Chapter 1
Introduction

PREVIEW

Why Study Statistics?

What are four important reasons why knowledge of statistics is essential for anyone majoring in psychology, sociology, or education?

Descriptive and Inferential Statistics

What is the difference between descriptive and inferential statistics?

Why must behavioral science researchers use inferential statistics?

Populations, Samples, Parameters, and Statistics

What is the difference between a population and a sample?

Why is it important to specify clearly the population from which a sample is drawn?

What is the difference between a parameter and a statistic?

Measurement Scales

What types of scales are used to measure variables in the behavioral sciences?

Independent and Dependent Variables

What is the difference between observational and experimental studies?

Summation Notation

Why is summation notation used by statisticians?

What are the eight rules involving summation notation?

Ihno's Study

An example that provides a common thread tying together all of the subsequent chapters

Summary

Exercises

Thought Questions

(continued)

PREVIEW (*continued*)

Computer Exercises

Bridge to SPSS

Why Study Statistics?

This book is written primarily for undergraduates majoring in psychology, or one of the other behavioral sciences. There are four reasons why a knowledge of statistics is essential for those who wish to conduct or consume behavioral science research:

1. ***To understand the professional literature.*** Most professional literature in the behavioral sciences includes results that are based on statistical analyses. Therefore, you will be unable to understand important articles in scientific journals and books unless you understand statistics. It is possible to seek out secondhand reports that are designed for the statistically uninformed, but those who prefer this alternative to obtaining firsthand information probably should not be majoring in the fields of behavioral science.
2. ***To understand and evaluate statistical claims made in the popular media.*** This is a relatively new reason to acquire quantitative reasoning skills: The ''Information Age'' of the early 21st century has ushered in an explosion of quantitative claims in newspapers, television, and the Internet. Unfortunately, the difference between claims that are statistically sound and those that merely appear that way can be difficult to detect without some formal training. You will make many important life decisions over the next decade that will require weighing probabilities under conditions of uncertainty. A competence with basic statistics will empower you to make maximal use of the available information and help protect you from those who may wish to mislead you with pretty graphs and numbers.
3. ***To understand the rationale underlying research in the behavioral sciences.*** Statistics is not just a catalog of procedures and formulas. It offers the rationale on which much of behavioral science research is based—namely, drawing inferences about a population based on data obtained from a sample. Those familiar with statistics understand that research consists of a series of educated guesses and fallible decisions, not right or wrong answers. Those without a knowledge of statistics, by contrast, cannot understand the strengths and weaknesses of the techniques used by behavioral scientists to collect information and draw conclusions.
4. ***To carry out behavioral science research.*** In order to contribute competent research to the behavioral sciences, it is necessary to design

the statistical analysis *before* the data are collected. Otherwise, the research procedures may be so poorly planned that not even an expert statistician can make any sense out of the results. To be sure, it is possible (and often advisable) to consult someone more experienced in statistics for assistance. Without some statistical knowledge of your own, however, you will find it difficult or impossible to convey your needs to someone else and to understand the replies.

Save for these introductory remarks, we do not regard it as our task to persuade you that statistics is important in psychology and other behavioral sciences. If you are seriously interested in any of these fields, you will find this out for yourself. Accordingly, this book does not rely on documented examples selected from the professional literature to prove to you that statistics really is used in these fields. Instead, we have devised several artificial, but realistic, examples with numerical values that reveal the processes and issues involved in statistical analyses as clearly as possible.

One example we use throughout this text is based on a hypothetical study performed by a new student named Carrie on the relative friendliness of four dormitory halls on her campus. We present the data for this study in the first exercise at the end of this chapter and return to it in many of the subsequent chapters.

We have tried to avoid a ''cookbook'' approach that places excessive emphasis on computational recipes. Instead, the various statistical procedures and the essential underlying concepts have been explained at length, and insofar as possible in standard English, so that you will know not only what to do but *why* you are doing it. Do not, however, expect to learn the material in this book from a single reading; the concepts involved in statistics, especially inferential statistics, are sufficiently challenging that it is often said that the only way to completely understand the material is to teach it (or write a book about it). Having said that, however, there is no reason to approach statistics with fear and trembling. You certainly do not need any advanced understanding of mathematics to obtain a good working knowledge of basic statistics. What *is* needed is mathematical comprehension sufficient to cope with beginning high school algebra and a willingness to work at new concepts until they are understood, which requires, in turn, a willingness to spend some time working through at least half of the exercises at the end of each chapter (and, perhaps, some additional exercises from the online Study Guide).

Descriptive and Inferential Statistics

One purpose of statistics is to summarize or describe the characteristics of a set of data in a clear and convenient fashion. This is accomplished by what are called **descriptive statistics**. For example, your grade point average serves as a convenient summary of all of the grades that you have received in college. Part I of this book is devoted to descriptive statistics.

A second function of statistics is to make possible the solution of an extremely important problem. Behavioral scientists can never measure *all* of the cases in which they are interested. For example, a clinical psychologist studying the effects of various kinds of therapies cannot obtain data on every single mental health patient in the world; a social psychologist studying gender differences in attitudes cannot measure all of the millions of men and women in the United States; a cognitive psychologist cannot observe the reading behavior of all literate adults. Behavioral scientists want to know what is happening in a given **population**—a large group (theoretically an infinitely large group) of people, animals, objects, or responses that are alike in at least one respect (e.g., all women in the United States). They cannot measure the entire population, however, because it is so large that it would be much too time-consuming and expensive to do so. What to do?

Turns out there's a reasonably simple solution: Measure just a relatively small number of cases drawn from the population (i.e., a **sample**), and use inferential statistics to make educated guesses about the population. **Inferential statistics** makes it possible to draw inferences about what is happening in the population based on what is observed in a sample from that population. (This point is discussed at greater length in Chapter 5.) The subsequent parts of this book are devoted to inferential statistics, which makes frequent use of some of the descriptive statistics discussed in Part I.

Populations, Samples, Parameters, and Statistics

As the above discussion indicates, the term **population** as used in statistics does not necessarily refer to people. For example, the population of interest may be that of all white rats of a given genetic strain or all responses of a single subject's eyelid in a conditioning experiment.

Whereas the population consists of all of the cases of interest, a **sample** consists of any subgroup drawn from the specified population. It is important that the population be clearly specified. For example, a group of 100 Macalester College freshmen might be a well-drawn sample from the population of all Macalester freshmen or a poorly drawn sample from the population of all undergraduates in the United States (poorly drawn because it probably will not be representative of all U.S. undergraduates). It is strictly proper to apply (i.e., *generalize*) the research results only to the specified population from which the sample was drawn. (A researcher *may* justifiably argue that her results are more widely generalizable, but she is on her own if she does so, because the rules of statistical inference do not justify this.)

A **statistic** is a numerical quantity (such as an average) that summarizes some characteristic of a sample. A **parameter** is the corresponding value of that characteristic in the population. For example, if the average studying time of a sample of 100 New York University freshmen is 7.4 hours per week, then 7.4 is a statistic. If the average studying time

of the population of all NYU freshmen is 9.6 hours per week, then 9.6 is the corresponding population parameter. Usually the values of population parameters are unknown because the population is too large to measure in its entirety, and appropriate techniques of inferential statistics are therefore used to estimate the values of population parameters from sample statistics. If the sample is properly selected, the sample statistics will often give good estimates of the parameters of the population from which the sample was drawn; if the sample is poorly chosen, erroneous conclusions are likely to occur. Whether you are doing your own research or reading about that of someone else, you should always check to be sure that the population to which the results are generalized is proper in light of the sample from which the results were obtained.

Measurement Scales

You may have noticed that we have used the term *data* several times without talking about where the data come from. It should come as no surprise that in the behavioral sciences, the data generally come from measuring some aspect of the behavior of a human or animal. Unlike physics, in which there are quite a few important **constants** (values that are always the same, such as the speed of light or the mass of an electron), the behavioral sciences deal mainly with the measurement of **variables**, which can take on a range of different values. An additional complication faced by the behavioral scientist is that some of the variables of most interest are difficult to measure (e.g., anxiety). In this section we discuss the measurement scales most commonly used in the behavioral sciences. It is important to know which scale you are using because that choice often determines which statistical technique is appropriate. The scales differ with respect to how finely they can distinguish differences among instances. For example, the nominal scale can only distinguish whether an item is in one category or another, but the categories have no inherent order (e.g., the color of your iPhone). The ordinal scale involves measurements that can distinguish order on a set of values (e.g., bigger, smaller). Interval scales add the capability to measure quantities on a scale that have equal intervals between units (e.g., inches; temperature), and ratio scales are interval scales for which a value of zero means that *absolutely none* of the variable being measured is present. We describe them below in order of complexity.

Nominal Scales

The crudest form of measurement is to classify items by assigning names to them (categorization), which does not involve any numerical precision at all. Such a scale is called a **nominal scale**. For example, a person's occupation can only be ''measured'' on a nominal scale (e.g., accountant, lawyer, carpenter, sales clerk). We can count the number of people who

fall into each category, but (unlike ordinal, interval, or ratio scales) there is no obvious order to the categories, and certainly no regular intervals between them. We refer to such categorical data as being **qualitative**, as distinguished from data measured on one of the **quantitative** scales described next.

Ordinal Scales

Sometimes it is possible to order your categories, even though the intervals are not precise. The most common example of this in psychological research is called a Likert scale (after its creator, Rensis Likert), on which respondents rate their agreement with some statement by choosing, for instance, among ''strongly agree,'' ''agree,'' ''uncertain,'' ''disagree,'' and ''strongly disagree.'' The order of the categories is clear, but because there is no way to be sure that they are equally spaced (is the psychological distance between ''strongly agree'' and ''agree'' the same as between ''agree'' and ''uncertain''?), this type of scale lacks the interval property and is therefore called an **ordinal scale**. Although it is a somewhat controversial practice, many behavioral researchers simply assign numbers to the categories (e.g., strongly agree is 1, agree is 2, etc.) and then treat the data as though they came from an interval scale.

Another, less common, way that an ordinal scale can be created is by rank ordering. It may not be possible to measure, in a precise way, the creativity of paintings produced by students in an art class, but a panel of judges could rank them from most to least creative, with perhaps a few paintings tied at the same rank.

It is important to distinguish between the *characteristics of the variable* we are measuring, on one hand, and the particular *scale with which we choose to measure that variable*, on the other. For example, suppose a teacher wishes to measure how often middle school students raise their hands in class. Hand-raising frequency is a variable that can be measured quantitatively by just counting how often each student raises a hand. However, when the teacher uses these data to understand hand-raising behavior in his class, his purposes may be best served by organizing the students into three broad, but ordinal, groups: (1) those who never raised their hand (most of the group); (2) those who did so just a few times; and (3) those who did so many times. In that case, he will want to create an ordinal scale to present data that was originally measured on an interval scale (described next). In other words, just because you are using a variable that has the *potential* to be measured on a particular scale does not necessarily mean that the data you collect will possess or display the characteristics of that scale. Sometimes researchers have quantitative measurements that vary in such an odd way that it becomes more useful just to rank them all and use the ranks in place of the original measurements. We explain this somewhat unusual practice when we describe statistical tests based on ordinal data in Chapter 8.

Interval/Ratio Scales

The most precise scales are the kinds that are used for physical measurement. For instance, the temperature of the skin at your fingertips can be related to the amount of stress that you experience (high stress can cause the constriction of peripheral blood vessels, resulting in a decrease in skin temperature). Using either the Fahrenheit or Celsius temperature scale allows a precise measurement of skin temperature. Because degrees on either scale are fixed units that are always the same size, you can be sure that the interval between, say, 32 and 33 degrees Celsius is exactly the same as the interval between 18 and 19 degrees Celsius. These two temperature scales are therefore called *interval scales*. Another desirable property that scales may have is the ratio property, which requires that the scale have a true zero point, that is, a measurement of zero on the scale indicates that there is really nothing left of what is being measured. For example, a measurement of zero pounds indicates the total absence of weight. If the scale has a true zero point in addition to the interval property, a measurement of 6 units, for instance, will actually indicate twice as much of what is being measured as compared to only 3 units. Thus, it is appropriate to call these scales *ratio scales*. A measurement of zero degrees on the Kelvin temperature scale means that there is no temperature at all (i.e., absolute zero), which is not the case for the Fahrenheit or Celsius scales. Therefore, Kelvin is a ratio scale, whereas the other two are just interval scales (zero degrees Celsius, e.g., does not indicate the complete absence of temperature—it just means that water will freeze into ice).

It is only the interval property that is needed for precise measurement, so for our purposes it is common to make no distinction between interval and ratio scales, referring instead to interval/ratio data, or simply *quantitative data*, as opposed to qualitative data. It's worth noting that the quantitative/qualitative distinction may best be thought of as a dimension on which the clearest cases are at the extremes, with fuzzier cases falling toward the middle. For example, when members of one category can be distinguished from members of another based only on *qualitative features of the items* (e.g., shape or color), not on the quantity of such features, then you have categorical (nominal) data; it makes sense that data of this sort are called *qualitative.* By contrast, variables that assign values based on the *quantity of the variable measured* (e.g., weight, length), rather than the type measured, are reasonably called *quantitative.* But what if you have ranked paintings on the basis of creativity, as in the example from the section on ordinal variables above? Is the difference in creativity between the #1 ranked painting and the #2 ranked painting only qualitative (one painting is in a different category of creativity from the other), or quantitative (one exhibits more creativity than the other)? It may not be clear. Arguably both qualitative *and* quantitative differences create the separation between the two paintings. What about survey data based on participants' answers to questions on a 10-point scale? As we stated earlier,

social scientists often behave as though these data possess interval-scale properties and therefore analyze them quantitatively. So, are ordinal data qualitative, quantitative, or both? As our examples illustrate, the quantitative/qualitative distinction may not be as useful for characterizing ordinal data as it is for categorical and interval/ratio data. In any case, in this text, we describe specific methods for dealing with ordinal data, and at some point you can decide for yourself when to treat your ordinal data as though it came from an interval or ratio scale (looking, of course, at what your colleagues are doing with similar data).

Independent and Dependent Variables

Most behavioral research can be classified into one of two categories: **observational** or **experimental**.

In the simplest experiment, a researcher creates two conditions. The participants assigned to one of the conditions get some form of treatment, such as a pill intended to reduce depression. Those assigned to the other condition get something that superficially resembles the treatment, such as a pill that is totally ineffective (i.e., a placebo); they are part of a control group. These two conditions are the two different levels of an **independent variable**, or one that is created by the experimenter. Commonly, an independent variable is one whose levels are qualitative (e.g., a pill that contains medicine versus a placebo that does not). The **dependent variable**, or the variable that is measured by the experimenter and is expected to change from one level of the independent variable to another, is usually quantitative (such as a self-rating of depression). We will begin to describe such experiments in Part II of this text, when we turn to inferential statistics.

True independent variables are under the control of the experimenter, as in the example above. **Quasi-independent variables** are those that are *not* under the control of the experimenter. Quasi-independent variables are also referred to as subject variables, selected variables, or grouping variables. An obvious example is gender. An experimenter cannot (ethically) randomly assign some participants to be in the ''male'' group and others to be in the ''female'' group. The experimenter just *selects* participants who have already identified themselves as male or female to be in her study. Other examples of quasi-independent variables are religion, marital status, and age. In these cases, the experimenter cannot decide whether you will be Lutheran or Jewish or whether you will be single or married; participants' group memberships have already been determined before they arrive at the study.

Why does it matter whether an independent variable is ''true'' or ''quasi''? Inferences concerning causality depend on this distinction. True independent variables allow experimenters to draw a causal link between the independent variable and the dependent variable (assuming the rest of the experiment was conducted properly). Establishing causality is important to progress in science. In the hypothetical antidepressant experiment

described previously, the researcher can establish whether the active ingredient in the pill really *caused* people to become less depressed. With a quasi-independent variable, such causal conclusions cannot be made (e.g., if married people are less depressed, is it because being married is causing a reduction in depression, or because those who are less depressed are more likely to get married?).

Some independent variables can be truly experimental or quasi-experimental depending on the design of the study. For example, suppose you wanted to know whether exercising leads to better health (as measured by number of sick days taken at work). You could ask a sample of workers about their exercise habits and number of sick days over a 6-month period. If the data reveal that those who exercise the most have the fewest sick days, the researcher is *not* entitled to draw a causal relationship between the two. In this example, exercise habit is a subject variable; the experimenter did not manipulate how many hours per week each person spends in the gym. That lack of control leaves open other possible explanations. Maybe the causal arrow points the other way and people who are healthy tend to (and are able to) work out more. Perhaps some third variable like optimism drives both health and exercise. You just cannot be sure which conclusion is correct.

Now imagine that the researcher had assigned people at random to work out 0, 5, 10, or 15 hours per week at a gym and that after 6 months she got the same inverse association between hours spent exercising and number of sick days. In this situation, the researcher *can* claim that exercise causes fewer sick days. Other possibilities such as those mentioned earlier have been controlled for. Optimism cannot explain the results because random assignment would have distributed the optimistic people evenly among the four workout groups. Note also that true experiments are usually more difficult to conduct compared to quasi-experimental versions of the same studies. Therefore, a trade-off often exists between effort and expense, on one hand, and the sort of conclusions that can be drawn, on the other. In fact, behavioral researchers often study the relationships among variables that are not convenient, or even possible, to ever manipulate. If one simply measures the relationship between two dependent variables (e.g., self-esteem in teenagers and their family's annual income to see if those from more affluent families tend to have higher—or lower—self-esteem), this is an observational (i.e., quasi-experimental) study.

Methods for dealing with studies in which both the independent and dependent variables (or two dependent variables) are quantitative will be presented in Chapter 9. Research in which both variables being related are qualitative will be discussed in the last chapter of this text. Simple studies, in which one variable has only one or two qualitative levels, while the other is continuous and therefore has many (theoretically, infinite) possible levels, will be analyzed in Chapters 6 and 7. For now, it is important to begin developing the tools you will need to work with quantitative data. One of the most basic of such tools is the summation procedure, to which we turn next.

Summation Notation

Mathematical formulas and symbols can appear forbidding, if you're not used to working with them. Once you get to know them, however, you see that they actually simplify matters. They are just convenient ways to clearly and concisely convey information that would be much more awkward to express in words, much in the same way that ''LOL'' and ''OMG'' simplify expression in email and text messages. In statistics, a particularly important symbol is the one used to represent the *sum* of a set of numbers—that is, the value obtained by adding up all of the numbers.

To illustrate the use of summation notation, let us suppose that eight students take a 10-point quiz. Letting X stand for the variables in question (quiz scores), let us further suppose that the results are as follows:

$$X_1 = 7 \quad X_2 = 9 \quad X_3 = 6 \quad X_4 = 10$$
$$X_5 = 6 \quad X_6 = 5 \quad X_7 = 3 \quad X_8 = X_N = 4$$

Notice that X_1 represents the score of the first student on X; X_2 stands for the score of the second student; and so on. Also, the *number of scores* is denoted by N; in this example, $N = 8$. The last score may be represented by either X_8 or X_N. The *sum of all the X scores* is represented by

$$\sum_{i=1}^{N} X_i$$

where $\sum$, the Greek capital letter sigma, stands for ''the sum of'' and is called the *summation sign*. The subscript below the summation sign indicates that the sum begins with the first score (X_i where $i = 1$), and the superscript above the summation sign indicates that the sum continues up to and including the last score (X_i where $i = N$ or, in this example, 8). Thus,

$$\begin{aligned}\sum_{i=1}^{N} X_i &= X_1 + X_2 + X_3 + X_4 + X_5 + X_6 + X_7 + X_8 \\ &= 7 + 9 + 6 + 10 + 6 + 5 + 3 + 4 \\ &= 50\end{aligned}$$

Most of the time, however, the sum of *all* the scores is needed in the statistical analysis. In such situations it is customary to omit the indices i and N from the notation, as follows:

$$\sum X = \text{sum of all the } X \text{ scores}$$

The absence of written indication as to where to begin and end the summation is taken to mean that all the X scores are to be summed.

Summation Rules

Certain rules involving summation notation will prove useful in subsequent chapters. But first, we'd like to flash back to junior high when you

learned the order in which mathematical operations should be computed. Do you recall learning the order of operations, PEMDAS, together with a phrase to help you remember it like ''Please Excuse My Dear Aunt Sally''? Or perhaps ''Pandas Eat: Mustard on Dumplings, and Apples with Spice''? ''PEMDAS'' stands for ''Parentheses, Exponents, Multiplication and Division, and Addition and Subtraction.'' It describes the order in which you should perform mathematical operations: Parentheses are taken care of before exponents, followed by multiplication and division (which are at the same rank), followed by addition and subtraction (also at the same rank). As you will see, these summation rules often involve operations beyond addition, so Please Excuse My Dear Aunt Sally. Let us suppose that the eight students previously mentioned take a second quiz, denoted by Y. The results of both quizzes can be summarized conveniently as follows:

Students (S)	Quiz 1 (X)	Quiz 2 (Y)
1	7	8
2	9	6
3	6	4
4	10	10
5	6	5
6	5	10
7	3	9
8	4	8

We have already seen that $\sum X = 50$. The sum of the scores on the second quiz is equal to

$$\begin{aligned}\sum Y &= Y_1 + Y_2 + Y_3 + Y_4 + Y_5 + Y_6 + Y_7 + Y_8 \\ &= 8 + 6 + 4 + 10 + 5 + 10 + 9 + 8 \\ &= 60\end{aligned}$$

The following rules are illustrated using the small set of data just shown, and you should verify each one carefully.

Rule 1. $\boldsymbol{\sum(X + Y) = \sum X + \sum Y}$

Illustration	S	X	Y	$X + Y$
	1	7	8	15
	2	9	6	15
	3	6	4	10
	4	10	10	20
	5	6	5	11
	6	5	10	15
	7	3	9	12
	8	4	8	12
		$\sum X = 50$	$\sum Y = 60$	$\sum(X + Y) = 110$
		$\sum X + \sum Y = 110$		

If you remember your Aunt Sally, this rule should be intuitively obvious; the same total should be reached regardless of the order in which the scores are added.

Rule 2. $\sum(X - Y) = \sum X - \sum Y$

Illustration	S	X	Y	$X - Y$
	1	7	8	−1
	2	9	6	3
	3	6	4	2
	4	10	10	0
	5	6	5	1
	6	5	10	−5
	7	3	9	−6
	8	4	8	−4
		$\sum X = 50$	$\sum Y = 60$	$\sum(X - Y) = -10$
		$\sum X - \sum Y = -10$		

As with the first rule, it makes no difference whether you subtract first and then sum $\left[\sum(X - Y)\right]$ or obtain the sums of X and Y first and then subtract $\left(\sum X - \sum Y\right)$.

Unfortunately, matters are not so simple when multiplication and squaring are involved.

Rule 3. $\sum XY \neq \sum X \sum Y$

That is, first multiplying each X score by the corresponding Y score and then summing $\left(\sum XY\right)$ is *not* equal to summing the X scores $\left(\sum X\right)$ and summing the Y scores $\left(\sum Y\right)$ first and then multiplying once $\left(\sum X \sum Y\right)$.

Illustration	S	X	Y	XY
	1	7	8	56
	2	9	6	54
	3	6	4	24
	4	10	10	100
	5	6	5	30
	6	5	10	50
	7	3	9	27
	8	4	8	32
		$\sum X = 50$	$\sum Y = 60$	$\sum XY = 373$
		$\sum X \sum Y = (50)(60) = 3{,}000$		

Observe that $\sum XY = 373$, while $\sum X \sum Y = 3{,}000$.

Rule 4. $\sum X^2 \neq (\sum X)^2$

That is, first squaring all of the X values and then summing $\left(\sum X^2\right)$ is *not* equal to summing first and then squaring that single quantity, $\left[\left(\sum X\right)^2\right]$.

Illustration	S	X	X^2
	1	7	49
	2	9	81
	3	6	36
	4	10	100
	5	6	36
	6	5	25
	7	3	9
	8	4	16
		$\sum X = 50$	$\sum X^2 = 352$
		$\left(\sum X\right)^2 = (50)^2 = 2{,}500$	

Here, $\sum X^2 = 352$, while $\left(\sum X\right)^2 = 2{,}500$.

Rule 5. If k is a *constant* (a fixed numerical value), then

$$\sum k = Nk$$

	Suppose that $k = 3$. Then,	
Illustration	S	k
	1	3
	2	3
	3	3
	4	3
	5	3
	6	3
	7	3
	8	3
		$\sum k = 24$
	$Nk = (8)(3) = 24$	

Rule 6. If k is a constant,

$$\sum (X + k) = \sum X + Nk$$

		Suppose that $k = 5$. Then,		
Illustration	S	X	k	$X + k$
	1	7	5	12
	2	9	5	14
	3	6	5	11
	4	10	5	15
	5	6	5	11
	6	5	5	10
	7	3	5	8
	8	4	5	9
		$\sum X = 50$	$\sum k = Nk = 40$	$\sum (X + k) = 90$
		$\sum X + Nk = 50 + (8)(5) = 90$		

This rule follows directly from Rules 1 and 5.

Rule 7. If k is a constant,

$$\sum(X - k) = \sum X - Nk$$

The illustration of this rule is similar to that of Rule 6 and is left to the reader as an exercise.

Rule 8. If k is a constant,

$$\sum kX = k\sum X$$

	Suppose that $k = 2$. Then,			
Illustration	***S***	***X***	***k***	***kX***
	1	7	2	14
	2	9	2	18
	3	6	2	12
	4	10	2	20
	5	6	2	12
	6	5	2	10
	7	3	2	6
	8	4	2	8
		$\sum X = 50$		$\sum kX = 100$
		$k\sum X = (2)(50) = 100$		

Ihno's Study

The process of summation can be considered the ''workhorse'' of statistics; you will see it used in one form or another in all of the statistical procedures described later in this text. Specifically, the summation sign is an important part of most statistical formulas, and we think it is vital that you understand how statistical formulas work. Although we expect, and certainly hope, that all statistics students will eventually become proficient at performing statistical analyses by computer, we believe that there is important educational value in asking students to apply basic statistical formulas directly to small sets of numbers to see how the formulas work and to thus gain a greater conceptual understanding of the statistical results being generated by computer programs. At the same time, we want to help you become familiar with the use of statistical software, especially the statistical package known as SPSS, so to that end, we created a fairly large data set (100 participants and more than a dozen variables), which is included in a table in the appendix of this book and is also available in several electronic formats online. In addition to exercises at the end of each chapter, which ask the reader to apply the formulas of that chapter to small sets of data, we have also included a set of computer exercises in each chapter, and each of those exercises refers to the same large data set in the Appendix, which is described next.

The data in the Appendix comes from a hypothetical study performed by Ihno (pronounced ''Eee-know''), an advanced doctoral student, who is the teaching assistant for several sections of a statistics course. The 100 participants are the students who were enrolled in Ihno's sections and who attended one of two review classes that she conducted each week as the teaching assistant (TA) for the course. (Of course, all of her students voluntarily signed proper informed consent forms, and her study was approved by the appropriate review board at her hypothetical school.) Her data were collected on two different days. On the first day of classes, the students who came to either Ihno's morning or afternoon review session filled in a brief background questionnaire on which they provided contact information, some qualitative data (gender, undergrad major, why they had enrolled in statistics, and whether they regularly drink coffee) and some quantitative data (number of math courses already completed, the score they received on a diagnostic math background quiz they were all required to take before registering for statistics, and a rating of their math phobia on a scale from 0 to 10).

The rest of Ihno's data were collected as part of an experiment that she conducted on one day in the middle of the semester. The combined results of the two class sessions on that day add up to a total of 100 students who participated in the experiment. (Due to late registration and other factors, not all of Ihno's students took the diagnostic math background quiz.) At the beginning of the experiment, Ihno explained how each student could take his or her own pulse. She then provided a one-minute interval during which they counted the number of beats and wrote down that number as their (baseline) heart rate in beats per minute (bpm). Then each student reported how many cups of coffee they had drunk since waking up that morning and filled out an anxiety questionnaire consisting of 10 items, each rated (0 to 4) on a 5-point Likert scale. The questionnaire items inquired about anxiety and how the student was feeling at the present time (e.g., ''Would you say that you are now feeling tense and restless? Circle one: Not at all; Somewhat; Moderately; Quite a bit; Extremely''). Total scores could range from 0 to 40, and provided a measure of baseline anxiety.

Next, Ihno announced a pop quiz. She handed out a page containing 11 multiple-choice statistics questions on material covered during the preceding two weeks, and asked the students to keep this page face-down while taking and recording their (prequiz) pulse and filling out an anxiety questionnaire for a second time (i.e., prequiz). Then Ihno told the students they had 15 minutes to take the fairly difficult quiz. She also told them that the first 10 questions were worth 1 point each but that the 11th question was worth 3 points of extra credit. Ihno's experimental manipulation consisted of varying the difficulty of the 11th question. Twenty-five quizzes were distributed at each level of difficulty of the final question: easy, moderate, difficult, and impossible to solve. After the quizzes were collected at the end of the 15 minutes, Ihno asked the students to provide heart rate and anxiety data (postquiz) one more

time. Finally, Ihno explained the experiment, adding that the 11th quiz question would not be scored and that, although the students would get back their quizzes with their score for the first 10 items, that score would not influence their grade for the statistics course. All of the data from Ihno's experiment is printed out in the Appendix at the end of this book.

The computer exercises in this text can be solved by any major statistical package, and because the data for Ihno's study is also available on the Web as both a tab-delimited text file and an Excel 2007 file (go to www.psych.nyu.edu/cohen/statstext.html), most packages will be able to read in the data directly from the electronic file. However, because SPSS is currently the most popular statistical package among behavioral researchers, we have also included Ihno's data as an SPSS file. Moreover, Bridge to SPSS sections at the end of each chapter can give you further assistance with computer analysis, unless, of course, your statistics instructor has assigned different statistical software. However, note that our Bridge to SPSS sections are not meant to replace a basic guide to the SPSS program (one advantage of using such a popular package is that there are many good guides available at a variety of different levels and prices); they are designed to help you bridge the gaps between the labels SPSS uses to present results in its output files and the statistical terms and notation used in this text. The Bridge to SPSS sections also help to ensure that you will know how to use SPSS to solve the computer exercises in the corresponding chapters.

Summary

Descriptive statistics are used to summarize and make large quantities of data understandable. *Inferential statistics* are used to draw inferences about numerical quantities (called *parameters*) concerning *populations* based on numerical quantities (called *statistics*) obtained from *samples*. Some behavioral variables cannot be measured quantitatively (e.g., choice of religion) but can only be measured qualitatively using categories (e.g., Catholic, Jewish, Buddhist), which comprise a *nominal* scale. If the categories can be placed in order (e.g., the belts awarded for different levels of skill in the martial arts—black belt, brown belt, etc.), an *ordinal* scale has been created. If the scale involves precise measurement resulting in units of equal size, the data are considered to be *quantitative*, whether the scale has a true zero point (*ratio* scale) or not (*interval* scale). (Ordinal scales are often treated as quantitative in behavioral research; nominal scales always produce *qualitative* data.) Experiments involve measuring dependent variables that are expected to vary somewhat as a function of the different levels of one or more independent variables created by the researcher. *Observational* research involves comparing dependent variables to each other because no variables are being manipulated, nor are participants being assigned to experimental conditions.

The summation sign, $\sum$, is used to indicate ''the sum of'' and occurs frequently in statistical work. Remember that $\sum X$ is a shorthand version of

$$\sum_{i=1}^{N} X_i$$

(where N = number of subjects or cases).

Summation Rules:

1. $\sum(X + Y) = \sum X + \sum Y$
2. $\sum(X - Y) = \sum X - \sum Y$
3. $\sum XY$ (multiply first, then add) $\neq \sum X \sum Y$ (add first, then multiply)
4. $\sum X^2$ (square first, then add) $\neq (\sum X)^2$ (add first, then square)
5. If k is a constant, $\sum k = Nk$
6. $\sum(X + k) = \sum X + Nk$
7. $\sum(X - k) = \sum X - Nk$
8. $\sum kX = k \sum X$

Exercises

A first-year student from Oklahoma named Carrie is trying to decide which of four dormitories to list as her first choice. She wants to live in the dorm with the friendliest students, so she makes up a friendliness questionnaire, on which a score of 20 indicates the friendliest a student can be and zero, the least friendly. She manages to get 15 students each from two of the dorms to fill out her questionnaire, 10 from a third dorm, and 5 from a fourth. The individual friendliness scores are shown in Table 1.1. Exercise 1 in this chapter refers to the set of data just described. We will return to this data set several times in later chapters.

Table 1.1 **Hypothetical Scores on a 20-Point Friendliness Measure for Students From Four Different Dormitories at One Midwestern College**

Turck Hall ($N = 15$)	17	18	6	13	9
	17	11	16	5	15
	11	13	10	20	14
Kirk Hall ($N = 15$)	17	8	12	12	3
	12	7	14	1	11
	9	8	6	11	7
Dupre Hall ($N = 10$)	9	11	6	5	4
	9	10	4	5	7
Doty Hall ($N = 5$)	14	8	17	6	10

1. Compute the following (these values will not by themselves help Carrie to make a decision about the dormitories, but we will use these values as steps in future exercises to answer Carrie's concern):
   (a) For Turck Hall:

$$\sum X = \_\_\_\_ \quad \sum X^2 = \_\_\_\_ \quad \left(\sum X\right)^2 = \_\_\_\_$$

   (b) For Kirk Hall:

$$\sum X = \_\_\_\_ \quad \sum X^2 = \_\_\_\_ \quad \left(\sum X\right)^2 = \_\_\_\_$$

   (c) For Dupre Hall:

$$\sum X = \_\_\_\_ \quad \sum X^2 = \_\_\_\_ \quad \left(\sum X\right)^2 = \_\_\_\_$$

   (d) For Doty Hall:

$$\sum X = \_\_\_\_ \quad \sum X^2 = \_\_\_\_ \quad \left(\sum X\right)^2 = \_\_\_\_$$

2. Express the following words in symbols.
   (a) Add up all the scores on test X, then add up all the scores on test Y, and then add the two sums together.
   (b) Add up all the scores on test G. To this, add the following: the sum obtained by squaring all the scores on test P and then adding them up.
   (c) Square all the scores on test X. Add them up. From this, subtract 6 times the sum you get when you multiply each score on X by the corresponding score on Y and add them up. To this, add 4 times the quantity obtained by adding up all the scores on test X and squaring the result. To this, add twice the sum obtained by squaring each Y score and then adding them up. (Compare the amount of space needed to express this equation in words with the amount of space needed to express it in symbols. Do you see why summation notation is necessary?)
3. Five students are enrolled in an advanced course in psychology. Two quizzes are given early in the semester, each worth a total of 10 points. The results are as follows:

Student	Quiz 1 (X)	Quiz 2 (Y)
1	0	2
2	2	6
3	1	7
4	3	6
5	4	9

(a) Compute each of the following:

$\sum X =$ _____ $(\sum X)^2 =$ _____ $\sum (X - Y) =$ _____
$\sum Y =$ _____ $(\sum Y)^2 =$ _____ $\sum X - \sum Y =$ _____
$\sum X^2 =$ _____ $\sum (X + Y) =$ _____ $\sum XY =$ _____
$\sum Y^2 =$ _____ $\sum X + \sum Y =$ _____ $\sum X \sum Y =$ _____

(b) Using the results of part (a), show that each of the following rules listed in this chapter is true:

Rule 1: _____ = _____
Rule 2: _____ = _____
Rule 3: _____ ≠ _____
Rule 4: _____ ≠ _____ (X data)
_____ ≠ _____ (Y data)

(c) After some consideration, the instructor decides that Quiz 1 was excessively difficult and decides to add 4 points to each student's score. This can be represented in symbols by using k to stand for the constant amount in question, 4 points.

Using Rule 6, compute $\sum (X + k) =$ _____ + _____ = _____.
Compute $\sum X + k =$ _____ + _____ = _____. (Note that this result is different from the preceding one.)
Now add 4 points to each student's score on Quiz 1 and obtain the sum of these new scores.

(d) Had the instructor been particularly uncharitable, he might have decided that Quiz 2 was too easy and subtracted 3 points from each student's score on that quiz. Although this is a new problem, the letter k can again be used to represent the constant; here, $k = 3$.

Using Rule 7, compute $\sum (Y - k) =$ _____ − _____ = _____.
Compute $\sum Y - k =$ _____ − _____ = _____. (Note that this result is different from the preceding one.)
Now subtract 3 points from each student's score on Quiz 2 and obtain the sum of these new scores.

(e) Suppose that the instructor decides to double all of the original scores on Quiz 1.

Using Rule 8, compute $\sum kX =$ _____ × _____ = _____.
Now double each student's score on Quiz 1 and obtain the sum of these new scores.

4. For each of the following (separate) sets of data, compute the values needed in order to fill in the answer spaces. Then answer the additional questions that follow.

Data set 1:

S	*X*	*Y*	$N =$ ____	
1	1	2	$\sum X =$ ____	$\sum Y =$ ____
2	3	5	$\sum X^2 =$ ____	$\sum Y^2 =$ ____
3	1	0	$(\sum X)^2 =$ ____	$(\sum Y)^2 =$ ____
4	0	1	$\sum XY =$ ____	$\sum X \sum Y =$ ____
5	2	3	$\sum(X + Y) =$ ____	$\sum(X - Y) =$ ____

Data set 2:

S	*X*	*Y*	$N =$ ____	
1	7.14	0	$\sum X =$ ____	$\sum Y =$ ____
2	8.00	2.60	$\sum X^2 =$ ____	$\sum Y^2 =$ ____
3	0	4.32	$(\sum X)^2 =$ ____	$(\sum Y)^2 =$ ____
4	4.00	2.00	$\sum XY =$ ____	$\sum X \sum Y =$ ____
5	4.00	6.00	$\sum(X + Y) =$ ____	$\sum(X - Y) =$ ____
6	1.00	1.15		
7	2.25	1.00		
8	10.00	3.00		

	set 1	*set 2*
If every X score is multiplied by 3.2, what is the new $\sum X$ in each set?	____	____
If 7 is subtracted from every Y score, what is the new $\sum Y$ in each set?	____	____
If 1.8 is added to every X score, what is the new $\sum X$ in each set?	____	____
If every Y score is divided by 4, what is the new $\sum Y$ in each set?	____	____

(Hint: Use the appropriate summation rule in each case so as to make the calculations easier.)

5. Compute the values needed to fill in the blanks.

Data set 3:

S	*X*	*Y*	$N =$ ____	
1	97	89	$\sum X =$ ____	$\sum Y =$ ____
2	68	57	$\sum X^2 =$ ____	$\sum Y^2 =$ ____
3	85	87	$(\sum X)^2 =$ ____	$(\sum Y)^2 =$ ____
4	74	76	$\sum XY =$ ____	$\sum X \sum Y =$ ____
5	92	97	$\sum(X + Y) =$ ____	$\sum(X - Y) =$ ____
6	92	79		
7	100	91		
8	63	50		
9	85	85		
10	87	84		
11	81	91		
12	93	91		
13	77	75		
14	82	77		

Thought Questions

1. What is the difference between (a) a population and a sample? (b) a parameter and a statistic? (c) descriptive statistics and inferential statistics?
2. What is the difference between a constant and a variable?
3. What is the difference between a ratio scale and an interval scale?
4. What important property do interval scales have that ordinal and nominal scales do *not* have?
5. If we are classifying psychiatric patients as having one or another of five different psychological disorders—Major Depressive Disorder, Bipolar Disorder, Generalized Anxiety Disorder, Obsessive-Compulsive Disorder, and Phobic Disorder—what kind of measurement scale are we using?
6. A poll of sportswriters ranks the 25 best college football teams in the country, where #1 is the best team, #2 is the second best team, and so on. What kind of measurement scale is this?
7. What kind of measurement scales are used for each of the following variables in Ihno's study? (a) The gender of each student. (b) The undergraduate major of each student. (c) The number of math courses each student has already completed. (d) Score on math background quiz.
8. What is the difference between an independent variable and a dependent variable?

Computer Exercises

1. Read Ihno's data into your statistical software package. For those not using the Statistical Package for the Social Sciences (SPSS), we have provided the data in two convenient formats: a tab-delimited text file and an Excel spreadsheet (Microsoft Office 2007). For the convenience of SPSS users, we have also included the data as an SPSS.sav file, though your instructor may want you to know how to read text or Excel files into SPSS.
2. Label the values of the categorical (i.e., qualitative) variables according to the following codes: For gender, 1 = Female and 2 = Male; for undergrad major, 1 = Psychology, 2 = Pre-med, 3 = Biology, 4 = Sociology, and 5 = Economics. Your instructor may ask you to fill in missing-value codes for any data that are missing (e.g., blank cell in the Excel spreadsheet).
3. A good many statistical functions can be performed in Excel. As a first step, use the Sum function to add up the scores for each of the quantitative variables in the Excel file of Ihno's data.
4. Create a new variable that adds 10 points to everyone's math background quiz score. How does the sum of this variable compare to the

sum of the original variable? What general rule is being illustrated by this comparison?

5. Create a new variable that is 10 times the statistics quiz score. How does the sum of this variable compare to the sum of the original variable? What general rule is being illustrated by this comparison?

Bridge to SPSS

SPSS is certainly not the only statpack available, and your instructor may prefer to teach you how to use SAS, Minitab, *R*, or another statistical program. However, because SPSS is the most popular statistical software package in the social and behavioral sciences, and is particularly easy to learn for introductory statistics students, we provide this section after the computer exercises in each chapter. (This should help students translate the terminology used by SPSS into the language that we are using to describe the same topics in this text.) We also show you briefly how to obtain from SPSS the statistics discussed in the chapter, along with a few tricks and shortcuts. However, we understand that even though you may be using SPSS in your statistics course, you may not be using the latest version, which will probably be 20.0 by the time this text is printed. Therefore, we only describe aspects of SPSS that have not changed between version 14 and the latest version.

The rightmost column in SPSS's Variable View spreadsheet is labeled **Measure**, and it allows you to classify each of your variables as being measured by one of the following three types of scales: nominal, ordinal, or scale. The first two terms are used the same way by SPSS that we have defined them in this chapter. *Scale* is SPSS's term for interval/ratio data. In general, numerical data are set to **scale** by default, whereas string data—which contain letters instead of, or in addition to, numbers—are set to **nominal**. In practice, these scale designations are not important, because SPSS uses them only to determine the way some charts are displayed.

For simplicity, Ihno's data set is presented entirely in terms of numbers, even for the categorical variables of gender and undergrad major. To assign meaningful labels to the arbitrary numbers we have used to represent the different levels of the categorical variables, go down the Values column of Variable View until you reach the row for a categorical variable. Then click in the right side of that cell to open the Value Labels box. For gender, you would type **1** for Value, and then tab to Value Label, where you can type **female**. Click the **Add** button, and provide a label (male) for value **2**, then **Add** again, and **OK**. The process is similar for undergrad major. Note that if you use the Missing column to define a particular value of a variable, say 99, as meaning that the value is missing, rather than just leaving the cell blank (e.g., you could use 99 to mean ''missing'' for the math background quiz, because none of the real values can be that high), you can then attach a value label to that value, such as ''never took the math quiz.''

To create new variables that are based on ones already in your spreadsheet, which some of our computer exercises will ask you to do, click on the **Transform** menu, then **Compute**. In the Compute Variable box that opens up, Target Variable is a name that you make up (and type into that box) for the new variable (but no more than eight characters and no embedded spaces); when you have filled in a Numeric Expression and then click **OK**, the new variable will suddenly appear in the rightmost column of your Data View spreadsheet. We will leave it to your instructor, or an SPSS guidebook, to teach you various ways to create numeric expressions that transform your existing variables into new ones.