EXPLORING STATISTICS

TALES OF DISTRIBUTIONS

12TH EDITION

CHRIS SPATZ

12th Edition

Exploring Statistics

Tales of Distributions

Chris Spatz

Outcrop Publishers Conway, Arkansas

Online study guide available at
http://exploringstatistics.com/studyguide.php

Exploring Statistics: Tales of Distributions
12th Edition
Chris Spatz

Cover design: Grace Oxley
Answer Key: Jill Schmidlkofer
Webmaster & Ebook: Fingertek Web Design, Tina Haggard
Managers: Justin Murdock, Kevin Spatz

Outcrop Publishers
615 Davis Street
Conway, AR 72034
Email: info@outcroppublishers.com
Website: outcroppublishers.com
Library of Congress Control Number: [Applied for]

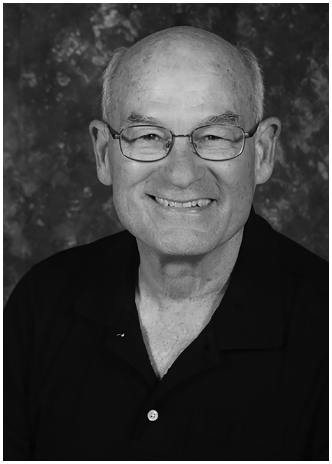
ISBN-13 (hardcover): 978-0-9963392-2-3
ISBN-13 (ebook): 978-0-9963392-3-0
ISBN-13 (study guide): 978-0-9963392-4-7

Ordering information:
Students and professors – visit exploringstatistics.com
Bookstores – email info@outcroppublishers.com

Photo Credits – Chapter 1
Karl Pearson – Courtesy of Wellcomeimages.org
Ronald A. Fisher – R.A. Fisher portrait, 0006973, Special Collections Research Center, North Carolina State University Libraries, Raleigh, North Carolina
Jerzy Neyman – Paul R. Halmos Photograph Collection, e_ph 0223_01, Dolph Briscoe Center for American History, The University of Texas at Austin
Jacob Cohen – New York University Archives, Records of the NYU Photo Bureau

Printed in the United States of America by Walsworth ®
1 2 3 4 5 6 7 24 23 22 21 20 19 18

About The Author



Chris Spatz is at Hendrix College where he twice served as chair of the Psychology Department. Dr. Spatz's undergraduate education was at Hendrix and his PhD in experimental psychology is from Tulane University in New Orleans. He subsequently completed postdoctoral fellowships in animal behavior at the University of California, Berkeley, and the University of Michigan. Before returning to Hendrix to teach, Spatz held positions at The University of the South and the University of Arkansas at Monticello.

Spatz served as a reviewer for the journal Teaching of Psychology for more than 20 years. He co-authored a research methods textbook, wrote several chapters for edited books, and was a section editor for the *Encyclopedia of Statistics in Behavioral Science.*

In addition to writing and publishing, Dr. Spatz enjoys the outdoors, especially canoeing, camping, and gardening. He swims several times a week (mode = 3). Spatz has been an opponent of high textbook prices for years, and he is happy to be part of a new wave of authors who provide high-quality textbooks to students at affordable prices.

With love and affection,

this textbook is dedicated to

Thea Siria Spatz, Ed.D., CHES

CHAPTER

1

Introduction

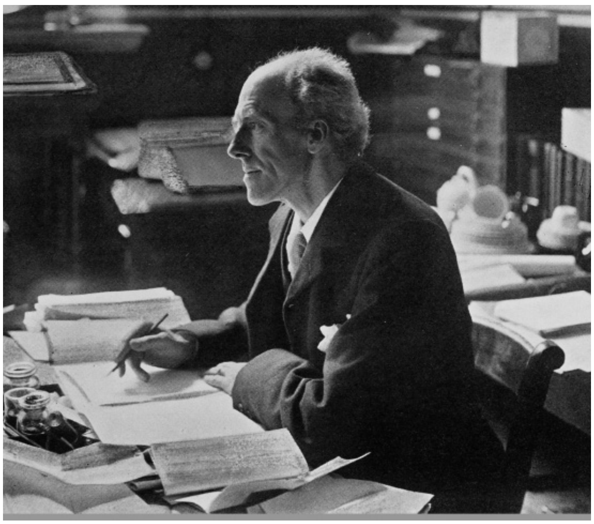
OBJECTIVES FOR CHAPTER 1

After studying the text and working the problems in this chapter, you should be able to:

1. Distinguish between descriptive and inferential statistics
2. Define population, sample, parameter, statistic, and variable as they are used in statistics
3. Distinguish between quantitative and categorical variables
4. Distinguish between continuous and discrete variables
5. Identify the lower and upper limits of a continuous variable
6. Identify four scales of measurement and distinguish among them
7. Distinguish between statistics and experimental design
8. Define independent variable, dependent variable, and extraneous variable and identify them in experiments
9. Describe statistics' place in epistemology
10. List actions to take to analyze a data set
11. Identify a few events in the history of statistics

WE BEGIN OUR exploration of statistics with a trip to London. The year is 1900.

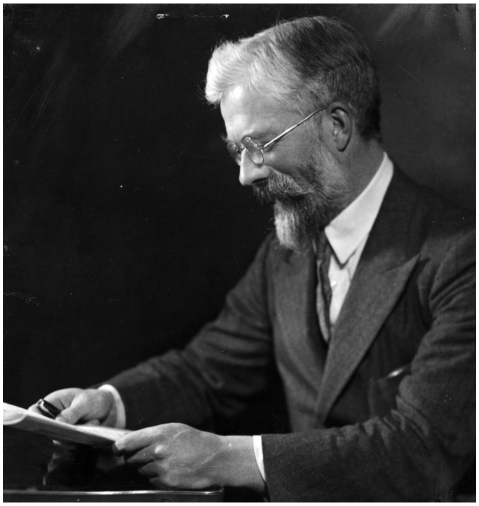
Walking into an office at University College London, we meet a tall, well-dressed man about 40 years old. He is Karl Pearson, Professor of Applied Mathematics and Mechanics. I ask him to tell us a little about himself and why he is an important person. He seems authoritative, glad to talk about himself. As a young man, he says, he wrote essays, a play, and a novel, and he also worked for women's suffrage. These days, he is excited about this new branch of biology called genetics. He says he supervises lots of data gathering.



Karl Pearson

Pearson, warming to our group, lectures us about the major problem in science—there is no agreement on how to decide among competing theories. Fortunately, he just published a new statistical method that provides an objective way to decide among competing theories, regardless of the discipline. The method is called chi square.[1] Pearson says, "Now, arguments will be much fewer. Gather a thousand data points and calculate a chi square test. The result gives everyone an objective way to determine whether or not the data fit the theory."

Exploration Notes from a student: Exploration off to good start. Hit on a nice, easy-to-remember date to start with, visited a founder of statistics, and had a statistic called chi square described as a big deal.



Ronald A. Fisher

Our next stop is Rothamsted Experiment Station just north of London. Now the year is 1925. There are fields all around the agricultural research facility, each divided into many smaller plots. The growth in the fields seems quite variable.

Arriving at the office, the atmosphere is congenial. The staff is having tea. There are two topics—a new baby and a new book. We get introduced to Ronald Fisher, the chief statistician. Fisher is a small man with thick glasses and red hair.

He tells us about his new child[2] and then motions to a book on the table. Sneaking a peek, we read the title: *Statistical Methods for Research Workers*. Fisher becomes focused on his book, holding forth in an authoritative way.

He says the book explains how to conduct experiments and that an experiment is just a comparison of two or more conditions. He tells us we don't need a thousand data points. He says that small samples, randomly selected, are the way for science to progress. "With an experiment and my technique of analysis of variance," he exclaims, "you can determine why that field out there"—here he waves toward the window—"is so variable. We can find out what makes some plots lush and some mimsy." *Analysis of variance*,[3] he says, works in any discipline, not just agriculture.

Exploration Notes: Looks like statistics had some controversy in it.[4] Also looks like progress. Statistics is used for experiments, too, and not just for testing theories. And Fisher says experiments can be used to compare anything. If that's right, I can use statistics no matter what I major in.

---

[1] Chi square, which is explained in this book in Chapter 14, has been called one of the 20 most important inventions in the 20th century (Hacking, 1984).
[2] (in what will become a family with eight children).
[3] explained in Chapters 11-13
[4] The slight sniping I've built into this story is just a hint of the strong animosity between Fisher and Pearson.

Next we go to Poland to visit Jerzy Neyman at his office at the University of Warsaw. It is 1933. As we walk in, he smiles, seems happy we've arrived, and makes us feel completely welcome.


Jerzy Neyman

Motioning to an envelope on his desk, he tells us it holds a manuscript that he and Egon Pearson[5] wrote. "The problem with Fisher's analysis of variance test is that it focuses exclusively on finding a difference between groups. Suppose the statistical test doesn't detect a difference. Does that prove there is no difference? No, of course not. It may be that the test was just not sensitive enough to detect the difference. Right?"

At his question, a few of us nod in agreement. Seeing uncertainty, he notes, "Maybe a larger sample is needed to find the difference, you see? Anyway, what we've done is expand statistics to cover not just finding a difference, but also what it means when the test doesn't find a difference. Our approach is what you people in your time will call *null hypothesis significance testing.*"

Exploration Notes: Statistics seems like a work in progress. Changing. Now it is not just about finding a difference but also about what it means not to find a difference. Also, looks like null hypothesis significance testing is a phrase that might turn up on tests.

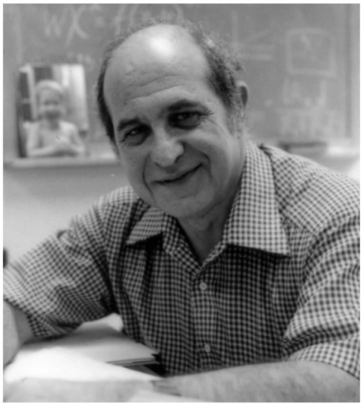
Our next trip is to libraries, say, anytime between 1940 and 2000. For this exploration, the task is to examine articles in professional journals published in various disciplines. The disciplines include anthropology, biology, chemistry, defense strategy, education, forestry, geology, health, immunology, jurisprudence, manufacturing, medicine, neurology, ophthalmology, political science, psychology, sociology, zoology, and others. I'm sure you get the idea—the whole range of disciplines that use quantitative measures in their research. What this exploration produces is the discovery that all of these disciplines rely on a data analysis technique called *null hypothesis significance testing* (NHST).[6] Many different statistical tests are employed. However, for all the tests in all the disciplines, the phrase, "$p < .05$" turns up frequently.

Exploration Notes: It seems that all that earlier controversy has subsided and scientists in all sorts of disciplines have agreed that NHST is the way to analyze quantitative data. All of them seem to think that if there is a comparison to be made, applying NHST is a necessary step to get correct conclusions. All of them use "$p < .05$," so I'll have to be sure to find out exactly what that means.

---

[5] Egon Pearson was Karl Pearson's son.

[6] Null hypothesis significance testing is first explained in Chapters 9 and 10.



Jacob Cohen

Our next excursion is a 1962 visit with Jacob Cohen at New York University in New York City. He is holding his article about studies published in the *Journal of Abnormal and Social Psychology*, a leading psychology journal. He tells us that the NHST technique has problems. Also, he says we should be calculating an effect size statistic, which will show whether the differences observed in our experiments are large or small.

Exploration Notes: The idea of an effect size index makes a lot of sense. Just knowing there is a difference isn't enough. How big is the difference? Wonder what "problems with NHST" is all about.

Back to the library for a final excursion to check out recent events. We come across a 2014 article by Geoff Cumming on the "new statistics." We find things like, "avoid NHST and use better techniques" (p. 26) and "we should not trust any p value" (p. 13). This seems like awfully strong advice. Are researchers taking this advice? Looking through more of today's research in journals in several fields, we find that most statistical analyses use NHST and there are many instances of "$p < .05$."

Exploration Notes, Conclusion: These days, it looks like statistics is in transition again. There's a lot of controversy out there about how to analyze data from experiments. The NHST approach is still very common, though, so it's clear I must learn it. But I want to be prepared for changes. I hope knowing NHST will be helpful for the future.[7]

Welcome to statistics at a time when the discipline is once again in transition. A well-established tradition (null hypothesis significance testing) has been in place for almost a century but is now under attack. New ways of thinking about data analysis are emerging, and along with them, a collection of statistics that do not include the traditional NHST approach. As for the immediate future, though, NHST remains the method most widely used by researchers in many fields. In addition, much of the thinking required for NHST is required for other approaches.

Our exploration tour is over, so I'll quit supplying notes; they are your responsibility now. As your own experience probably shows, making up your own summary notes improves retention of what you read. In addition, I have a suggestion. Adopt a mindset that thinks *growth*. A student with a growth mindset expects to learn new things. When challenges arise, as they

[7] Not only helpful, but necessary, I would say.

inevitably do, acknowledge them and figure out how to meet the challenge. A growth mindset treats ability as something to be developed (see Dweck, 2016). If you engage yourself in this course, you can expect to use what you learn for the rest of your life.

The main title of this book is "Exploring Statistics." *Exploring* conveys the idea of uncovering something that was not apparent before. An attitude of *searching, wondering, checking*, and so forth is what I want to encourage. (Those who object to traditional NHST procedures are driven by this exploration motivation.) As for this book's subtitle, "Tales of Distributions," I'll have more to say about it as we go along.

Disciplines that Use Quantitative Data

Which disciplines use quantitative data? The list is long and more variable than the list I gave earlier. The examples and problems in this textbook, however, come from psychology, biology, sociology, education, medicine, politics, business, economics, forestry, and everyday life. Statistics is a powerful method for getting answers from data, and this makes it popular with investigators in a wide variety of fields.

Statistics is used in areas that might surprise you. As examples, statistics has been used to determine the effect of cigarette taxes on smoking among teenagers, the safety of a new surgical anesthetic, and the memory of young school-age children for pictures (which is as good as that of college students). Statistics show which diseases have an inheritance factor, how to improve short-term weather forecasts, and why giving intentional walks in baseball is a poor strategy. All these examples come from Statistics: *A Guide to the Unknown*, a book edited by Judith M. Tanur and others (1989). Written for those "without special knowledge of statistics," this book has 29 essays on topics as varied as those above.

In American history, the authorship of 12 of *The Federalist* papers was disputed for a number of years. (*The Federalist* papers were 85 short essays written under the pseudonym "Publius" and published in New York City newspapers in 1787 and 1788. Written by James Madison, Alexander Hamilton, and John Jay, the essays were designed to persuade the people of the state of New York to ratify the Constitution of the United States.) To determine authorship of the 12 disputed papers, each was graded with a quantitative *value analysis* in which the importance of such values as national security, a comfortable life, justice, and equality was assessed. The value analysis scores were compared with value analysis scores of papers known to have been written by Madison and Hamilton (Rokeach, Homant, & Penner, 1970). Another study, by Mosteller and Wallace, analyzed *The Federalist* papers using the frequency of words such as *by* and *to* (reported in Tanur et al., 1989). Both studies concluded that Madison wrote all 12 essays.

Here is an example from law. Rodrigo Partida was convicted of burglary in Hidalgo County, a border county in southern Texas. A grand jury rejected his motion for a new trial. Partida's attorney filed suit, claiming that the grand jury selection process discriminated against Mexican-Americans. In the end (*Castaneda v. Partida*, 430 U.S. 482 [1976]), Justice Harry

Blackmun of the U.S. Supreme Court wrote, regarding the number of Mexican-Americans on grand juries, "If the difference between the expected and the observed number is greater than two or three standard deviations, then the hypothesis that the jury drawing was random (is) suspect." In Partida's case, the difference was approximately 12 standard deviations, and the Supreme Court ruled that Partida's attorney had presented *prima facie* evidence. (*Prima facie* evidence is so good that one side wins the case unless the other side rebuts the evidence, which in this case did not happen.) *Statistics: A Guide to the Unknown* includes two essays on the use of statistics by lawyers.

Gigerenzer et al. (2007), in their public interest article on health statistics, point out that lack of statistical literacy among both patients and physicians undermines the information exchange necessary for informed consent and shared decision making. The result is anxiety, confusion, and undue enthusiasm for testing and treatment.

Whatever your current interests or thoughts about your future as a statistician, I believe you will benefit from this course. A successful statistics course teaches you to identify questions a set of data can answer; determine the statistical procedures that will provide the answers; carry out the procedures; and then, using plain English and graphs, tell the story the data reveal.

The best way for you to acquire all these skills (especially the part about telling the story) is to *engage* statistics. Engaged students are easily recognized; they are prepared for exams, are not easily distracted while studying, and generally finish assignments on time. *Becoming* an engaged student may not be so easy, but many have achieved it. Here are my recommendations. Read with the goal of understanding. Attend class. Do all the assignments (on time). Write down questions. Ask for explanations. Expect to understand. (Disclaimer: I'm not suggesting that you marry statistics, but just engage for this one course.)

Are you uncertain about whether your background skills are adequate for a statistics course? For most students, this is an unfounded worry. Appendix A, Getting Started, should help relieve your concerns.

What Do You Mean, "Statistics"?

Descriptive statistic
A number that conveys a particular characteristic of a set of data.

Mean
Arithmetic average; sum of scores divided by number of scores.

Inferential statistics
Method that uses sample evidence and probability to reach conclusions about unmeasurable populations.

The *Oxford English Dictionary* says that the word *statistics* came into use almost 250 years ago. At that time, *statistics* referred to a country's quantifiable political characteristics—characteristics such as population, taxes, and area. Statistics meant "state numbers." Tables and charts of those numbers turned out to be a very satisfactory way to compare different countries and to make projections about the future. Later, tables and charts proved useful to people studying trade (economics) and natural phenomena (science). Statistical thinking spread because it helped. Today, two different techniques are called *statistics*.

Descriptive statistics[8] produce a number or a figure that summarizes or describes a set of data. You are already familiar with some descriptive statistics. For example, you know about the arithmetic average, called the **mean**. You have probably known how to compute a mean since elementary school—just add up the numbers and divide the total by the number of entries. As you already know, the mean describes the central tendency of a set of numbers. The basic idea of descriptive statistics is simple: They summarize a set of data with one number or graph. This book covers about a dozen descriptive statistics.

The other statistical technique is **inferential statistics**. Inferential statistics use measurements from a sample to reach conclusions about a larger, *unmeasured* population. There is, of course, a problem with samples.

Samples always depend *partly* on the luck of the draw; chance helps determine the particular measurements you get.

If you have the measurements for the entire population, chance doesn't play a part—all the variation in the numbers is "true" variation. But with samples, some of the variation is the true variation in the population and some is just the chance ups and downs that go with a sample. Inferential statistics was developed as a way to account for the effects of chance that come with sampling. This book will cover about a dozen and a half inferential statistics.

Here is a textbook definition: Inferential statistics is a method that takes chance factors into account when samples are used to reach conclusions about populations. Like most textbook definitions, this one condenses many elements into a short sentence. Because the idea of using samples to understand populations is perhaps the *most important concept* in this course, please pay careful attention when elements of inferential statistics are explained.

Inferential statistics has proved to be a very useful method in scientific disciplines. Many other fields use inferential statistics, too, so I selected examples and problems from a variety of disciplines for this text and its auxiliary materials. *Null hypothesis significance testing*, which had a prominent place in our exploration tour, is an inferential statistics technique.

Here is an example from psychology that uses the NHST technique. Today, there is a lot of evidence that people remember the tasks they fail to complete better than the tasks they complete. This is known as the *Zeigarnik effect*. Bluma Zeigarnik asked participants in her experiment to do about 20 tasks, such as work a puzzle, make a clay figure, and construct a box from cardboard.[9] For each participant, half the tasks were interrupted before completion. Later, when the participants were asked to recall the tasks they worked on, they listed more of the interrupted tasks (average about 7) than the completed tasks (about 4).

One good question to start with is, "Did interrupting make a big difference or a small difference?" In this case, interruption produced about three additional memory items compared to the completion condition. This is a 75% difference, which seems like a big change, given our experience with tests of memory. The question of "How big is the difference?" can often be answered by calculating an *effect size index*.

---

[8] Boldface words and phrases are defined in the margin and also in Appendix D, Glossary of Words.

[9] A summary of this study can be found in Ellis (1938). The complete reference and all others in the text are listed in the References section at the back of the book.

clue to the future

> Calculating effect size indexes is first addressed in Chapter 5. It is also a topic in Chapters 9-14.

So, should you conclude that interruption improves memory? Not yet. It might be that interruption actually has no effect but that several *chance factors* happened to favor the interrupted tasks in Zeigarnik's particular experiment. One way to meet this objection is to conduct the experiment again. Similar results would lend support to the conclusion that interruption improves memory. A less expensive way to meet the objection is to use inferential statistics such as NHST.

NHST begins with the actual data from the experiment. It ends with a probability—the probability of obtaining data like those actually obtained if it is true that interruption has *no* effect on memory. If the probability is very small, you can conclude that interruption *does* affect memory. For Zeigarnik's data, the probability was tiny.

Now for the conclusion. One version might be, "After completing about 20 tasks, memory for interrupted tasks (average about 7) was greater than memory for completed tasks (average about 4). The approximate 75% difference cannot be attributed to chance because chance by itself would rarely produce a difference between two samples as large as this one." The words *chance* and *rarely* tell you that probability is an important element of inferential statistics.

My more complete answer to what I mean by "statistics" is Chapter 6 in *21st Century Psychology: A Reference Handbook* (Spatz, 2008). This 8-page chapter summarizes in words (no formulas) the statistical concepts usually covered in statistics courses. This chapter can orient you as you begin your study of statistics and later provide a review after you finish your course.

clue to the future

> The first part of this book is devoted to descriptive statistics (Chapters 2–6) and the second part to inferential statistics (Chapters 7–15). Inferential statistics is the more comprehensive of the two because it combines descriptive statistics, probability, and logic.

Statistics: A Dynamic Discipline

Many people continue to think of statistics as a collection of techniques that were developed long ago, that have not changed, and that will be the same in the future. That view is mistaken. Statistics is a dynamic discipline characterized by more than a little controversy. New techniques in both descriptive and inferential statistics continue to be developed. Controversy continues too, as you saw at the end of our exploration tour. To get a feel for the issues when the controversy entered the mainstream, see Dillon (1999) or Spatz (2000) for nontechnical summaries. For more technical explanations, see Nickerson (2000). To read about current approaches, see Erceg-Hurn and Mirosevich (2008), Kline (2013), or Cumming (2014).

In addition to controversy over techniques, attitudes toward data analysis shifted in recent years. The shift has been toward the idea of exploring data to see what it reveals and away from using statistical analyses to nail down a conclusion. This shift owes much of its impetus to John Tukey (1915–2000), who promoted Exploratory Data Analysis (Lovie, 2005). Tukey invented techniques such as the boxplot (Chapter 5) that reveal several characteristics of a data set simultaneously.

Today, statistics is used in a wide variety of fields. Researchers start with a phenomenon, event, or process that they want to understand better. They make measurements that produce numbers. The numbers are manipulated according to the rules and conventions of statistics. Based on the outcome of the statistical analysis, researchers draw conclusions and then write the story of their new understanding of the phenomenon, event, or process. Statistics is just one tool that researchers use, but it is often an essential tool.

Some Terminology

Like most courses, statistics introduces you to many new words. In statistics, most of the terms are used over and over again. Your best move, when introduced to a new term, is to *stop, read* the definition carefully, and *memorize* it. As the term continues to be used, you will become more and more comfortable with it. Making notes is helpful.

Populations and Samples

A **population** consists of all the scores of some specified group. A **sample** is a subset of a population. The population is the thing of interest. It is defined by the investigator and includes all cases. The following are some populations:

Family incomes of college students in the fall of 2017
Weights of crackers eaten by obese male students
Depression scores of Alaskans
Gestation times for human beings
Memory scores of human beings[10]

Population
All measurements of a specified group.

Sample
Measurements of a subset of a population.

[10] I didn't pull these populations out of thin air; they are all populations that researchers have gathered data on. Studies of these populations will be described in this book.

Investigators are always interested in populations. However, as you can determine from these examples, populations can be so large that not all the members can be studied. The investigator must often resort to measuring a sample that is small enough to be manageable. A sample taken from the population of incomes of families of college students might include only 40 students. From the last population on the list, Zeigarnik used a sample of 164.

Most authors of research articles carefully explain the characteristics of the samples they use. Often, however, they do not identify the *population*, leaving that task to the reader.

The answer to the question "What is the population?" depends on the specifics of a research area, but many researchers generalize generously. For example, for some topics it is reasonable to generalize from the results of a study on rats to "all mammals." In all cases, however, the reason for gathering data from a sample is to generalize the results to a larger population even though sampling introduces some uncertainty into the conclusions.

Parameters and Statistics

Parameter
Numerical or nominal characteristic of a population.

Statistic
Numerical or nominal characteristic of a sample.

A **parameter** is some numerical (number) or nominal (name) characteristic of a population. An example is the mean reading readiness score of all first-grade pupils in the United States. A **statistic** is some numerical or nominal characteristic of a sample. The mean reading readiness score of 50 first-graders is a statistic, and so is the observation that 45% are girls. A parameter is constant; it does not change unless the population itself changes. The mean of a population is exactly one number. Unfortunately, the parameter often cannot be computed because the population is unmeasurable. So, a statistic is used as an estimate of the parameter, although, as suggested before, statistics tend to differ from one sample to another. If you have five samples from the same population, you will probably have five different sample means. In sum, parameters are constant; statistics are variable.

Variables

Variable
Something that exists in more than one amount or in more than one form.

A **variable** is something that exists in more than one amount or in more than one form. Height and eye color are both variables. The notation *67 inches* is a numerical way to identify a group of persons who are similar in height. Of course, there are many other groups, each with an identifying number. Blue and brown are common eye colors, which might be assigned the numbers *0* and *1*. All participants represented by *0* have the same eye color. I will often refer to numbers like *67* and *0* as *scores* or *test scores*. A score is simply the result of measuring a variable.

Quantitative Variables

Scores on **quantitative variables** tell you the *degree* or *amount* of the thing being measured. At the very least, a larger score indicates more of the variable than a smaller score does.

Quantitative variable
Variable whose scores indicate different amounts.

Continuous Variables.

Continuous variables are quantitative variables whose scores can be any value or intermediate value over the variable's possible range. The continuous memory scores in Zeigarnik's experiment make up a quantitative, continuous variable. *Number of tasks recalled* scores come in whole numbers such as 4 or 7, but it seems reasonable to assume that the thing being measured, memory, is a continuous variable. Thus, of two participants who both scored 7, one just barely got 7 and the other almost scored 8. Picture the continuous variable, *recall*, as **Figure 1.1**.

Continuous variable
A quantitative variable whose scores can be any amount.

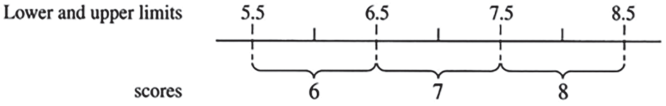


FIGURE 1.1 The lower and upper limits of recall scores of 6, 7, and 8

Figure 1.1 shows that a score of 7 is used for a range of possible recall values—the range from 6.5 to 7.5. The number 6.5 is the **lower limit** and 7.5 is the **upper limit** of the score of 7. The idea is that recall can be any value between 6.5 and 7.5, but that all the recall values in this range are expressed as 7. In a similar way, a charge indicator value of 62% on your cell phone stands for all the power values between 61.5% (the lower limit) and 62.5% (the upper limit).

Lower limit
Bottom of the range of possible values that a measurement on a continuous variable can have.

Upper limit
Top of the range of possible values that a measurement on a continuous variable can have.

Sometimes scores are expressed in tenths, hundredths, or thousandths. Like integers, these scores have lower and upper limits that extend halfway to the next value on the quantitative scale.

Discrete Variables.

Some quantitative variables are classified as **discrete variables** because intermediate values are not possible. The number of siblings you have, the number of times you've been hospitalized, and how many pairs of shoes you have are examples. Intermediate scores such as 2½ just don't make sense.

Discrete variable
Variable for which intermediate values between scores are not meaningful.

Categorical Variables

Categorical variable
Variable whose scores differ in kind, not amount.

Categorical variables (also called *qualitative variables*) produce scores that differ in *kind* and not *amount*. Eye color is a categorical variable. Scores might be expressed as *blue* and *brown* or as *0* and *1*, but substituting a number for a name does not make eye color a quantitative variable. American political affiliation is a categorical variable with values of Democrat, Republican, Independent, and Other. College major is another categorical variable.

Some categorical variables have the characteristic of *order*. College standing has ordered measurements of senior, junior, sophomore, and freshman. Military rank is a categorical variable with scores such as sergeant, corporal, and private. Categorical variables such as color and gender do not have an inherent order. All categorical variables produce discrete scores, but not all discrete scores are from a categorical variable.

Problems and Answers

At the beginning of this chapter, I urged you to engage statistics. Have you? For example, did you read the footnotes? Have you looked up any words you weren't sure of? (How near are you to dictionary definitions when you study?) Have you read a paragraph a second time, wrinkled your brow in concentration, made notes in the book margin, or promised yourself to ask your instructor or another student about something you aren't sure of? *Engagement* shows up as activity. Best of all, the activity at times is a nod to yourself and a satisfied, "Now I understand."

From time to time, I will use my best engagement tactic: I'll give you a set of problems so that you can practice what you have just been reading about. Working these problems correctly is additional evidence that you have been engaged. You will find the answers at the end of the book in Appendix G. Here are some suggestions for *efficient* learning.

1. Buy yourself a notebook or establish a file for statistics. Save your work there. When you make an error, don't remove it—note the error and rework the problem correctly. Seeing your error later serves as a reminder of what not to do on a test. If you find that I have made an error, write to me with a reminder of what not to do in the next edition.
2. Never, never look at an answer before you have worked the problem (or at least tried twice to work the problem).
3. For each set of problems, work the first one and then immediately check your answer against the answer in the book. If you make an error, find out why you made it—faulty understanding, arithmetic error, or whatever.
4. Don't be satisfied with just doing the math. If a problem asks for an interpretation, write out your interpretation.
5. When you finish a chapter, go back over the problems immediately, reminding yourself of the various techniques you have learned.
6. Use any blank spaces near the end of the book for your special notes and insights.

Now, here is an opportunity to see how actively you have been reading.

PROBLEMS

1.1. The history-of-statistics tour began with what easy-to-remember date?
1.2. The dominant approach to inferential statistics that is under attack is called ___________.
1.3. Identify each number below as coming from a quantitative variable or a categorical variable.
  a. 65 – seconds to work a puzzle
  b. 319 – identification number for intellectual disability in the American Psychiatric Association manual
  c. 3 – group identification for small-cup daffodils
  d. 4 – score on a high school advanced placement exam
  e. 81 – milligrams of aspirin
1.4. Place lower and upper limits beside the continuous variables. Write *discrete* beside the others.
  a. ______________________ 20, seconds to work a puzzle
  b. ______________________ 14, number of concerts attended
  c. ______________________ 3, birth order
  d. ______________________ 10, speed in miles per hour
1.5. Write a paragraph that gives the definitions of *population, sample, parameter,* and *statistic* and the relationships among them.
1.6. Two kinds of statistics are ____________ statistics and ____________ statistics. Fill each blank with the correct adjective.
  a. To reach a conclusion about an unmeasured population, use ___________ statistics.
  b. ____________ statistics take chance into account to reach a conclusion.
  c. ____________ statistics are numbers or graphs that summarize a set of data.

Scales of Measurement

Numbers mean different things in different situations. Consider three answers that appear to be identical but are not:

What number were you wearing in the race?	"5"
What place did you finish in?	"5"
How many minutes did it take you to finish?	"5"

The three 5s all look the same. However, the three variables (identification number, finish place, and time) are quite different. Because of the difference in what the variables measure, each 5 has a different interpretation.

To illustrate this difference, consider another person whose answers to the same three questions were 10, 10, and 10. If you take the first question by itself and know that the two people had scores of 5 and 10, what can you say? You can say that the first runner was different from the second, but *that is all*. (Think about this until you agree.) On the second question, with scores of 5 and 10, what can you say? You can say that the first runner was faster than the second and, of course, that they are different.

Comparing the 5 and 10 on the third question, you can say that the first runner was twice as fast as the second runner (and, of course, was faster and different).

The point of this discussion is to draw the distinction between the *thing* you are interested in and the *number* that stands for the thing. Much of your experience with numbers has been with pure numbers or quantitative measures such as time, length, and amount. Four and two have a relationship of *twice as much* and *half as much*. And, for distance and seconds, four *is* twice two; for amounts, two is half of four. But these relationships do not hold when numbers are used to measure some things. For example, for political race finishes, *twice* and *half* are not helpful. Second place is not half or twice anything compared to fourth place.

S. S. Stevens (1946) identified four different *scales of measurement*, each of which carries a different set of information. Each scale uses numbers, but the information that can be inferred from the numbers differs. The four scales are *nominal*, *ordinal, interval,* and *ratio*.

Nominal scale
Measurement scale in which numbers serve only as labels and do not indicate any quantitative relationship.

In the **nominal scale**, numbers are used simply as names and have no real quantitative value. Numerals on sports uniforms are an example. Thus, 45 is *different* from 32, but that is all you can say. The person represented by 45 is not "more than" the person represented by 32, and certainly it would be meaningless to calculate the mean of the two numbers. Examples of nominal variables include psychological diagnoses, personality types, and political parties. Psychological diagnoses, like other nominal variables, consist of a set of categories. People are assessed and then classified into one of the categories. The categories have both a name (such as posttraumatic stress disorder or autism spectrum disorder) and a number (309.81 and 299.00, respectively). On a nominal scale, the numbers mean only that the categories are different. In fact, for a nominal scale variable, the numbers could be assigned to categories at random. Of course, all things that are alike must have the same number.

Ordinal scale
Measurement scale in which numbers are ranks; equal differences between numbers do not represent equal differences between the things measured.

The **ordinal scale** has the characteristic of the nominal scale (different numbers mean different things) plus the characteristic of indicating *greater than* or *less than*. In the ordinal scale, the object with the number 3 has less or more of something than the object with the number 5. Finish places in a race are an example of an ordinal scale. The runners finish in rank order, with 1 assigned to the winner, 2 to the runner-up, and so on. Here, 1 means less time than 2. Judgments about anxiety, quality, and recovery often correspond to an ordinal scale. "Much improved," "improved," "no change," and "worse" are levels of an ordinal recovery variable. Ordinal scales are characterized by *rank order*.

The third kind of scale is the **interval scale**, which has the properties of both the nominal and ordinal scales plus the additional property that *intervals between the numbers are equal.* "Equal interval" means that the distance between the things represented by 2 and 3 is the same as the distance between the things represented by 3 and 4. Temperature is measured on an interval scale. The difference in temperature between 10°C and 20°C is the same as the difference between 40°C and 50°C. The Celsius thermometer, like all interval scales, has an arbitrary zero point. On the Celsius thermometer, this zero point is the freezing point of water at sea level. Zero degrees on this scale does not mean the complete absence of heat; it is simply a convenient starting point. With interval data, there is one restriction: You may not make simple ratio statements. You may not say that 100° is twice as hot as 50° or that a person with an IQ of 60 is half as intelligent as a person with an IQ of 120.[11]

Interval scale
Measurement scale in which equal differences between numbers represent equal differences in the thing measured. The zero point is arbitrarily defined.

The fourth kind of scale, the **ratio scale**, has all the characteristics of the nominal, ordinal, and interval scales plus one other: It has a *true zero point*, which indicates a complete absence of the thing measured. On a ratio scale, zero means "none." Height, weight, and time are measured with ratio scales. Zero height, zero weight, and zero time mean that no amount of these variables is present. With a true zero point, you can make ratio statements such as 16 kilograms is four times heavier than 4 kilograms.[12] **Table 1.1** summarizes the major differences among the four scales of measurement.

Ratio scale
Measurement scale with characteristics of interval scale; also, zero means that none of the thing measured is present.

TABLE 1.1 Characteristics of the four scales of measurement

	Scale characteristics			
Scale of measurement	Different numbers for different things	Numbers convey greater than and less than	Equal differences mean equal amounts	Zero means none of what was measured was detected
Nominal	Yes	No	No	No
Ordinal	Yes	Yes	No	No
Interval	Yes	Yes	Yes	No
Ratio	Yes	Yes	Yes	Yes

[11] Convert 100°C and 50°C to Fahrenheit (F = 1.8C + 32) and suddenly the "twice as much" relationship disappears.

[12] Convert 16 kilograms and 4 kilograms to pounds (1 kg = 2.2 lbs) and the "four times heavier" relationship is maintained.

Knowing the distinctions among the four scales of measurement will help you in two tasks in this course. The kind of *descriptive statistics* you can compute from numbers depends, in part, on the scale of measurement the numbers represent. For example, it is senseless to compute a mean of numbers on a nominal scale. Calculating a mean Social Security number, a mean telephone number, or a mean psychological diagnosis is either a joke or evidence of misunderstanding numbers.

Understanding scales of measurement is sometimes important in choosing the kind of *inferential statistic* that is appropriate for a set of data. If the dependent variable (see next section) is a nominal variable, then a chi square analysis is appropriate (Chapter 14). If the dependent variable is a set of ranks (ordinal data), then a nonparametric statistic is required (Chapter 15). Most of the data analyzed with the techniques described in Chapters 7–13 are interval and ratio scale data.

The topic of scales of measurement is controversial among statisticians. Part of the controversy involves viewpoints about the underlying thing you are interested in and the number that represents the thing (Wuensch, 2005). In addition, it is sometimes difficult to classify some of the variables used in the social and behavioral sciences. Often they appear to fall between the ordinal scale and the interval scale. For example, a score may provide more information than simply rank, but equal intervals cannot be proven. Examples include aptitude and ability tests, personality measures, and intelligence tests. In such cases, researchers generally treat the scores as if they were interval scale data.

Statistics and Experimental Design

Here is a story that will help you distinguish between statistics (applying straight logic) and experimental design (observing what actually happens). This is an excerpt from a delightful book by E. B. White, *The Trumpet of the Swan* (1970, pp. 63–64).

> The fifth-graders were having a lesson in arithmetic, and their teacher, Miss Annie Snug, greeted Sam with a question.
>
> "Sam, if a man can walk three miles in one hour, how many miles can he walk in four hours?" "It would depend on how tired he got after the first hour," replied Sam. The other pupils roared. Miss Snug rapped for order.
>
> "Sam is quite right," she said. "I never looked at the problem that way before. I always supposed that man could walk twelve miles in four hours, but Sam may be right: that man may not feel so spunky after the first hour. He may drag his feet. He may slow up."
>
> Albert Bigelow raised his hand. "My father knew a man who tried to walk twelve miles, and he died of heart failure," said Albert.
>
> "Goodness!" said the teacher. "I suppose *that* could happen, too."

“Anything can happen in four hours,” said Sam. “A man might develop a blister on his heel. Or he might find some berries growing along the road and stop to pick them. That would slow him up even if he wasn’t tired or didn’t have a blister.”

“It would indeed,” agreed the teacher. “Well, children, I think we have all learned a great deal about arithmetic this morning, thanks to Sam Beaver.”

Everyone had learned how careful you have to be when dealing with figures.

Statistics involves the manipulation of numbers and the conclusions based on those manipulations (Miss Snug). Experimental design (also called research methods) deals with all the things that influence the numbers you get (Sam and Albert). **Figure 1.2** illustrates these two approaches to getting an answer. This text could have been a “pure” statistics book, from which you would learn to analyze numbers without knowing where they came from or what they referred to. You would learn about statistics, but such a book would be dull, dull, dull. On the other hand, to describe procedures for collecting numbers is to teach experimental design—and this book is for a statistics course. My solution to this conflict is generally to side with Miss Snug but to include some aspects of experimental design throughout the book. Knowing experimental design issues is especially important when it comes time to interpret a statistical analysis. Here’s a start on experimental design.

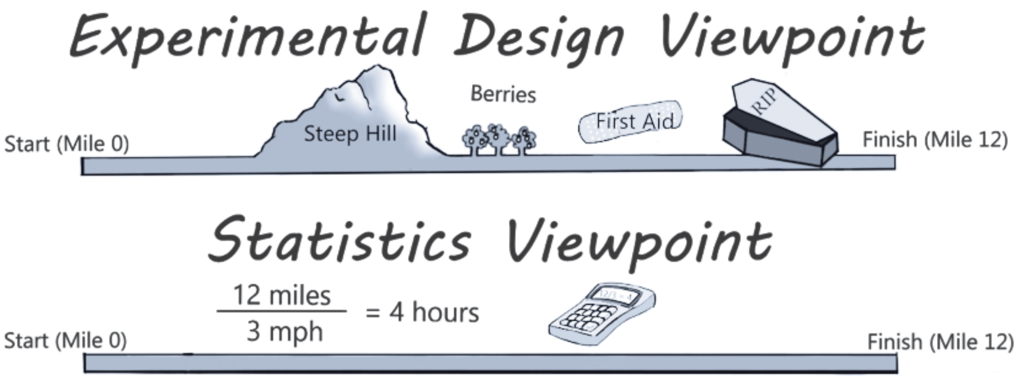


FIGURE 1.2 **Travel time from an experimental design viewpoint and a statistical viewpoint**

Experimental Design Variables

The overall task of an experimenter is to discover relationships among variables. Variables are things that vary, and researchers have studied personality, health, gender, anger, caffeine, memory, beliefs, age, skill.... (I’m sure you get the picture—almost anything can be a variable.)

Independent and Dependent Variables

Independent variable
Variable controlled by the researcher; changes in this variable may produce changes in the dependent variable.

Dependent variable
Observed variable that is expected to change as a result of changes in the independent variable in an experiment.

Level
One value of the independent variable.

Treatment
One value (or level) of the independent variable.

A simple experiment has two major variables, the **independent variable** and the **dependent variable**. In the simplest experiment, the researcher selects two values of the independent variable for investigation. Values of the independent variable are usually called **levels** and sometimes called **treatments**.

The basic idea is that the researcher finds or creates two groups of participants that are similar except for the independent variable. These individuals are measured on the dependent variable. The question is whether the data will allow the experimenter to claim that the values on the dependent variable *depend* on the level of the independent variable.

The values of the dependent variable are found by *measuring* or *observing* participants in the investigation. The dependent variable might be scores on a personality test, number of items remembered, or whether or not a passerby offered assistance. For the independent variable, the two groups might have been selected because they were already different—in age, gender, personality, and so forth. Alternatively, the experimenter might have produced the difference in the two groups by an experimental manipulation such as creating different amounts of anxiety or providing different levels of practice.

An example might help. Suppose for a moment that as a budding gourmet cook you want to improve your spaghetti sauce. One of your buddies suggests adding marjoram. To investigate, you serve spaghetti sauce at two different gatherings. For one group of guests, the sauce is spiced with marjoram; for the other it is not. At both gatherings, you count the number of favorable comments about the spaghetti sauce. *Stop reading; identify the independent and the dependent variables.*

The dependent variable is the number of favorable comments, which is a measure of the taste of the sauce. The independent variable is marjoram, which has two levels: present and absent.

Extraneous Variables

Extraneous variable
Variable other than the independent variable that may affect the dependent variable.

One of the pitfalls of experiments is that every situation has other variables besides the independent variable that might possibly be responsible for the changes in the dependent variable. These other variables are called **extraneous variables**. In the story, Sam and Albert noted several extraneous variables that could influence the time to walk 12 miles.

Are there any extraneous variables in the spaghetti sauce example? Oh yes, there are many, and just one is enough to raise suspicion about a conclusion that relates the taste of spaghetti sauce to marjoram. Extraneous variables include the amount and quality of the other ingredients in the sauce, the spaghetti itself, the "party moods" of the two groups, and how hungry everyone was. If any of these extraneous variables was actually operating, it weakens the claim that a difference in the comments about the sauce is the result of the presence or absence of marjoram.

The simplest way to remove an extraneous variable is to be sure that all participants are equal on that variable. For example, you can ensure that the sauces are the same except for marjoram by mixing up the ingredients, dividing it into two batches, adding marjoram to one batch but not the other, and then cooking. The "party moods" variable can be controlled (equalized) by conducting the taste test in a laboratory. Controlling extraneous variables is a complex topic covered in courses that focus on research methods and experimental design.

In many experiments, it is impossible or impractical to control all the extraneous variables. Sometimes researchers think they have controlled them all, only to find that they did not. The effect of an uncontrolled extraneous variable is to prevent a simple cause-and-effect conclusion. Even so, if the dependent variable changes when the independent variable changes, something is going on. In this case, researchers can say that the two variables are *related*, but that other variables may play a part, too.

At this point, you can test your understanding by engaging yourself with these questions: What were the independent and dependent variables in the Zeigarnik experiment? How many levels of the independent variable were there?[13]

How well did Zeigarnik control extraneous variables? For one thing, each participant was tested at both levels of the independent variable. That is, the recall of each participant was measured for interrupted tasks and for completed tasks. One advantage of this technique is that it naturally controls many extraneous variables. Thus, extraneous variables such as age and motivation were exactly the same for tasks that were interrupted as for tasks that were not because the same people contributed scores to both levels.

At various places in the following chapters, I will explain experiments and the statistical analyses using the terms *independent* and *dependent variables*. These explanations usually assume that all extraneous variables were controlled; that is, you may assume that the experimenter knew how to design the experiment so that changes in the dependent variable could be attributed correctly to changes in the independent variable. However, I present a few investigations (like the spaghetti sauce example) that I hope you recognize as being so poorly designed that conclusions cannot be drawn about the relationship between the independent variable and dependent variable. Be alert.

---

[13] Try for answers. Then, if need be, here's a hint: First, identify the dependent variable; for the dependent variable, you don't know values until data are gathered. Next, identify the independent variable; you can tell what the values of the independent variable are just from the description of the design.

Here's my summary of the relationship between statistics and experimental design. Researchers suspect that there is a relationship between two variables. They design and conduct an experiment; that is, they choose the levels of the independent variable (treatments), control the extraneous variables, and then measure the participants on the dependent variable. The measurements (data) are analyzed using statistical procedures. Finally, the researcher tells a story that is consistent with the results obtained and the procedures used.

Statistics and Philosophy

The two previous sections directed your attention to the relationship between statistics and experimental design; this section will direct your thoughts to the place of statistics in the grand scheme of things.

Epistemology
The study or theory of the nature of knowledge.

Explaining the grand scheme of things is the task of philosophy and, over the years, many schemes have been advanced. For a scheme to be considered a grand one, it has to address **epistemology**—that is, to propose answers to the question: How do we acquire knowledge?

Both *reason* and *experience* have been popular answers among philosophers.[14] For those who emphasize the importance of reason, mathematics has been a continuing source of inspiration. Classical mathematics starts with axioms that are assumed to be true. Theorems are thought up and are then proved by giving axioms as reasons. Once a theorem is proved, it can be used as a reason in a proof of other theorems.

Statistics has its foundations in mathematics and thus, a statistical analysis is based on reason. As you go about the task of calculating $\overline{X}$ or $\hat{s}$, finding confidence intervals, and telling the story of what they mean, know deep down that you are engaged in logical reasoning. (Experimental design is more complex; it includes experience and observation as well as reasoning.)

In the 19th century, science concentrated on observation and description. The variation that always accompanied a set of observations was thought to be due to imprecise observing, imprecise instruments, or failure of nature to "hit the mark." During the 20th century, however, statistical methods such as NHST revolutionized the philosophy of science by focusing on the variation that was always present in data (Salsburg, 2001; Gould, 1996). Focusing on variation allowed changes in the data to be associated with particular causes.

As the 21st century approached, flaws in the logic of NHST statistics began to be recognized. In addition to logical flaws, the practices of researchers and journal editors (such as requiring a statistical analysis to show that $p < .05$) came under scrutiny. This concern with how science is conducted has led to changes in how data are analyzed and how information is shared. The practice of statistics is in transition.

---

[14] In philosophy, those who emphasize reason are rationalists and those who emphasize experience are empiricists.

Let's move from formal descriptions of philosophy to a more informal one. A very common task of most human beings can be described as *trying to understand*. Statistics has helped many in their search for better understanding, and it is such people who have recommended (or demanded) that statistics be taught in school. A reasonable expectation is that you, too, will find statistics useful in your future efforts to understand and persuade.

Speaking of persuasion, you have probably heard it said, "You can prove anything with statistics." The implied message is that a conclusion based on statistics is suspect because statistical methods are unreliable. Well, it just isn't true that statistical methods are unreliable, but it is true that people can misuse statistics (just as any tool can be misused). One of the great advantages of studying statistics is that you get better at recognizing statistics that are used improperly.

Statistics: Then and Now

Statistics began with counting, which, of course, was prehistory. The origin of the mean is almost as obscure. It was in use by the early 1700s, but no one is credited with its discovery. Graphs, however, began when J. H. Lambert, a Swiss-German scientist and mathematician, and William Playfair, an English political economist, invented and improved graphs in the period 1765 to 1800 (Tufte, 2001).

The Royal Statistical Society was established in 1834 by a group of Englishmen in London. Just 5 years later, on November 27, 1839, at 15 Cornhill in Boston, a group of Americans founded the American Statistical Society. Less than 3 months later, for a reason that you can probably figure out, the group changed its name to the American Statistical Association, which continues today (*www.amstat.org*).

According to Walker (1929), the first university course in statistics in the United States was probably "Social Science and Statistics," taught at Columbia University in 1880. The professor was a political scientist, and the course was offered in the economics department. In 1887, at the University of Pennsylvania, separate courses in statistics were offered by the departments of psychology and economics. By 1891, Clark University, the University of Michigan, and Yale had been added to the list of schools that taught statistics, and anthropology had been added to the list of departments. Biology was added in 1899 (Harvard) and education in 1900 (Columbia).

You might be interested in when statistics was first taught at your school and in what department. College catalogs are probably your most accessible source of information.

This course provides you with the opportunity to improve your ability to understand and use statistics. Kirk (2008) identifies four levels of statistical sophistication:

Category 1—those who understand statistical presentations
Category 2—those who understand, select, and apply statistical procedures
Category 3—applied statisticians who help others use statistics
Category 4—mathematical statisticians who develop new statistical techniques and discover new characteristics of old techniques

I hope that by the end of your statistics course, you will be well along the path to becoming a Category 2 statistician.

How to Analyze a Data Set

The end point of analyzing a data set is a story that explains the relationships among the variables in the data set. I recommend that you analyze a data set in three steps. The first step is exploratory. Read all the information and examine the data. Calculate descriptive statistics and focus on the differences that are revealed. In this textbook, descriptive statistics are emphasized in Chapters 2 through 6 and include graphs, means, and effect size indexes. Calculating descriptive statistics helps you develop preliminary ideas for your story (Step 3). The second step is to answer the question, What are the effects that chance could have on the descriptive statistics I calculated? An answer requires inferential statistics (Chapter 7 through Chapter 15). The third step is to write the story the data reveal. Incorporate the descriptive and inferential statistics to support the conclusions in the story. Of course, the skills you've learned and taught yourself about composition will be helpful as you compose and write your story. Don't worry about length; most good statistical stories about simple data sets can be told in one paragraph.

Write your story using journal style, which is quite different from textbook style. Textbook style, at least this textbook, is chatty, redundant, and laced with footnotes.[15] Journal style, on the other hand, is terse, formal, and devoid of footnotes. Paragraphs labeled Interpretation in Appendix G, this textbook's answer section, are examples of journal style. And for guidance in writing up an entire study, see Appelbaum et al. (2018).

Helpful Features of This Book

At various points in this chapter, I encouraged your engagement in statistics. Your active participation is *necessary* if you are to learn statistics. For my part, I worked to organize this book and write it in a way that encourages active participation. Here are some of the features you should find helpful.

Objectives

Each chapter begins with a list of skills the chapter is designed to help you acquire. Read this list of objectives first to find out what you are to learn to do. Then thumb through the chapter and read the headings. Next, study the chapter, working all the problems *as you come to them*. Finally, reread the objectives. Can you meet each one? If so, put a check mark beside that objective.

---

[15] You are reading the footnotes, aren't you? Your answer — "Well, yes, it seems I am."

Problems and Answers

The problems in this text are in small groups within the chapter rather than clumped together at the end. This encourages you to read a little and work problems, followed by more reading and problems. Psychologists call this pattern spaced practice. Spaced practice patterns lead to better performance than massed practice patterns. The problems come from a variety of disciplines; the answers are in Appendix G.

Some problems are conceptual and do not require any arithmetic. Think these through and write your answers. Being able to calculate a statistic is almost worthless if you cannot explain in English what it means. Writing reveals how thoroughly you understand. To emphasize the importance of explanations, I highlited **Interpretation** in the answers in Appendix G. On occasion, problems or data sets are used again, either later in that chapter or in another. If you do not work the problem when it is first presented, you are likely to be frustrated when it appears again. To alert you, I have put an asterisk (*) beside problems that are used again.

At the end of many chapters, comprehensive problems are marked with a CP Working these problems requires knowing most of the material in the chapter. For most students, it is best to work all the problems, but be sure you can work those marked with a CP

clue to the future

Often a concept is presented that will be used again in later chapters. These ideas are separated from the rest of the text in a box labeled "Clue to the Future." You have already seen two of these "Clues" in this chapter. Attention to these concepts will pay dividends later in the course.

error detection

I have boxed in, at various points in the book, ways to detect errors. Some of these "Error Detection" tips will also help you better understand the concept. Because many of these checks can be made early, they can prevent the frustrating experience of getting an impossible answer when the error could have been caught in Step 2.

Figure and Table References

Sometimes the words Figure and Table are in boldface print. This means that you should examine the figure or table at that point. Afterward, it will be easy for you to return to your place in the text—just find the boldface type.

Transition Passages

At six places in this book, there are major differences between the material you just finished and the material in the next section. "Transition Passages," which describe the differences, separate these sections.

Glossaries

This book has three separate glossaries of words, symbols, and formulas.

1. *Words*. The first time an important word is used in the text, it appears in boldface type accompanied by a definition in the margin. In later chapters, the word may be boldfaced again, but margin definitions are not repeated. Appendix D is a complete glossary of words (p. 401). I suggest you mark this appendix.
2. *Symbols*. Statistical symbols are defined in Appendix E (p. 405). Mark it too.
3. *Formulas*. Formulas for all the statistical techniques used in the text are printed in Appendix F (p. 407), in alphabetical order according to the name of the technique.

Computers, Calculators, and Pencils

Computer software, calculators, and pencils with erasers are all tools used at one time or another by statisticians. Any or all of these devices may be part of the course you are taking. Regardless of the calculating aids that you use, however, your task is the same:

- Read a problem.
- Decide what statistical procedure to use.
- Apply that procedure using the tools available to you.
- Write an interpretation of the results.

Pencils, calculators, and software represent, in ascending order, tools that are more and more error-free. People who routinely use statistics routinely use computers. You may or may not use one at this point. Remember, though, whether you are using a software program or not, your principal task is to understand and describe.

For many of the worked examples in this book, I included the output of a popular statistical software program, IBM SPSS.[16] If your course includes IBM SPSS, these tables should help familiarize you with the program.

[16] The original name of the program was Statistical Package for the Social Sciences.

Concluding Thoughts

The first leg of your exploration of statistics is complete. Some of the ideas along the path are familiar and perhaps a few are new or newly engaging. As your course progresses, you will come to understand what is going on in many statistical analyses, and you will learn paths to follow if you analyze data that you collect yourself.

This book is a fairly complete introduction to elementary statistics. Of course, there is lots more to statistics, but there is a limit to what you can do in one term. Even so, exploration of paths not covered in a textbook can be fun. Encyclopedias, both general and specialized, often reward such exploration. Try the *Encyclopedia of Statistics in Behavioral Sciences or the International Encyclopedia of the Social and Behavioral Sciences*. Also, when you finish this course (but before any final examination), I recommend Chapter 16, the last chapter in this book. It is an overview/integrative chapter.

Most students find that this book works well as a textbook in their statistics course. I recommend that you keep the book after the course is over to use as a *reference* book. In courses that follow statistics and even after leaving school, many find themselves looking up a definition or reviewing a procedure.[17] Thus, a familiar textbook becomes a valued guidebook that serves for years into the future. For me, exploring statistics and using them to understand the world is quite satisfying. I hope you have a similar experience.

PROBLEMS

1.7. Name the four scales of measurement identified by S. S. Stevens.

1.8. Give the properties of each of the scales of measurement.

1.9. Identify the scale of measurement in each of the following cases.

- **a.** Geologists have a "hardness scale" for identifying different rocks, called Mohs' scale. The hardest rock (diamond) has a value of 10 and will scratch all others. The second hardest will scratch all but the diamond, and so on. Talc, with a value of 1, can be scratched by every other rock. (A fingernail, a truly handy field-test instrument, has a value between 2 and 3.)
- **b.** The volumes of three different cubes are 40, 64, and 65 cubic inches.
- **c.** Three different highways are identified by their numbers: 40, 64, and 65.
- **d.** Republicans, Democrats, Independents, and Others are identified on the voters' list with the numbers 1, 2, 3, and 4, respectively.
- **e.** The winner of the Miss America contest was Miss New York; the two runners-up were Miss Ohio and Miss California.[18]
- **f.** The prices of the three items are \$3.00, \$10.00, and \$12.00.
- **g.** She earned three degrees: B.A., M.S., and Ph.D.

---

[17] This book's index is unusually extensive. If you make margin notes, they will help too.

[18] Contest winners have come most frequently from these states, which have had six winners each.

1.10. Undergraduate students conducted the three studies that follow. For each study, identify the dependent variable, the independent variable, the number of levels of the independent variable, and the names of the levels of the independent variable.

- **a.** Becca had students in a statistics class rate a résumé, telling them that the person had applied for a position that included teaching statistics at their college. The students rated the résumé on a scale of 1 (not qualified) to 10 (extremely qualified). All the students received identical résumés, except that the candidate's first name was Jane on half the résumés and John on the other half.
- **b.** Michael's participants filled out the Selfism scale, which measures narcissism. (Narcissism is neurotic self-love.) In addition, students were classified as first-born, second-born, and later-born.
- **c.** Johanna had participants read a description of a crime and "Mr. Anderson," the person convicted of the crime. For some participants, Mr. Anderson was described as a janitor. For others, he was described as a vice president of a large corporation. For still others, no occupation was given. After reading the description, participants recommended a jail sentence (in months) for Mr. Anderson.

1.11. Researchers who are now well known conducted the three classic studies that follow. For each study, identify the dependent variable, the independent variable, and the number and names of the levels of the independent variable. Complete items i and ii.

- **a.** Theodore Barber hypnotized 25 people, giving each a series of suggestions. The suggestions included arm rigidity, hallucinations, color blindness, and enhanced memory. Barber counted the number of suggestions the participants complied with (the mean was 4.8). For another 25 people, he simply asked them to achieve the best score they could (but no hypnosis was used). This second group was given the same suggestions, and the number complied with was counted (the mean was 5.1). (See Barber, 1976.)
    - **i.** Identify a nominal variable and a statistic.
    - **ii.** In a sentence, describe what Barber's study shows.
- **b.** Elizabeth Loftus had participants view a film clip of a car accident. Afterward, some were asked, "How fast was the car going?" and others were asked, "How fast was the car going when it passed the barn?" (There was no barn in the film.) A week later, Loftus asked the participants, "Did you see a barn?" If the barn had been mentioned earlier, 17% said yes; if it had not been mentioned, 3% said yes. (See Loftus, 1979.)
    - **i.** Identify a population and a parameter.
    - **ii.** In a sentence, describe what Loftus's study shows.
- **c.** Stanley Schachter and Larry Gross gathered data from obese male students for about an hour in the afternoon. At the end of this time, a clock on the wall was correct (5:30 p.m.) for 20 participants, slow (5:00 p.m.) for 20 others, and fast (6:00 p.m.) for 20 more. The actual time, 5:30, was the usual dinnertime for these students. While participants filled out a final questionnaire, Wheat Thins® were freely available. The weight of the crackers each student consumed was measured. The means were

as follows: 5:00 group—20 grams; 5:30 group—30 grams; 6:00 group—40 grams. (See Schachter and Gross, 1968.)

i. Identify a ratio scale variable.

ii. In a sentence, describe what this study shows.

1.12. There are uncontrolled extraneous variables in the study described here. Name as many as you can. Begin by identifying the dependent and independent variables.

An investigator concluded that Textbook A was better than Textbook B after comparing the exam scores of two statistics classes. One class met MWF at 10:00 a.m. for 50 minutes, used Textbook A, and was taught by Dr. X. The other class met for 2.5 hours on Wednesday evening, used Textbook B, and was taught by Dr. Y. All students took the same comprehensive test at the end of the term. The mean score for Textbook A students was higher than the mean score for Textbook B students.

1.13. In philosophy, the study of the nature of knowledge is called _____.

1.14. a. The two approaches to epistemology identified in the text are _____ and _____.

b. Statistics has its roots in _____.

1.15. Your textbook recommends a three-step approach to analyzing a data set. Summarize the steps.

1.16. Read the objectives at the beginning of this chapter. Responding to them will help you consolidate what you have learned.

KEY TERMS

Categorical variable (p. 12)
Continuous variable (p. 11)
Dependent variable (p. 18)
Descriptive statistics (p. 6)
Discrete variable (p. 11)
Epistemology (p. 20)
Extraneous variable (p. 18)
Independent variable (p. 18)
Inferential statistics (p. 6)
Interval scale (p. 15)
Level (p. 18)
Lower limit (p. 11)
Mean (p. 7)
Nominal scale (p. 14)
Ordinal scale (p. 14)
Parameter (p. 10)
Population (p. 9)
Qualitative variable (p. 12)
Quantitative variable (p. 11)
Ratio scale (p. 15)
Sample (p. 9)
Statistic (p. 10)
Treatment (p. 18)
Upper limit (p. 11)
Variable (p. 10)

Transition Passage

To Descriptive Statistics

STATISTICAL TECHNIQUES ARE often categorized as descriptive statistics and inferential statistics. The next five chapters are about descriptive statistics. You are already familiar with some of these descriptive statistics, such as the mean, range, and bar graphs. Others may be less familiar—the correlation coefficient, effect size index, and boxplot. All of these and others that you study will be helpful in your efforts to understand data.

The phrase *Exploring Data* appears in three of the chapter titles that follow. This phrase is a reminder to approach a data set with the attitude of an explorer, an attitude of *What can I find here?* Descriptive statistics are especially valuable in the early stages of an analysis as you explore what the data have to say. Later, descriptive statistics are essential when you convey your story of the data to others. In addition, many descriptive statistics have important roles in the inferential statistical techniques that are covered in later chapters. Let's get started.