OBSERVATIONS

DEFINITIONS OF STATISTICS

By one definition, *statistics* consist of facts and figures such as average income, crime rate, birth rate, baseball batting averages, and so on. These statistics are usually informative and time saving because they condense large quantities of information into a few simple figures. Later in this chapter we return to the notion of calculating statistics (facts and figures) but, for now, we concentrate on a much broader definition of statistics. Specifically, we use the term statistics to refer to a set of mathematical procedures. In this case, we are using the term *statistics* as a shortened version of *statistical procedures*. For example, you are probably using this book for a statistics course in which you will learn about the statistical techniques that are used for research in the behavioral sciences.

Research in psychology (and other fields) involves gathering information. To determine, for example, whether violence on TV has any effect on children's behavior, you would need to gather information about children's behaviors and the TV programs they watch. When researchers finish the task of gathering information, they typically find themselves with pages and pages of measurements such as IQ scores, personality scores, reaction time scores, and so on. In this book, we present the statistics that researchers use to analyze and interpret the information that they gather. Specifically, statistics serve two general purposes:

1. Statistics are used to organize and summarize the information so that the researcher can see what happened in the research study and can communicate the results to others.
2. Statistics help the researcher to answer the questions that initiated the research by determining exactly what general conclusions are justified based on the specific results that were obtained.

DEFINITION

The term **statistics** refers to a set of mathematical procedures for organizing, summarizing, and interpreting information.

Statistical procedures help to ensure that the information or observations are presented and interpreted in an accurate and informative way. In somewhat grandiose terms, statistics help researchers bring order out of chaos. In addition, statistics provide researchers with a set of standardized techniques that are recognized and understood throughout the scientific community. Thus, the statistical methods used by one researcher are familiar to other researchers, who can accurately interpret the statistical analyses with a full understanding of how the analysis was done and what the results signify.

1.2 POPULATIONS AND SAMPLES

WHAT ARE THEY?

Research in the behavioral sciences typically begins with a general question about a specific group (or groups) of individuals. For example, a researcher may be interested in the effect of divorce on the self-esteem of preteen children. Or a researcher may want to examine the amount of time spent in the bathroom for men compared to women. In the first example, the researcher is interested in the group of *preteen children*. In the second example, the researcher wants to compare the group of *men* with the group of *women*. In statistical terminology, the entire group that a researcher wishes to study is called a *population*.

DEFINITION

A **population** is the set of all the individuals of interest in a particular study.

As you can well imagine, a population can be quite large—for example, the entire set of women on the planet Earth. A researcher might be more specific, limiting the population for study to women who are registered voters in the United States. Perhaps the investigator would like to study the population consisting of women who are heads of state. Populations can obviously vary in size from extremely large to very small, depending on how the researcher defines the population. The population being studied should always be identified by the researcher. In addition, the population need not consist of people—it could be a population of rats, corporations, parts produced in a factory, or anything else a researcher wants to study. In practice, populations are typically very large, such as the population of college sophomores in the United States or the population of small businesses.

Because populations tend to be very large, it usually is impossible for a researcher to examine every individual in the population of interest. Therefore, researchers typically select a smaller, more manageable group from the population and limit their studies to the individuals in the selected group. In statistical terms, a set of individuals selected from a population is called a *sample*. A sample is intended to be representative of its population, and a sample should always be identified in terms of the population from which it was selected.

DEFINITION

A **sample** is a set of individuals selected from a population, usually intended to represent the population in a research study.

Just as we saw with populations, samples can vary in size. For example, one study might examine a sample of only 10 students in a graduate program, and another study might use a sample of more than 1,000 registered voters representing the population of a major city.

So far we have talked about a sample being selected from a population. However, this is actually only half of the full relationship between a sample and its population. Specifically, when a researcher finishes examining the sample, the goal is to generalize the results back to the entire population. Remember that the research started with a general question about the population. To answer the question, a researcher studies a sample and then generalizes the results from the sample to the population. The full relationship between a sample and a population is shown in Figure 1.1.

VARIABLES AND DATA

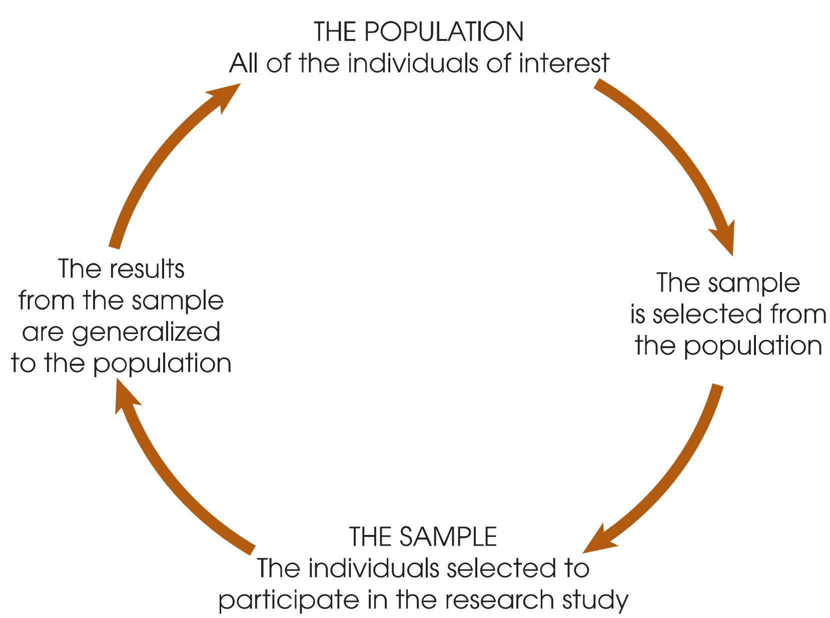
Typically, researchers are interested in specific characteristics of the individuals in the population (or in the sample), or they are interested in outside factors that may influence the individuals. For example, a researcher may be interested in the influence of the weather on people's moods. As the weather changes, do people's moods also change? Something that can change or have different values is called a *variable*.

DEFINITION

A **variable** is a characteristic or condition that changes or has different values for different individuals.

FIGURE 1.1

The relationship between a population and a sample.



Once again, variables can be characteristics that differ from one individual to another, such as height, weight, gender, or personality. Also, variables can be environmental conditions that change, such as temperature, time of day, or the size of the room in which the research is being conducted.

To demonstrate changes in variables, it is necessary to make measurements of the variables being examined. The measurement obtained for each individual is called a *datum* or, more commonly, a *score* or *raw score*. The complete set of scores is called the *data set* or simply the *data*.

DEFINITIONS

Data (plural) are measurements or observations. A **data set** is a collection of measurements or observations. A **datum** (singular) is a single measurement or observation and is commonly called a **score** or **raw score**.

Before we move on, we should make one more point about samples, populations, and data. Earlier, we defined populations and samples in terms of *individuals*. For example, we discussed a population of college sophomores and a sample of preschool children. Be forewarned, however, that we will also refer to populations or samples of *scores*. Because research typically involves measuring each individual to obtain a score, every sample (or population) of individuals produces a corresponding sample (or population) of scores.

PARAMETERS AND STATISTICS

When describing data, it is necessary to distinguish whether the data come from a population or a sample. A characteristic that describes a population—for example, the average score for the population—is called a *parameter*. A characteristic that describes a sample is called a *statistic*. Thus, the average score for a sample is an example of a statistic. Typically, the research process begins with a question about a population parameter. However, the actual data come from a sample and are used to compute sample statistics.

DEFINITIONS

A **parameter** is a value, usually a numerical value, that describes a population. A parameter is usually derived from measurements of the individuals in the population.

A **statistic** is a value, usually a numerical value, that describes a sample. A statistic is usually derived from measurements of the individuals in the sample.

Every population parameter has a corresponding sample statistic, and most research studies involve using statistics from samples as the basis for answering questions about population parameters. As a result, much of this book is concerned with the relationship between sample statistics and the corresponding population parameters. In Chapter 7, for example, we examine the relationship between the mean obtained for a sample and the mean for the population from which the sample was obtained.

DESCRIPTIVE AND INFERENTIAL STATISTICAL METHODS

Although researchers have developed a variety of different statistical procedures to organize and interpret data, these different procedures can be classified into two general categories. The first category, *descriptive statistics*, consists of statistical procedures that are used to simplify and summarize data.

DEFINITION

Descriptive statistics are statistical procedures used to summarize, organize, and simplify data.

Descriptive statistics are techniques that take raw scores and organize or summarize them in a form that is more manageable. Often the scores are organized in a table or a graph so that it is possible to see the entire set of scores. Another common technique is to summarize a set of scores by computing an average. Note that even if the data set has hundreds of scores, the average provides a single descriptive value for the entire set.

The second general category of statistical techniques is called *inferential statistics*. Inferential statistics are methods that use sample data to make general statements about a population.

DEFINITION

Inferential statistics consist of techniques that allow us to study samples and then make generalizations about the populations from which they were selected.

Because populations are typically very large, it usually is not possible to measure everyone in the population. Therefore, a sample is selected to represent the population. By analyzing the results from the sample, we hope to make general statements about the population. Typically, researchers use sample statistics as the basis for drawing conclusions about population parameters.

One problem with using samples, however, is that a sample provides only limited information about the population. Although samples are generally *representative* of their populations, a sample is not expected to give a perfectly accurate picture of the whole population. There usually is some discrepancy between a sample statistic and the corresponding population parameter. This discrepancy is called *sampling error*, and it creates the fundamental problem that inferential statistics must always address (Box 1.1).

DEFINITION

Sampling error is the naturally occurring discrepancy, or error, that exists between a sample statistic and the corresponding population parameter.

The concept of sampling error is illustrated in Figure 1.2. The figure shows a population of 1,000 college students and two samples, each with 5 students, that have been selected from the population. Notice that each sample contains different individuals who have different characteristics. Because the characteristics of each sample depend on the specific people in the sample, statistics vary from one sample to another. For example, the five students in sample 1 have an average age of 19.8 years and the students in sample 2 have an average age of 20.4 years.

BOX 1.1

THE MARGIN OF ERROR BETWEEN STATISTICS AND PARAMETERS

One common example of sampling error is the error associated with a sample proportion. For example, in newspaper articles reporting results from political polls, you frequently find statements such as this:

> Candidate Brown leads the poll with 51% of the vote. Candidate Jones has 42% approval, and the remaining 7% are undecided. This poll was taken from a sample of registered voters and has a margin of error of plus-or-minus 4 percentage points.

The *margin of error* is the sampling error. In this case, the percentages that are reported were obtained from a sample and are being generalized to the whole population. As always, you do not expect the statistics from a sample to be perfect. There is always some margin of error when sample statistics are used to represent population parameters.

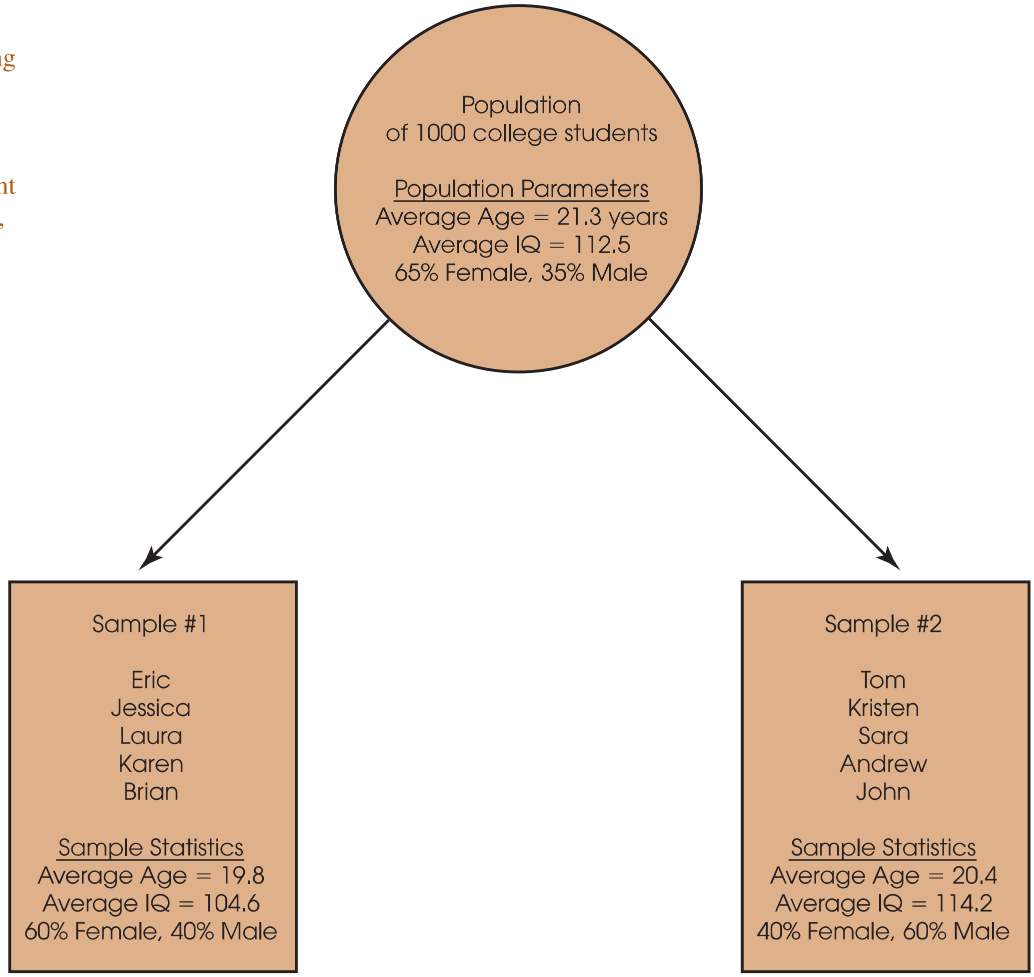


FIGURE 1.2

A demonstration of sampling error. Two samples are selected from the same population. Notice that the sample statistics are different from one sample to another, and all of the sample statistics are different from the corresponding population parameters. The natural differences that exist, by chance, between a sample statistic and a population parameter are called sampling error.

It is also very unlikely that the statistics obtained for a sample are identical to the parameters for the entire population. In Figure 1.2, for example, neither sample has statistics that are exactly the same as the population parameters. You should also realize that Figure 1.2 shows only two of the hundreds of possible samples. Each sample would contain different individuals and would produce different statistics. This is the basic concept of sampling error: sample statistics vary from one sample to another and typically are different from the corresponding population parameters.

As a further demonstration of sampling error, imagine that your statistics class is separated into two groups by drawing a line from front to back through the middle of the room. Now imagine that you compute the average age (or height, or IQ) for each group. Will the two groups have exactly the same average? Almost certainly they will not. No matter what you chose to measure, you will probably find some difference between the two groups.

However, the difference you obtain does not necessarily mean that there is a systematic difference between the two groups. For example, if the average age for students on the right-hand side of the room is higher than the average for students on the left, it is unlikely that some mysterious force has caused the older people to gravitate to the right side of the room. Instead, the difference is probably the result of random factors such as chance. The unpredictable, unsystematic differences that exist from one sample to another are an example of sampling error.

STATISTICS IN THE CONTEXT OF RESEARCH

The following example shows the general stages of a research study and demonstrates how descriptive statistics and inferential statistics are used to organize and interpret the data. At the end of the example, note how sampling error can affect the interpretation of experimental results, and consider why inferential statistical methods are needed to deal with this problem.

EXAMPLE 1.1

Figure 1.3 shows an overview of a general research situation and demonstrates the roles that descriptive and inferential statistics play. The purpose of the research study

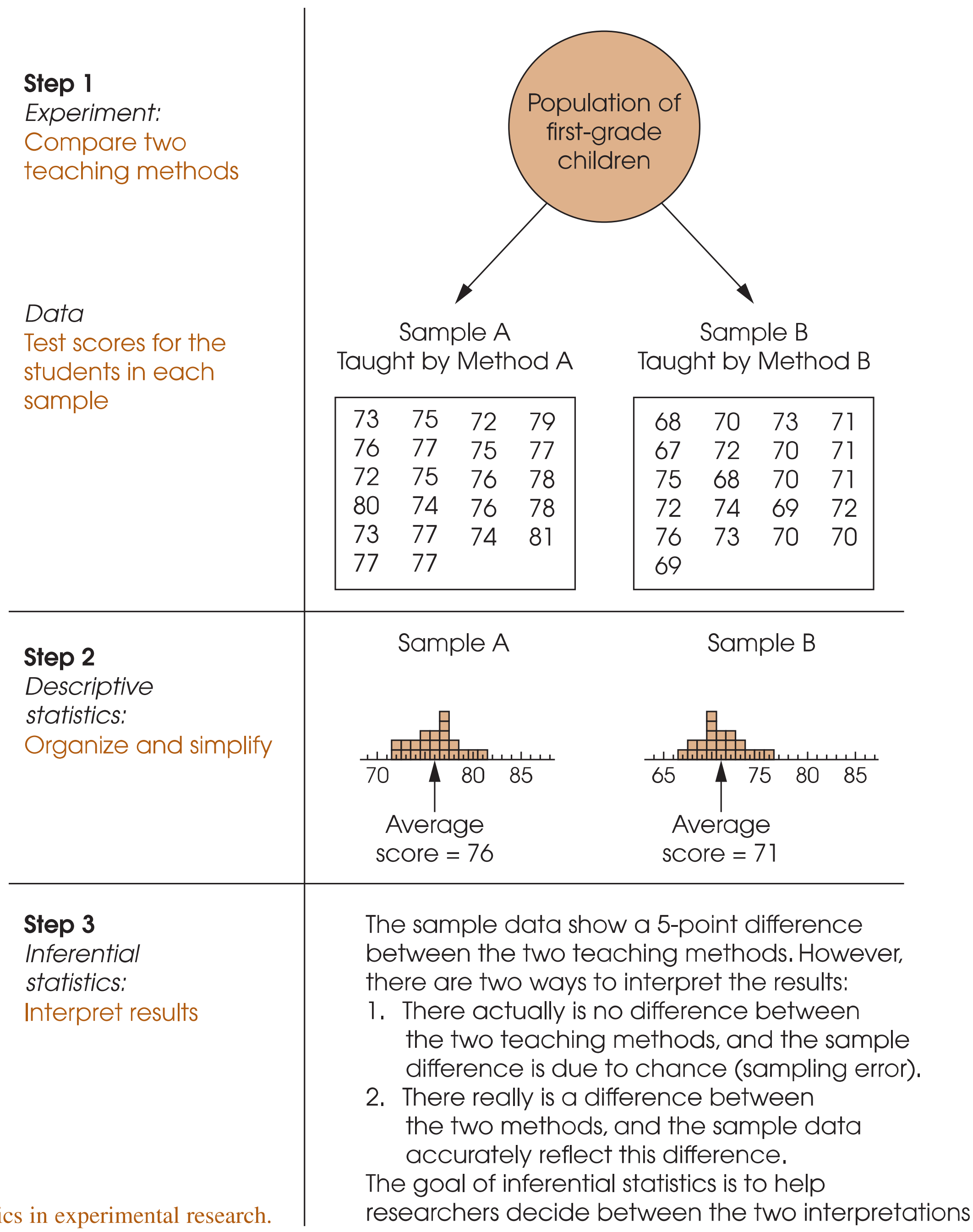


FIGURE 1.3

The role of statistics in experimental research.

is to evaluate the difference between two methods for teaching reading to first-grade children. Two samples are selected from the population of first-grade children. The children in sample A are assigned to teaching method A and the children in sample B are assigned to method B. After 6 months, all of the students are given a standardized reading test. At this point, the researcher has two sets of data: the scores for sample A and the scores for sample B (see Figure 1.3). Now is the time to begin using statistics.

First, descriptive statistics are used to simplify the pages of data. For example, the researcher could draw a graph showing the scores for each sample or compute the average score for each sample. Note that descriptive methods provide a simplified, organized description of the scores. In this example, the students taught by method A averaged 76 on the standardized test, and the students taught by method B averaged only 71.

Once the researcher has described the results, the next step is to interpret the outcome. This is the role of inferential statistics. In this example, the researcher has found a difference of 5 points between the two samples (sample A averaged 76 and sample B averaged 71). The problem for inferential statistics is to differentiate between the following two interpretations:

1. There is no real difference between the two teaching methods, and the 5-point difference between the samples is just an example of sampling error (like the samples in Figure 1.2).
2. There really is a difference between the two teaching methods, and the 5-point difference between the samples was caused by the different methods of teaching.

In simple English, does the 5-point difference between samples provide convincing evidence of a difference between the two teaching methods, or is the 5-point difference just chance? The purpose of inferential statistics is to answer this question.

LEARNING CHECK

1. A researcher is interested in the texting habits of high school students in the United States. If the researcher measures the number of text messages that each individual sends each day and calculates the average number for the entire group of high school students, the average number would be an example of a ____________.
2. A researcher is interested in how watching a reality television show featuring fashion models influences the eating behavior of 13-year-old girls.
   a. A group of 30 13-year-old girls is selected to participate in a research study. The group of 30 13-year-old girls is an example of a ____________.
   b. In the same study, the amount of food eaten in one day is measured for each girl and the researcher computes the average score for the 30 13-year-old girls. The average score is an example of a ___________.
3. Statistical techniques are classified into two general categories. What are the two categories called, and what is the general purpose for the techniques in each category?
4. Briefly define the concept of sampling error.

ANSWERS

1. parameter
2. **a.** sample
   b. statistic

3. The two categories are descriptive statistics and inferential statistics. Descriptive techniques are intended to organize, simplify, and summarize data. Inferential techniques use sample data to reach general conclusions about populations.

4. Sampling error is the error, or discrepancy, between the value obtained for a sample statistic and the value for the corresponding population parameter.

1.3 DATA STRUCTURES, RESEARCH METHODS, AND STATISTICS

INDIVIDUAL VARIABLES

Some research studies are conducted simply to describe individual variables as they exist naturally. For example, a college official may conduct a survey to describe the eating, sleeping, and study habits of a group of college students. When the results consist of numerical scores, such as the number of hours spent studying each day, they are typically described by the statistical techniques that are presented in Chapters 3 and 4. Non-numerical scores are typically described by computing the proportion or percentage in each category. For example, a recent newspaper article reported that 61% of the adults in the United States currently drink alcohol.

RELATIONSHIPS BETWEEN VARIABLES

Most research, however, is intended to examine relationships between two or more variables. For example, is there a relationship between the amount of violence that children see on television and the amount of aggressive behavior they display? Is there a relationship between the quality of breakfast and level of academic performance for elementary school children? Is there a relationship between the number of hours of sleep and grade point average for college students? To establish the existence of a relationship, researchers must make observations—that is, measurements of the two variables. The resulting measurements can be classified into two distinct data structures that also help to classify different research methods and different statistical techniques. In the following section we identify and discuss these two data structures.

I. Measuring Two Variables for Each Individual: The Correlational Method
One method for examining the relationship between variables is to observe the two variables as they exist naturally for a set of individuals. That is, simply measure the two variables for each individual. For example, research has demonstrated a relationship between sleep habits, especially wake-up time, and academic performance for college students (Trockel, Barnes, and Egget, 2000). The researchers used a survey to measure wake-up time and school records to measure academic performance for each student. Figure 1.4 shows an example of the kind of data obtained in the study. The researchers then look for consistent patterns in the data to provide evidence for a relationship between variables. For example, as wake-up time changes from one student to another, is there also a tendency for academic performance to change?

Consistent patterns in the data are often easier to see if the scores are presented in a graph. Figure 1.4 also shows the scores for the eight students in a graph called a scatter plot. In the scatter plot, each individual is represented by a point so that the horizontal position corresponds to the student's wake-up time and the vertical position corresponds to the student's academic performance score. The scatter plot shows a clear relationship between wake-up time and academic performance: as wake-up time increases, academic performance decreases.

(a)

Student	Wake-up Time	Academic Performance
A	11	2.4
B	9	3.6
C	9	3.2
D	12	2.2
E	7	3.8
F	10	2.2
G	10	3.0
H	8	3.0

(b)

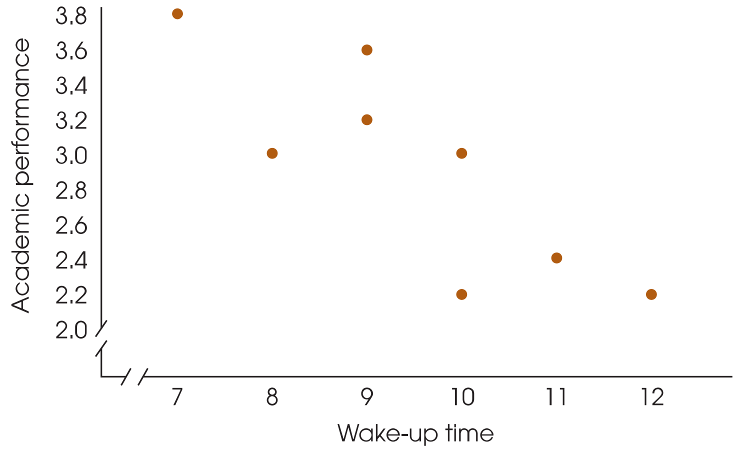


FIGURE 1.4

One of two data structures for studies evaluating the relationship between variables. Note that there are two separate measurements for each individual (wake-up time and academic performance). The same scores are shown in a table (a) and in a graph (b).

A research study that simply measures two different variables for each individual and produces the kind of data shown in Figure 1.4 is an example of the *correlational method*, or the *correlational research strategy.*

DEFINITION

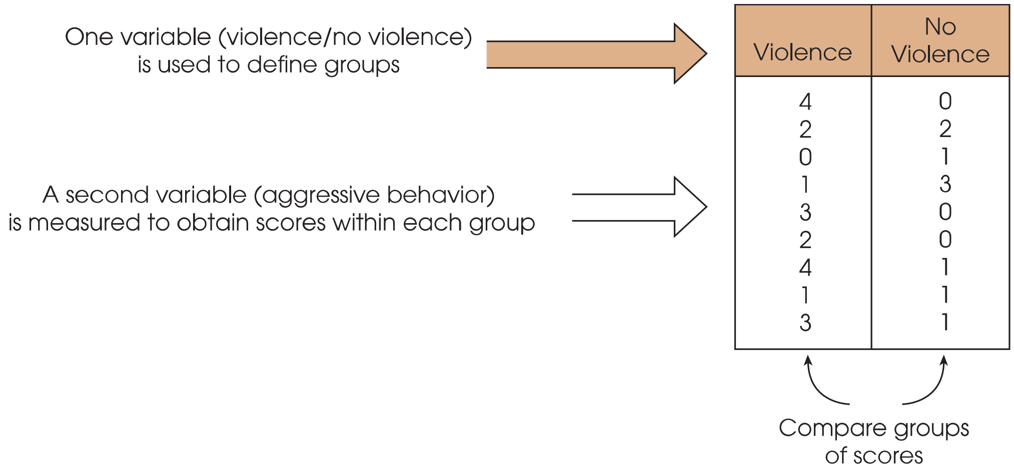
> In the **correlational method**, two different variables are observed to determine whether there is a relationship between them.

Limitations of the Correlational Method The results from a correlational study can demonstrate the existence of a relationship between two variables, but they do not provide an explanation for the relationship. In particular, a correlational study cannot demonstrate a cause-and-effect relationship. For example, the data in Figure 1.4 show a systematic relationship between wake-up time and academic performance for a group of college students; those who sleep late tend to have lower performance scores than those who wake early. However, there are many possible explanations for the relationship and we do not know exactly what factor (or factors) is responsible for late sleepers having lower grades. In particular, we cannot conclude that waking students up earlier would cause their academic performance to improve, or that studying more would cause students to wake up earlier. To demonstrate a cause-and-effect relationship between two variables, researchers must use the experimental method, which is discussed next.

II. Comparing Two (or More) Groups of Scores: Experimental and Nonexperimental Methods The second method for examining the relationship between two variables involves the comparison of two or more groups of scores. In this situation, the relationship between variables is examined by using one of the variables to define the groups, and then measuring the second variable to obtain scores for each group. For example, one group of elementary school children is shown a 30-minute action/adventure television program involving numerous instances of violence, and a second group is shown a 30-minute comedy that includes no violence. Both groups are

FIGURE 1.5

The second data structure for studies evaluating the relationship between variables. Note that one variable is used to define the groups and the second variable is measured to obtain scores within each group.



Violence	No Violence
4	0
2	2
0	1
1	3
3	0
2	0
4	1
1	1
3	1

then observed on the playground and a researcher records the number of aggressive acts committed by each child. An example of the resulting data is shown in Figure 1.5. The researcher compares the scores for the violence group with the scores for the no-violence group. A systematic difference between the two groups provides evidence for a relationship between viewing television violence and aggressive behavior for elementary school children.

THE EXPERIMENTAL METHOD

One specific research method that involves comparing groups of scores is known as the *experimental method* or the *experimental research strategy*. The goal of an experimental study is to demonstrate a cause-and-effect relationship between two variables. Specifically, an experiment attempts to show that changing the value of one variable causes changes to occur in the second variable. To accomplish this goal, the experimental method has two characteristics that differentiate experiments from other types of research studies:

1. **Manipulation** The researcher manipulates one variable by changing its value from one level to another. A second variable is observed (measured) to determine whether the manipulation causes changes to occur.
2. **Control** The researcher must exercise control over the research situation to ensure that other, extraneous variables do not influence the relationship being examined.

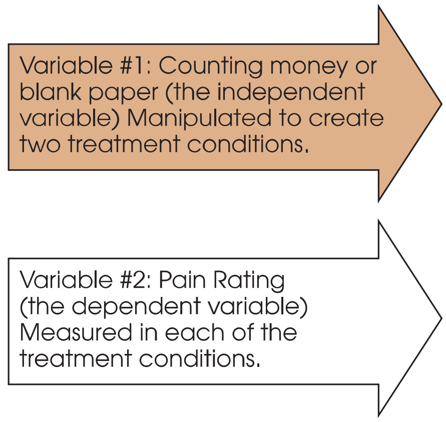
In more complex experiments, a researcher may systematically manipulate more than one variable and may observe more than one variable. Here we are considering the simplest case, in which only one variable is manipulated and only one variable is observed.

To demonstrate these two characteristics, consider an experiment in which researchers demonstrate the pain-killing effects of handling money (Zhou & Vohs, 2009). In the experiment, a group of college students was told that they were participating in a manual dexterity study. The researcher then manipulated the treatment conditions by giving half of the students a stack of money to count and the other half a stack of blank pieces of paper. After the counting task, the participants were asked to dip their hands into bowls of painfully hot water (122° F) and rate how uncomfortable it was. Participants who had counted money rated the pain significantly lower than those who had counted paper. The structure of the experiment is shown in Figure 1.6.

To be able to say that the difference in pain is caused by the money, the researcher must rule out any other possible explanation for the difference. That is, any other

FIGURE 1.6

The structure of an experiment. Participants are randomly assigned to one of two treatment conditions: counting money or counting blank pieces of paper. Later, each participant is tested by placing one hand in a bowl of hot (122° F) water and rating the level of pain. A difference between the ratings for the two groups is attributed to the treatment (paper versus money).



Money	Paper
7	8
4	10
5	8
6	9
6	8
8	10
6	7
5	8
5	8
6	7

Compare groups of scores

variables that might affect pain tolerance must be controlled. There are two general categories of variables that researchers must consider:

1. **Participant Variables** These are characteristics such as age, gender, and intelligence that vary from one individual to another. Whenever an experiment compares different groups of participants (one group in treatment A and a different group in treatment B), researchers must ensure that participant variables do not differ from one group to another. For the experiment shown in Figure 1.6, for example, the researchers would like to conclude that handling money instead of plain paper causes a change in the participants' perceptions of pain. Suppose, however, that the participants in the money condition were primarily females and those in the paper condition were primarily males. In this case, there is an alternative explanation for any difference in the pain ratings that exists between the two groups. Specifically, it is possible that the difference in pain was caused by the money, but it also is possible that the difference was caused by the participants' gender (females can tolerate more pain than males can). Whenever a research study allows more than one explanation for the results, the study is said to be *confounded* because it is impossible to reach an unambiguous conclusion.
2. **Environmental Variables** These are characteristics of the environment such as lighting, time of day, and weather conditions. A researcher must ensure that the individuals in treatment A are tested in the same environment as the individuals in treatment B. Using the money-counting experiment (see Figure 1.6) as an example, suppose that the individuals in the money condition were all tested in the morning and the individuals in the paper condition were all tested in the evening. Again, this would produce a confounded experiment because the researcher could not determine whether the differences in the pain ratings were caused by the money or caused by the time of day.

Researchers typically use three basic techniques to control other variables. First, the researcher could use *random assignment*, which means that each participant has an equal chance of being assigned to each of the treatment conditions. The goal of random assignment is to distribute the participant characteristics evenly between the two groups so that neither group is noticeably smarter (or older, or faster) than the other. Random assignment can also be used to control environmental variables. For example, participants could be assigned randomly for testing either in the morning or in the afternoon. Second, the researcher can use *matching* to ensure equivalent groups or equivalent environments. For example, the researcher could match groups by ensuring that every group has exactly 60% females and 40% males. Finally, the researcher can control variables by *holding them constant.* For example, if an experiment uses only 10-year-old children as participants (holding age constant), then the researcher can be certain that one group is not noticeably older than another.

DEFINITION

In the **experimental method**, one variable is manipulated while another variable is observed and measured. To establish a cause-and-effect relationship between the two variables, an experiment attempts to control all other variables to prevent them from influencing the results.

Terminology in the Experimental Method Specific names are used for the two variables that are studied by the experimental method. The variable that is manipulated by the experimenter is called the *independent variable*. It can be identified as the treatment conditions to which participants are assigned. For the example in Figure 1.6, money versus paper is the independent variable. The variable that is observed and measured to obtain scores within each condition is the *dependent variable*. For the example in Figure 1.6, the level of pain is the dependent variable.

DEFINITIONS

The **independent variable** is the variable that is manipulated by the researcher. In behavioral research, the independent variable usually consists of the two (or more) treatment conditions to which subjects are exposed. The independent variable consists of the *antecedent* conditions that were manipulated *prior* to observing the dependent variable.

The **dependent variable** is the variable that is observed to assess the effect of the treatment.

Control conditions in an experiment An experimental study evaluates the relationship between two variables by manipulating one variable (the independent variable) and measuring one variable (the dependent variable). Note that in an experiment only one variable is actually measured. You should realize that this is different from a correlational study, in which both variables are measured and the data consist of two separate scores for each individual.

Often an experiment will include a condition in which the participants do not receive any treatment. The scores from these individuals are then compared with scores from participants who do receive the treatment. The goal of this type of study is to demonstrate that the treatment has an effect by showing that the scores in the treatment condition are substantially different from the scores in the no-treatment condition. In this kind of research, the no-treatment condition is called the *control condition*, and the treatment condition is called the *experimental condition*.

DEFINITIONS

Individuals in a **control condition** do not receive the experimental treatment. Instead, they either receive no treatment or they receive a neutral, placebo treatment. The purpose of a control condition is to provide a baseline for comparison with the experimental condition.

Individuals in the **experimental condition** do receive the experimental treatment.

Note that the independent variable always consists of at least two values. (Something must have at least two different values before you can say that it is "variable.") For the money-counting experiment (see Figure 1.6), the independent variable is money versus plain paper. For an experiment with an experimental group and a control group, the independent variable is treatment versus no treatment.

NONEXPERIMENTAL METHODS: NONEQUIVALENT GROUPS AND PRE–POST STUDIES

In informal conversation, there is a tendency for people to use the term *experiment* to refer to any kind of research study. You should realize, however, that the term only applies to studies that satisfy the specific requirements outlined earlier. In particular, a real experiment must include manipulation of an independent variable and rigorous control of other, extraneous variables. As a result, there are a number of other research designs that are not true experiments but still examine the relationship between variables by comparing groups of scores. Two examples are shown in Figure 1.7 and are discussed in the following paragraphs. This type of research study is classified as nonexperimental.

The top part of Figure 1.7 shows an example of a *nonequivalent groups* study comparing boys and girls. Notice that this study involves comparing two groups of scores (like an experiment). However, the researcher has no ability to control which

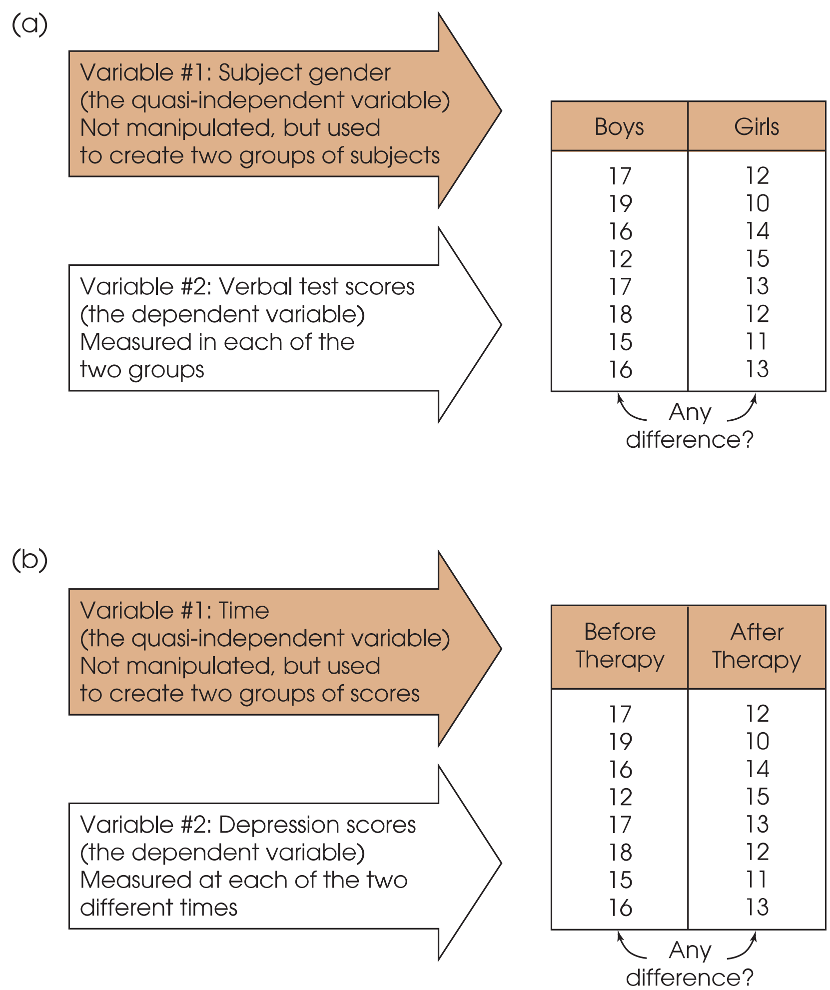


FIGURE 1.7

Two examples of nonexperimental studies that involve comparing two groups of scores. In (a) the study uses two preexisting groups (boys/girls) and measures a dependent variable (verbal scores) in each group. In (b) the study uses time (before/after) to define the two groups and measures a dependent variable (depression) in each group.

participants go into which group—all the males must be in the boy group and all the females must be in the girl group. Because this type of research compares preexisting groups, the researcher cannot control the assignment of participants to groups and cannot ensure equivalent groups. Other examples of nonequivalent group studies include comparing 8-year-old children and 10-year-old children, people with an eating disorder and those with no disorder, and comparing children from a single-parent home and those from a two-parent home. Because it is impossible to use techniques like random assignment to control participant variables and ensure equivalent groups, this type of research is not a true experiment.

Correlational studies are also examples of nonexperimental research. In this section, however, we are discussing non-experimental studies that compare two or more groups of scores.

The bottom part of Figure 1.7 shows an example of a *pre–post* study comparing depression scores before therapy and after therapy. The two groups of scores are obtained by measuring the same variable (depression) twice for each participant; once before therapy and again after therapy. In a pre–post study, however, the researcher has no control over the passage of time. The "before" scores are always measured earlier than the "after" scores. Although a difference between the two groups of scores may be caused by the treatment, it is always possible that the scores simply change as time goes by. For example, the depression scores may decrease over time in the same way that the symptoms of a cold disappear over time. In a pre–post study, the researcher also has no control over other variables that change with time. For example, the weather could change from dark and gloomy before therapy to bright and sunny after therapy. In this case, the depression scores could improve because of the weather and not because of the therapy. Because the researcher cannot control the passage of time or other variables related to time, this study is not a true experiment.

Terminology in nonexperimental research Although the two research studies shown in Figure 1.7 are not true experiments, you should notice that they produce the same kind of data that are found in an experiment (see Figure 1.6). In each case, one variable is used to create groups, and a second variable is measured to obtain scores within each group. In an experiment, the groups are created by manipulation of the independent variable, and the participants' scores are the dependent variable. The same terminology is often used to identify the two variables in nonexperimental studies. That is, the variable that is used to create groups is the independent variable and the scores are the dependent variable. For example, the top part of Figure 1.7, gender (boy/girl), is the independent variable and the verbal test scores are the dependent variable. However, you should realize that gender (boy/girl) is not a true independent variable because it is not manipulated. For this reason, the "independent variable" in a nonexperimental study is often called a *quasi-independent variable.*

DEFINITION

In a nonexperimental study, the "independent variable" that is used to create the different groups of scores is often called the **quasi-independent variable**.

DATA STRUCTURES AND STATISTICAL METHODS

The two general data structures that we used to classify research methods can also be used to classify statistical methods.

I. One Group with Two Variables Measured for each Individual Recall that the data from a correlational study consist of two scores, representing two different variables, for each individual. The scores can be listed in a table or displayed in a scatter plot as in Figure 1.5. The relationship between the two variables is usually measured and described using a statistic called a *correlation*. Correlations and the correlational method are discussed in detail in Chapters15 and 16.

Occasionally, the measurement process used for a correlational study simply classifies individuals into categories that do not correspond to numerical values. For example, a researcher could classify a group of college students by gender (male or female) and by cell-phone preference (talk or text). Note that the researcher has two scores for each individual but neither of the scores is a numerical value. This type of data is typically summarized in a table showing how many individuals are classified into each of the possible categories. Table 1.1 shows an example of this kind of summary table. The table shows, for example, that 30 of the males in the sample preferred texting to talking. This type of data can be coded with numbers (for example, male = 0 and female = 1) so that it is possible to compute a correlation. However, the relationship between variables for non-numerical data, such as the data in Table 1.1, is usually evaluated using a statistical technique known as a *chi-square test*. Chi-square tests are presented in Chapter 17.

II. Comparing Two or More Groups of Scores Most of the statistical procedures presented in this book are designed for research studies that compare groups of scores, like the experimental study in Figure 1.6 and the nonexperimental studies in Figure 1.7. Specifically, we examine descriptive statistics that summarize and describe the scores in each group, and we examine inferential statistics that allow us to use the groups, or samples, to generalize to the entire population.

When the measurement procedure produces numerical scores, the statistical evaluation typically involves computing the average score for each group and then comparing the averages. The process of computing averages is presented in Chapter 3, and a variety of statistical techniques for comparing averages are presented in Chapters 8–14. If the measurement process simply classifies individuals into non-numerical categories, the statistical evaluation usually consists of computing proportions for each group and then comparing proportions. In Table 1.1 we present an example of non-numerical data examining the relationship between gender and cell-phone preference. The same data can be used to compare the proportions for males with the proportions for females. For example, using text is preferred by 60% of the males compared to 50% of the females. As mentioned before, these data are evaluated using a chi-square test, which is presented in Chapter 17.

TABLE 1.1

Correlational data consisting of non-numerical scores. Note that there are two measurements for each individual: gender and cell phone preference. The numbers indicate how many people are in each category. For example, out of the 50 males, 30 prefer text over talk.

	Cell Phone Preference		
	Text	Talk	
Males	30	20	50
Females	25	25	50

LEARNING CHECK

1. Researchers have observed that high school students who watched educational television programs as young children tend to have higher grades than their peers who did not watch educational television. Is this study an example of an experiment? Explain why or why not.
2. What two elements are necessary for a research study to be an experiment?
3. Loftus and Palmer (1974) conducted an experiment in which participants were shown a video of an automobile accident. After the video, some participants were

asked to estimate the speed of the cars when they "smashed into" each other. Others were asked to estimate the speed when the cars "hit" each other. The "smashed into" group produced significantly higher estimates than the "hit" group. Identify the independent and dependent variables for this study.

ANSWERS

1. This study could be correlational or nonexperimental, but it is definitely not an example of a true experiment. The researcher is simply observing, not manipulating, the amount of educational television.
2. First, the researcher must manipulate one of the two variables being studied. Second, all other variables that might influence the results must be controlled.
3. The independent variable is the phrasing of the question and the dependent variable is the speed estimated by each participant.

1.4 VARIABLES AND MEASUREMENT

The scores that make up the data from a research study are the result of observing and measuring variables. For example, a researcher may finish a study with a set of IQ scores, personality scores, or reaction-time scores. In this section, we take a closer look at the variables that are being measured and the process of measurement.

CONSTRUCTS AND OPERATIONAL DEFINITIONS

Some variables, such as height, weight, and eye color are well-defined, concrete entities that can be observed and measured directly. On the other hand, many variables studied by behavioral scientists are internal characteristics that people use to help describe and explain behavior. For example, we say that a student does well in school because he or she is *intelligent*. Or we say that someone is *anxious* in social situations, or that someone seems to be *hungry*. Variables like intelligence, anxiety, and hunger are called *constructs*, and because they are intangible and cannot be directly observed, they are often called *hypothetical constructs*.

Although constructs such as intelligence are internal characteristics that cannot be directly observed, it is possible to observe and measure behaviors that are representative of the construct. For example, we cannot "see" intelligence but we can see examples of intelligent behavior. The external behaviors can then be used to create an operational definition for the construct. An *operational definition* defines a construct in terms of external behaviors that can be observed and measured. For example, your intelligence is measured and defined by your performance on an IQ test, or hunger can be measured and defined by the number of hours since last eating.

DEFINITIONS

Constructs are internal attributes or characteristics that cannot be directly observed but are useful for describing and explaining behavior.

An **operational definition** identifies a measurement procedure (a set of operations) for measuring an external behavior and uses the resulting measurements as a definition and a measurement of a hypothetical construct. Note that an operational definition has two components: First, it describes a set of operations for measuring a construct. Second, it defines the construct in terms of the resulting measurements.

DISCRETE AND CONTINUOUS VARIABLES

The variables in a study can be characterized by the type of values that can be assigned to them. A *discrete variable* consists of separate, indivisible categories. For this type of variable, there are no intermediate values between two adjacent categories. Consider the values displayed when dice are rolled. Between neighboring values—for example, five dots and six dots—no other values can ever be observed.

DEFINITION

A **discrete variable** consists of separate, indivisible categories. No values can exist between two neighboring categories.

Discrete variables are commonly restricted to whole, countable numbers—for example, the number of children in a family or the number of students attending class. If you observe class attendance from day to day, you may count 18 students one day and 19 students the next day. However, it is impossible ever to observe a value between 18 and 19. A discrete variable may also consist of observations that differ qualitatively. For example, people can be classified by gender (male or female), by occupation (nurse, teacher, lawyer, etc.), and college students can be classified by academic major (art, biology, chemistry, etc.). In each case, the variable is discrete because it consists of separate, indivisible categories.

On the other hand, many variables are not discrete. Variables such as time, height, and weight are not limited to a fixed set of separate, indivisible categories. You can measure time, for example, in hours, minutes, seconds, or fractions of seconds. These variables are called *continuous* because they can be divided into an infinite number of fractional parts.

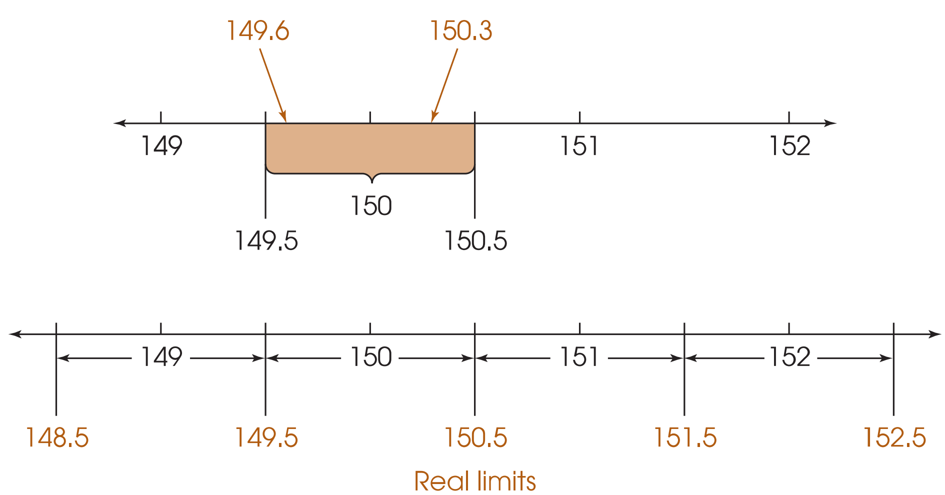
DEFINITION

For a **continuous variable**, there are an infinite number of possible values that fall between any two observed values. A continuous variable is divisible into an infinite number of fractional parts.

Suppose, for example, that a researcher is measuring weights for a group of individuals participating in a diet study. Because weight is a continuous variable, it can be pictured as a continuous line (Figure 1.8). Note that there are an infinite number of possible points on the line without any gaps or separations between neighboring points. For

FIGURE 1.8

When measuring weight to the nearest whole pound, 149.6 and 150.3 are assigned the value of 150 (top). Any value in the interval between 149.5 and 150.5 is given the value of 150.



any two different points on the line, it is always possible to find a third value that is between the two points.

Two other factors apply to continuous variables:

1. When measuring a continuous variable, it should be very rare to obtain identical measurements for two different individuals. Because a continuous variable has an infinite number of possible values, it should be almost impossible for two people to have exactly the same score. If the data show a substantial number of tied scores, then you should suspect that the measurement procedure is very crude or that the variable is not really continuous.
2. When measuring a continuous variable, each measurement category is actually an *interval* that must be defined by boundaries. For example, two people who both claim to weigh 150 pounds are probably not *exactly* the same weight. However, they are both around 150 pounds. One person may actually weigh 149.6 and the other 150.3. Thus, a score of 150 is not a specific point on the scale but instead is an interval (see Figure 1.8). To differentiate a score of 150 from a score of 149 or 151, we must set up boundaries on the scale of measurement. These boundaries are called *real limits* and are positioned exactly halfway between adjacent scores. Thus, a score of $X = 150$ pounds is actually an interval bounded by a *lower real limit* of 149.5 at the bottom and an *upper real limit* of 150.5 at the top. Any individual whose weight falls between these real limits will be assigned a score of $X = 150$.

DEFINITIONS

Real limits are the boundaries of intervals for scores that are represented on a continuous number line. The real limit separating two adjacent scores is located exactly halfway between the scores. Each score has two real limits. The **upper real limit** is at the top of the interval, and the **lower real limit** is at the bottom.

Technical Note: Students often ask whether a value of exactly 150.5 should be assigned to the $X = 150$ interval or the $X = 151$ interval. The answer is that 150.5 is the *boundary* between the two intervals and is not necessarily in one or the other. Instead, the placement of 150.5 depends on the rule that you are using for rounding numbers. If you are rounding up, then 150.5 goes in the higher interval ($X = 151$) but if you are rounding down, then it goes in the lower interval ($X = 150$).

The concept of real limits applies to any measurement of a continuous variable, even when the score categories are not whole numbers. For example, if you were measuring time to the nearest tenth of a second, the measurement categories would be 31.0, 31.1, 31.2, and so on. Each of these categories represents an interval on the scale that is bounded by real limits. For example, a score of $X = 31.1$ seconds indicates that the actual measurement is in an interval bounded by a lower real limit of 31.05 and an upper real limit of 31.15. Remember that the real limits are always halfway between adjacent categories.

Later in this book, real limits are used for constructing graphs and for various calculations with continuous scales. For now, however, you should realize that real limits are a necessity whenever you make measurements of a continuous variable.

Finally, we should warn you that the terms *continuous* and *discrete* apply to the variables that are being measured and not to the scores that are obtained from the measurement. For example, measuring people's heights to the nearest inch produces scores of 60, 61, 62, and so on. Although the scores may appear to be discrete numbers, the underlying variable is continuous. One key to determining whether a variable is continuous or discrete is that a continuous variable can be divided into any number of fractional parts. Height can be measured to the nearest inch, the nearest 0.5 inch, or the nearest 0.1 inch. Similarly, a professor evaluating students' knowledge could use a pass/fail system that classifies students into two broad categories. However, the professor could choose to use a 10-point quiz that divides student knowledge into 11 categories corresponding to quiz scores from 0 to 10. Or the professor could use a 100-point exam that potentially divides student knowledge into 101 categories from 0 to 100. Whenever you are free to choose the degree of precision or the number of categories for measuring a variable, the variable must be continuous.

SCALES OF MEASUREMENT

It should be obvious by now that data collection requires that we make measurements of our observations. Measurement involves assigning individuals or events to categories. The categories can simply be names such as male/female or employed/unemployed, or they can be numerical values such as 68 inches or 175 pounds. The categories used to measure a variable make up a *scale of measurement*, and the relationships between the categories determine different types of scales. The distinctions among the scales are important because they identify the limitations of certain types of measurements and because certain statistical procedures are appropriate for scores that have been measured on some scales but not on others. If you were interested in people's heights, for example, you could measure a group of individuals by simply classifying them into three categories: tall, medium, and short. However, this simple classification would not tell you much about the actual heights of the individuals, and these measurements would not give you enough information to calculate an average height for the group. Although the simple classification would be adequate for some purposes, you would need more sophisticated measurements before you could answer more detailed questions. In this section, we examine four different scales of measurement, beginning with the simplest and moving to the most sophisticated.

THE NOMINAL SCALE

The word *nominal* means "having to do with names." Measurement on a nominal scale involves classifying individuals into categories that have different names but are not related to each other in any systematic way. For example, if you were measuring the academic majors for a group of college students, the categories would be art, biology, business, chemistry, and so on. Each student would be classified in one category according to his or her major. The measurements from a nominal scale allow us to determine whether two individuals are different, but they do not identify either the direction or the size of the difference. If one student is an art major and another is a biology major we can say that they are different, but we cannot say that art is "more than" or "less than" biology and we cannot specify how much difference there is between art and biology. Other examples of nominal scales include classifying people by race, gender, or occupation.

DEFINITION

A **nominal scale** consists of a set of categories that have different names. Measurements on a nominal scale label and categorize observations, but do not make any quantitative distinctions between observations.

Although the categories on a nominal scale are not quantitative values, they are occasionally represented by numbers. For example, the rooms or offices in a building may be identified by numbers. You should realize that the room numbers are simply names and do not reflect any quantitative information. Room 109 is not necessarily bigger than Room 100 and certainly not 9 points bigger. It also is fairly common to use numerical values as a code for nominal categories when data are entered into computer programs. For example, the data from a survey may code males with a 0 and females with a 1. Again, the numerical values are simply names and do not represent any quantitative difference. The scales that follow do reflect an attempt to make quantitative distinctions.

THE ORDINAL SCALE

The categories that make up an *ordinal scale* not only have different names (as in a nominal scale) but also are organized in a fixed order corresponding to differences of magnitude.

DEFINITION

An **ordinal scale** consists of a set of categories that are organized in an ordered sequence. Measurements on an ordinal scale rank observations in terms of size or magnitude.

Often, an ordinal scale consists of a series of ranks (first, second, third, and so on) like the order of finish in a horse race. Occasionally, the categories are identified by verbal labels like small, medium, and large drink sizes at a fast-food restaurant. In either case, the fact that the categories form an ordered sequence means that there is a directional relationship between categories. With measurements from an ordinal scale, you can determine whether two individuals are different and you can determine the direction of difference. However, ordinal measurements do not allow you to determine the size of the difference between two individuals. For example, if Billy is placed in the low-reading group and Tim is placed in the high-reading group, you know that Tim is a better reader, but you do not know how much better. Other examples of ordinal scales include socioeconomic class (upper, middle, lower) and T-shirt sizes (small, medium, large). In addition, ordinal scales are often used to measure variables for which it is difficult to assign numerical scores. For example, people can rank their food preferences but might have trouble explaining "how much" they prefer chocolate ice cream to steak.

THE INTERVAL AND RATIO SCALES

Both an *interval scale* and a *ratio scale* consist of a series of ordered categories (like an ordinal scale) with the additional requirement that the categories form a series of intervals that are all exactly the same size. Thus, the scale of measurement consists of a series of equal intervals, such as inches on a ruler. Other examples of interval and ratio scales are the measurement of time in seconds, weight in pounds, and temperature in degrees Fahrenheit. Note that, in each case, one interval (1 inch, 1 second, 1 pound, 1 degree) is the same size, no matter where it is located on the scale. The fact that the intervals are all the same size makes it possible to determine both the size and the direction of the difference between two measurements. For example, you know that a measurement of 80° Fahrenheit is higher than a measure of 60°, and you know that it is exactly 20° higher.

The factor that differentiates an interval scale from a ratio scale is the nature of the zero point. An interval scale has an arbitrary zero point. That is, the value 0 is assigned to a particular location on the scale simply as a matter of convenience or reference. In particular, a value of zero does not indicate a total absence of the variable being measured. For example a temperature of 0° Fahrenheit does not mean that there is no temperature, and it does not prohibit the temperature from going even lower. Interval scales with an arbitrary zero point are relatively rare. The two most common examples are the Fahrenheit and Celsius temperature scales. Other examples include golf scores (above and below par) and relative measures such as above and below average rainfall.

A ratio scale is anchored by a zero point that is not arbitrary but rather is a meaningful value representing none (a complete absence) of the variable being measured. The existence of an absolute, nonarbitrary zero point means that we can measure the absolute amount of the variable; that is, we can measure the distance from 0. This makes it possible to compare measurements in terms of ratios. For example, an individual who requires 10 seconds to solve a problem (10 more than 0) has taken twice as much time as an individual who finishes in only 5 seconds (5 more than 0). With a ratio scale, we can measure the direction and the size of the difference between two measurements and we can describe the difference in terms of a ratio. Ratio scales are quite common and include physical measures such as height and weight, as well as variables such as reaction time or the number of errors on a test. The distinction between an interval scale and a ratio scale is demonstrated in Example 1.2.

DEFINITIONS

An **interval scale** consists of ordered categories that are all intervals of exactly the same size. Equal differences between numbers on a scale reflect equal differences in magnitude. However, the zero point on an interval scale is arbitrary and does not indicate a zero amount of the variable being measured.

A **ratio scale** is an interval scale with the additional feature of an absolute zero point. With a ratio scale, ratios of numbers do reflect ratios of magnitude.

EXAMPLE 1.2

A researcher obtains measurements of height for a group of 8-year-old boys. Initially, the researcher simply records each child's height in inches, obtaining values such as 44, 51, 49, and so on. These initial measurements constitute a ratio scale. A value of zero represents no height (absolute zero). Also, it is possible to use these measurements to form ratios. For example, a child who is 60 inches tall is one-and-a-half times taller than a child who is 40 inches tall.

Now suppose that the researcher converts the initial measurement into a new scale by calculating the difference between each child's actual height and the average height for this age group. A child who is 1 inch taller than average now gets a score of +1; a child 4 inches taller than average gets a score of +4. Similarly, a child who is 2 inches shorter than average gets a score of −2. On this scale, a score of zero corresponds to average height. Because zero no longer indicates a complete absence of height, the new scores constitute an interval scale of measurement.

Notice that original scores and the converted scores both involve measurement in inches, and you can compute differences, or distances, on either scale. For example, there is a 6-inch difference in height between two boys who measure 57 and 51 inches tall on the first scale. Likewise, there is a 6-inch difference between two boys who measure +9 and +3 on the second scale. However, you should also notice that ratio comparisons are not possible on the second scale. For example, a boy who measures +9 is not three times taller than a boy who measures +3.

STATISTICS AND SCALES OF MEASUREMENT

For our purposes, scales of measurement are important because they influence the kind of statistics that can and cannot be used. For example, if you measure IQ scores for a group of students, it is possible to add the scores together and calculate a mean score for the group. On the other hand, if you measure the academic major for each student, you cannot compute the mean. (What is the mean of three psychology majors, an English major, and two chemistry majors?) The vast majority of the statistical techniques presented in this book are designed for numerical scores from an interval or a ratio scale. For most statistical applications, the distinction between an interval scale and a ratio scale is not important because both scales produce numerical values that permit us to compute differences between scores, to add scores, and to calculate mean scores. On the other hand, measurements from nominal or ordinal scales are typically not numerical values and are not compatible with many basic arithmetic operations. Therefore, alternative statistical techniques are necessary for data from nominal or ordinal scales of measurement (for example, the median and the mode in Chapter 3, the Spearman correlation in Chapter 15, and the chi-square tests in Chapter 17). Additional statistical methods for measurements from ordinal scales are presented in Appendix E.

LEARNING CHECK

1. A survey asks people to identify their age, annual income, and marital status (single, married, divorced, etc.). For each of these three variables, identify the scale of measurement that probably is used and identify whether the variable is continuous or discrete.
2. An English professor uses letter grades (A, B, C, D, and F) to evaluate a set of student essays. What kind of scale is being used to measure the quality of the essays?
3. The teacher in a communications class asks students to identify their favorite reality television show. The different television shows make up a ______ scale of measurement.
4. A researcher studies the factors that determine the number of children that couples decide to have. The variable, number of children, is a ______________ (discrete/continuous) variable.
5. **a.** When measuring height to the nearest inch, what are the real limits for a score of 68 inches?
   b. When measuring height to the nearest half inch, what are the real limits for a score of 68 inches?

ANSWERS

1. Age and annual income are measured on ratio scales and are both continuous variables. Marital status is measured on a nominal scale and is a discrete variable.
2. ordinal
3. nominal
4. discrete
5. **a.** 67.5 and 68.5
   b. 67.75 and 68.25

1.5 STATISTICAL NOTATION

The measurements obtained in research studies provide the data for statistical analysis. Most of the statistical analyses use the same general mathematical operations, notation, and basic arithmetic that you have learned during previous years of school. In case you are unsure of your mathematical skills, there is a mathematics review section in Appendix A at the back of this book. The appendix also includes a skills-assessment exam (p. 678) to help you determine whether you need the basic mathematics review. In this section, we introduce some of the specialized notation that is used for statistical calculations. In later chapters, additional statistical notation is introduced as it is needed.

Measuring a variable in a research study typically yields a value or a score for each individual. Raw scores are the original, unchanged scores obtained in the study. Scores for a particular variable are represented by the letter X. For example, if performance in your statistics course is measured by tests and you obtain a 35 on the first test, then we could state that $X = 35$. A set of scores can be presented in a column that is headed by X. For example, a list of quiz scores from your class might be presented as shown in the margin (the single column on the left).

X
37
35
35
30
25
17
16

Scores	
X	Y
72	165
68	151
67	160
67	160
68	146
70	160
66	133

When observations are made for two variables, there will be two scores for each individual. The data can be presented as two lists labeled X and Y for the two variables. For example, observations for people's height in inches (variable X) and weight in pounds (variable Y) can be presented as shown in the double column in the margin. Each pair X, Y represents the observations made of a single participant.

The letter N is used to specify how many scores are in a set. An uppercase letter N identifies the number of scores in a population and a lowercase letter n identifies the number of scores in a sample. Throughout the remainder of the book you will notice that we often use notational differences to distinguish between samples and populations. For the height and weight data in the preceding table, $n = 7$ for both variables. Note that by using a lowercase letter n, we are implying that these data are a sample.

SUMMATION NOTATION

Many of the computations required in statistics involve adding a set of scores. Because this procedure is used so frequently, a special notation is used to refer to the sum of a set of scores. The Greek letter *sigma*, or Σ, is used to stand for summation. The expression ΣX means to add all the scores for variable X. The summation sign, Σ, can be read as "the sum of." Thus, ΣX is read "the sum of the scores." For the following set of quiz scores,

10, 6, 7, 4

$\Sigma X = 27$ and $N = 4$.

To use summation notation correctly, keep in mind the following two points:

1. The summation sign, Σ, is always followed by a symbol or mathematical expression. The symbol or expression identifies exactly which values are to be added. To compute ΣX, for example, the symbol following the summation sign is X, and the task is to find the sum of the X values. On the other hand, to compute $\Sigma(X - 1)^2$, the summation sign is followed by a relatively complex mathematical expression, so your first task is to calculate all of the $(X - 1)^2$ values and then add the results.
2. The summation process is often included with several other mathematical operations, such as multiplication or squaring. To obtain the correct answer, it is essential that the different operations be done in the correct sequence. Following is a list showing the correct *order of operations* for performing mathematical operations. Most of this list should be familiar, but you should note that we have inserted the summation process as the fourth operation in the list.

Order of Mathematical Operations

More information on the order of operations for mathematics is available in the Math Review appendix, page 679.

1. Any calculation contained within parentheses is done first.
2. Squaring (or raising to other exponents) is done second.
3. Multiplying and/or dividing is done third. A series of multiplication and/or division operations should be done in order from left to right.
4. Summation using the Σ notation is done next.
5. Finally, any other addition and/or subtraction is done.

The following examples demonstrate how summation notation is used in most of the calculations and formulas we present in this book.

EXAMPLE 1.3

A set of four scores consists of values 3, 1, 7, and 4. We will compute ΣX, ΣX^2, and $(\Sigma X)^2$ for these scores. To help demonstrate the calculations, we will use a computational table showing the original scores (the X values) in the first column. Additional columns can then be added to show additional steps in the series of operations. You should notice that the first three operations in the list (parentheses, squaring, and multiplying) all create a new column of values. The last two operations, however, produce a single value corresponding to the sum.

X	X^2
3	9
1	1
7	49
4	16

The table to the left shows the original scores (the X values) and the squared scores (the X^2 values) that are needed to compute ΣX^2.

The first calculation, ΣX, does not include any parentheses, squaring, or multiplication, so we go directly to the summation operation. The X values are listed in the first column of the table, and we simply add the values in this column:

$$\Sigma X = 3 + 1 + 7 + 4 = 15$$

To compute ΣX^2, the correct order of operations is to square each score and then find the sum of the squared values. The computational table shows the original scores and the results obtained from squaring (the first step in the calculation). The second step is to find the sum of the squared values, so we simply add the numbers in the X^2 column.

$$\Sigma X^2 = 9 + 1 + 49 + 16 = 75$$

The final calculation, $(\Sigma X)^2$, includes parentheses, so the first step is to perform the calculation inside the parentheses. Thus, we first find ΣX and then square this sum. Earlier, we computed $\Sigma X = 15$, so

$$(\Sigma X)^2 = (15)^2 = 225$$

EXAMPLE 1.4

Use the same set of four scores from Example 1.3 and compute $\Sigma(X - 1)$ and $\Sigma(X - 1)^2$. The following computational table will help demonstrate the calculations.

X	$(X - 1)$	$(X - 1)^2$
3	2	4
1	0	0
7	6	36
4	3	9

The first column lists the original scores. A second column lists the $(X - 1)$ values, and a third column shows the $(X - 1)^2$ values.

To compute $\Sigma(X - 1)$, the first step is to perform the operation inside the parentheses. Thus, we begin by subtracting one point from each of the X values. The resulting values are listed in the middle column of the table. The next step is to add the $(X - 1)$ values.

$$\Sigma(X - 1) = 2 + 0 + 6 + 3 + = 11$$

The calculation of $\Sigma(X + 1)^2$ requires three steps. The first step (inside parentheses) is to subtract 1 point from each X value. The results from this step are shown in the middle column of the computational table. The second step is to square each of the

$(X - 1)$ values. The results from this step are shown in the third column of the table. The final step is to add the $(X - 1)^2$ values to obtain

$$\Sigma(X - 1)^2 = 4 + 0 + 36 + 9 = 49$$

Notice that this calculation requires squaring before adding. A common mistake is to add the $(X - 1)$ values and then square the total. Be careful!

EXAMPLE 1.5

In both of the preceding examples, and in many other situations, the summation operation is the last step in the calculation. According to the order of operations, parentheses, exponents, and multiplication all come before summation. However, there are situations in which extra addition and subtraction are completed after the summation. For this example, use the same scores that appeared in the previous two examples, and compute $\Sigma X - 1$.

With no parentheses, exponents, or multiplication, the first step is the summation. Thus, we begin by computing ΣX. Earlier we found $\Sigma X = 15$. The next step is to subtract one point from the total. For these data,

$$\Sigma X - 1 = 15 - 1 = 14$$

EXAMPLE 1.6

For this example, each individual has two scores. The first score is identified as X, and the second score is Y. With the help of the following computational table, compute ΣX, ΣY, and ΣXY.

Person	X	Y	XY
A	3	5	15
B	1	3	3
C	7	4	28
D	4	2	8

To find ΣX, simply add the values in the X column.

$$\Sigma X = 3 + 1 + 7 + 4 = 15$$

Similarly, ΣY is the sum of the Y values.

$$\Sigma Y = 5 + 3 + 4 + 2 = 14$$

To compute ΣXY, the first step is to multiply X times Y for each individual. The resulting products (XY values) are listed in the third column of the table. Finally, we add the products to obtain

$$\Sigma XY = 15 + 3 + 28 + 8 = 54$$

LEARNING CHECK

1. Calculate each value requested for the following scores: 6, 2, 4, 2.
   - **a.** ΣX
   - **b.** ΣX^2
   - **c.** $(\Sigma X)^2$
   - **d.** $\Sigma(X - 2)$
   - **e.** $\Sigma(X - 2)^2$
2. Identify the first step in each of the following calculations.
   a. ΣX^2 **b.** $(\Sigma X)^2$ **c.** $\Sigma(X - 2)^2$
3. Use summation notation to express each of the following.
   - **a.** Add 4 points to each score and then add the resulting values.
   - **b.** Add the scores and then square the total.
   - **c.** Square each score, then add the squared values.

ANSWERS

1. **a.** 14
   b. 60
   c. 196
   d. 6
   e. 20
2. **a.** Square each score.
   b. Add the scores.
   c. Subtract 2 points from each score.
3. **a.** $\Sigma(X + 4)$
   b. $(\Sigma X)^2$
   c. ΣX^2

SUMMARY

1. The term *statistics* is used to refer to methods for organizing, summarizing, and interpreting data.
2. Scientific questions usually concern a population, which is the entire set of individuals one wishes to study. Usually, populations are so large that it is impossible to examine every individual, so most research is conducted with samples. A sample is a group selected from a population, usually for purposes of a research study.
3. A characteristic that describes a sample is called a statistic, and a characteristic that describes a population is called a parameter. Although sample statistics are usually representative of corresponding population parameters, there is typically some discrepancy between a statistic and a parameter. The naturally occurring difference between a statistic and a parameter is called sampling error.
4. Statistical methods can be classified into two broad categories: descriptive statistics, which organize and summarize data, and inferential statistics, which use sample data to draw inferences about populations.
5. The correlational method examines relationships between variables by measuring two different variables for each individual. This method allows researchers to measure and describe relationships, but cannot produce a cause-and-effect explanation for the relationship.
6. The experimental method examines relationships between variables by manipulating an independent variable to create different treatment conditions and then measuring a dependent variable to obtain a group of scores in each condition. The groups of scores are then compared. A systematic difference between groups provides evidence that changing the independent variable from one condition to another

also caused a change in the dependent variable. All other variables are controlled to prevent them from influencing the relationship. The intent of the experimental method is to demonstrate a cause-and-effect relationship between variables.

7. Nonexperimental studies also examine relationships between variables by comparing groups of scores, but they do not have the rigor of true experiments and cannot produce cause-and-effect explanations. Instead of manipulating a variable to create different groups, a nonexperimental study uses a preexisting participant characteristic (such as male/female) or the passage of time (before/after) to create the groups being compared.

8. A measurement scale consists of a set of categories that are used to classify individuals. A nominal scale consists of categories that differ only in name and are not differentiated in terms of magnitude or direction. In an ordinal scale, the categories are differentiated in terms of direction, forming an ordered series. An interval scale consists of an ordered series of categories that are all equal-sized intervals. With an interval scale, it is possible to differentiate direction and magnitude (or distance) between categories. Finally, a ratio scale is an interval scale for which the zero point indicates none of the variable being measured. With a ratio scale, ratios of measurements reflect ratios of magnitude.

9. A discrete variable consists of indivisible categories, often whole numbers that vary in countable steps. A continuous variable consists of categories that are infinitely divisible and each score corresponds to an interval on the scale. The boundaries that separate intervals are called real limits and are located exactly halfway between adjacent scores.

10. The letter X is used to represent scores for a variable. If a second variable is used, Y represents its scores. The letter N is used as the symbol for the number of scores in a population; n is the symbol for a number of scores in a sample.

11. The Greek letter sigma (Σ) is used to stand for summation. Therefore, the expression ΣX is read "the sum of the scores." Summation is a mathematical operation (like addition or multiplication) and must be performed in its proper place in the order of operations; summation occurs after parentheses, exponents, and multiplying/dividing have been completed.

KEY TERMS

statistics (5)
population (5)
sample (6)
variable (6)
data (7)
data set (7)
datum (7)
raw score (7)
parameter (7)
statistic (7)
descriptive statistics (7)
inferential statistics (8)
sampling error (8)
correlational method (13)
experimental method (14)
independent variable (16)
dependent variable (16)
control condition (16)
experimental condition (16)
nonequivalent groups study (17)
pre–post study (18)
quasi-independent variable (18)
construct (20)
operational definition (20)
discrete variable (21)
continuous variable (21)
real limits (22)
upper real limit (22)
lower real limit (22)
nominal scale (23)
ordinal scale (23)
interval scale (25)
ratio scale (25)

RESOURCES

Book Companion Website: www.cengage.com/psychology/gravetter

You can find practice quizzes and other learning aids for every chapter in this book on the book companion website, as well as a series of workshops and other resources corresponding to the main topic areas. In the left-hand column are a variety of learning exercises for Chapter 1, including a tutorial quiz. Also in the left-hand column, under

Book Resources, is a link to the workshops. For Chapter 1, there is a workshop that reviews the scales of measurement. To get there, click on the *Workshop* link, then click on *Scales of Measurement*. To find materials for other chapters, begin by selecting the desired chapter at the top of the page. Note that the workshops were not developed specifically for this book but are used by several different books written by different authors. As a result, you may find that some of the notation or terminology is different from that which you learned in this text.

At the end of each chapter we remind you about the Web resources. Again, there is a tutorial quiz for every chapter, and we notify you whenever there is a workshop that is related to the chapter content.

aplia

Improve your understanding of statistics with Aplia's auto-graded problem sets and immediate, detailed explanations for every question. To learn more, visit **www.aplia.com/statistics**.

CENGAGEbrain.com

Psychology CourseMate brings course concepts to life with interactive learning, study, and exam preparation tools that support the printed textbook. A textbook-specific website, Psychology CourseMate includes an integrated interactive eBook and other interactive learning tools including quizzes, flashcards, and more.

Visit www.cengagebrain.com to access your account and purchase materials.

SPSS

The Statistical Package for the Social Sciences, known as SPSS, is a computer program that performs most of the statistical calculations that are presented in this book, and is commonly available on college and university computer systems. Appendix D contains a general introduction to SPSS. In the Resource section at the end of each chapter for which SPSS is applicable, there are step-by-step instructions for using SPSS to perform the statistical operations presented in the chapter.

FOCUS ON PROBLEM SOLVING

It may help to simplify summation notation if you observe that the summation sign is always followed by a symbol or symbolic expression—for example, ΣX or $\Sigma(X + 3)$. This symbol specifies which values you are to add. If you use the symbol as a column heading and list all the appropriate values in the column, your task is simply to add up the numbers in the column. To find $\Sigma(X + 3)$ for example, start a column headed with $(X + 3)$ next to the column of Xs. List all the $(X + 3)$ values; then find the total for the column.

Often, summation notation is part of a relatively complex mathematical expression that requires several steps of calculation. The series of steps must be performed according to the order of mathematical operations (see page 27). The best procedure is to use a computational table that begins with the original X values listed in the first column. Except for summation, each step in the calculation creates a new column of values. For example, computing $\Sigma(X + 1)^2$ involves three steps and produces a computational table with three columns. The final step is to add the values in the third column (see Example 1.4).

DEMONSTRATION 1.1

SUMMATION NOTATION

A set of scores consists of the following values:

7 3 9 5 4

For these scores, compute each of the following:

ΣX
$(\Sigma X)^2$
ΣX^2
$\Sigma X + 5$
$\Sigma(X - 2)$

Compute ΣX To compute ΣX, we simply add all of the scores in the group.

$$\Sigma X = 7 + 3 + 9 + 5 + 4 = 28$$

Compute $(\Sigma X)^2$ The first step, inside the parentheses, is to compute ΣX. The second step is to square the value for ΣX.

$$\Sigma X = 28 \text{ and } (\Sigma X)^2 = (28)^2 = 784$$

X	X^2
7	49
3	9
9	81
5	25
4	16

Compute ΣX^2 The first step is to square each score. The second step is to add the squared scores. The computational table shows the scores and squared scores. To compute ΣX^2 we add the values in the X^2 column.

$$\Sigma X^2 = 49 + 9 + 81 + 25 + 16 = 180$$

Compute $\Sigma X + 5$ The first step is to compute ΣX. The second step is to add 5 points to the total.

$$\Sigma X = 28 \text{ and } \Sigma X + 5 = 28 + 5 = 33$$

X	$X - 2$
7	5
3	1
9	7
5	3
4	2

Compute $\Sigma(X - 2)$ The first step, inside parentheses, is to subtract 2 points from each score. The second step is to add the resulting values. The computational table shows the scores and the $(X - 2)$ values. To compute $\Sigma(X - 2)$, add the values in the $(X - 2)$ column

$$\Sigma(X - 2) = 5 + 1 + 7 + 3 + 2 = 18$$

PROBLEMS

*1. A researcher is investigating the effectiveness of a treatment for adolescent boys who are taking medication for depression. A group of 30 boys is selected and half receive the new treatment in addition to their medication and the other half continue to take their medication without any treatment. For this study,
   a. Identify the population.
   b. Identify the sample.

2. Define the terms parameter and statistic. Be sure that the concepts of population and sample are included in your definitions.

3. Statistical methods are classified into two major categories: descriptive and inferential. Describe the general purpose for the statistical methods in each category.

4. A researcher plans to compare two treatment conditions by measuring one sample in treatment 1 and a second sample in treatment 2. The researcher then compares the scores for the two treatments and finds a difference between the two groups.
   a. Briefly explain how the difference may have been caused by the treatments.
   b. Briefly explain how the difference simply may be sampling error.

5. Describe the data for a correlational research study. Explain how these data are different from the data obtained in experimental and nonexperimental studies, which also evaluate relationships between two variables.

6. Describe how the goal of an experimental research study is different from the goal for nonexperimental or correlational research. Identify the two elements that are necessary for an experiment to achieve its goal.

7. Strack, Martin, and Stepper (1988) found that people rated cartoons as funnier when holding a pen in their teeth (which forced them to smile) than when holding a pen in their lips (which forced them to frown). For this study, identify the independent variable and the dependent variable.

8. Judge and Cable (2010) found that thin women had higher incomes than heavier women. Is this an example of an experimental or a nonexperimental study?

9. Two researchers are both interested in the relationship between caffeine consumption and activity level for elementary school children. Each obtains a sample of $n = 20$ children.
   a. The first researcher interviews each child to determine the level of caffeine consumption. The researcher then records the level of activity for each child during a 30-minute session on the playground. Is this an experimental or a nonexperimental study? Explain your answer.
   b. The second researcher separates the children into two roughly equivalent groups. The children in one group are given a drink containing 300 mg of caffeine and the other group gets a drink with no caffeine. The researcher then records the level of activity for each child during a 30-minute session on the playground. Is this an experimental or a nonexperimental study? Explain your answer.

10. A researcher would like to evaluate the claim that large doses of vitamin C can help prevent the common cold. One group of participants is given a large dose of the vitamin (500 mg per day), and a second group is given a placebo (sugar pill). The researcher records the number of colds each individual experiences during the 3-month winter season.
   a. Identify the dependent variable for this study.
   b. Is the dependent variable discrete or continuous?
   c. What scale of measurement (nominal, ordinal, interval, or ratio) is used to measure the dependent variable?

11. A research study comparing alcohol use for college students in the United States and Canada reports that more Canadian students drink but American students drink more (Kuo, Adlaf, Lee, Gliksman, Demers, and Wechsler, 2002). Is this study an example of an experiment? Explain why or why not.

12. Oxytocin is a naturally occurring brain chemical that is nicknamed the "love hormone" because it seems to play a role in the formation of social relationships such as mating pairs and parent-child bonding. A recent study demonstrated that oxytocin appears to increase people's tendency to trust others (Kosfeld, Heinrichs, Zak, Fischbacher, and Fehr, 2005). Using an investment game, the study demonstrated that people who inhaled oxytocin were more likely to give their money to a trustee compared to people who inhaled an inactive placebo. For this experimental study, identify the independent variable and the dependent variable.

13. For each of the following, determine whether the variable being measured is discrete or continuous and explain your answer.
   a. Social networking (number of daily minutes on Facebook)
   b. Family size (number of siblings)

*Solutions for odd-numbered problems are provided in Appendix C.

c. Preference between digital or analog watch
d. Number of correct answers on a statistics quiz

14. Four scales of measurement were introduced in this chapter: nominal, ordinal, interval, and ratio.
a. What additional information is obtained from measurements on an ordinal scale compared to measurements on a nominal scale?
b. What additional information is obtained from measurements on an interval scale compared to measurements on an ordinal scale?
c. What additional information is obtained from measurements on a ratio scale compared to measurements on an interval scale?

15. In an experiment examining the effects of humor on memory, Schmidt (1994) showed participants a list of sentences, half of which were humorous and half were nonhumorous. The participants consistently recalled more of the humorous sentences than the nonhumorous sentences.
a. Identify the independent variable for this study.
b. What scale of measurement is used for the independent variable?
c. Identify the dependent variable for this study.
d. What scale of measurement is used for the dependent variable?

16. Explain why *shyness* is a hypothetical construct instead of a concrete variable. Describe how shyness might be measured and defined using an operational definition.

17. Ford and Torok (2008) found that motivational signs were effective in increasing physical activity on a college campus. Signs such as "Step up to a healthier lifestyle" and "An average person burns 10 calories a minute walking up the stairs" were posted by the elevators and stairs in a college building. Students and faculty increased their use of the stairs during times that the signs were posted compared to times when there were no signs.
a. Identify the independent and dependent variables for this study.
b. What scale of measurement is used for the independent variable?

18. For the following scores, find the value of each expression:
a. ΣX
b. ΣX^2
c. $(\Sigma X)^2$
d. $\Sigma(X - 1)$

X
4
2
1
5

19. For the following set of scores, find the value of each expression:
a. ΣX
b. ΣX^2
c. $\Sigma(X + 1)$
d. $\Sigma(X + 1)^2$

X
4
6
0
3
2

20. For the following set of scores, find the value of each expression:
a. ΣX
b. ΣX^2
c. $\Sigma(X + 4)$

X
−4
−2
0
−1
−1

21. Two scores, X and Y, are recorded for each of $n = 4$ subjects. For these scores, find the value of each expression.
a. ΣX
b. ΣY
c. ΣXY

Subject	X	Y
A	6	4
B	0	10
C	3	8
D	2	3

22. Use summation notation to express each of the following calculations:
a. Add 1 point to each score, then add the resulting values.
b. Add 1 point to each score and square the result, then add the squared values.
c. Add the scores and square the sum, then subtract 3 points from the squared value.

23. For the following set of scores, find the value of each expression:
a. ΣX^2
b. $(\Sigma X)^2$
c. $\Sigma(X - 2)$
d. $\Sigma(X - 2)^2$

X
1
0
5
2