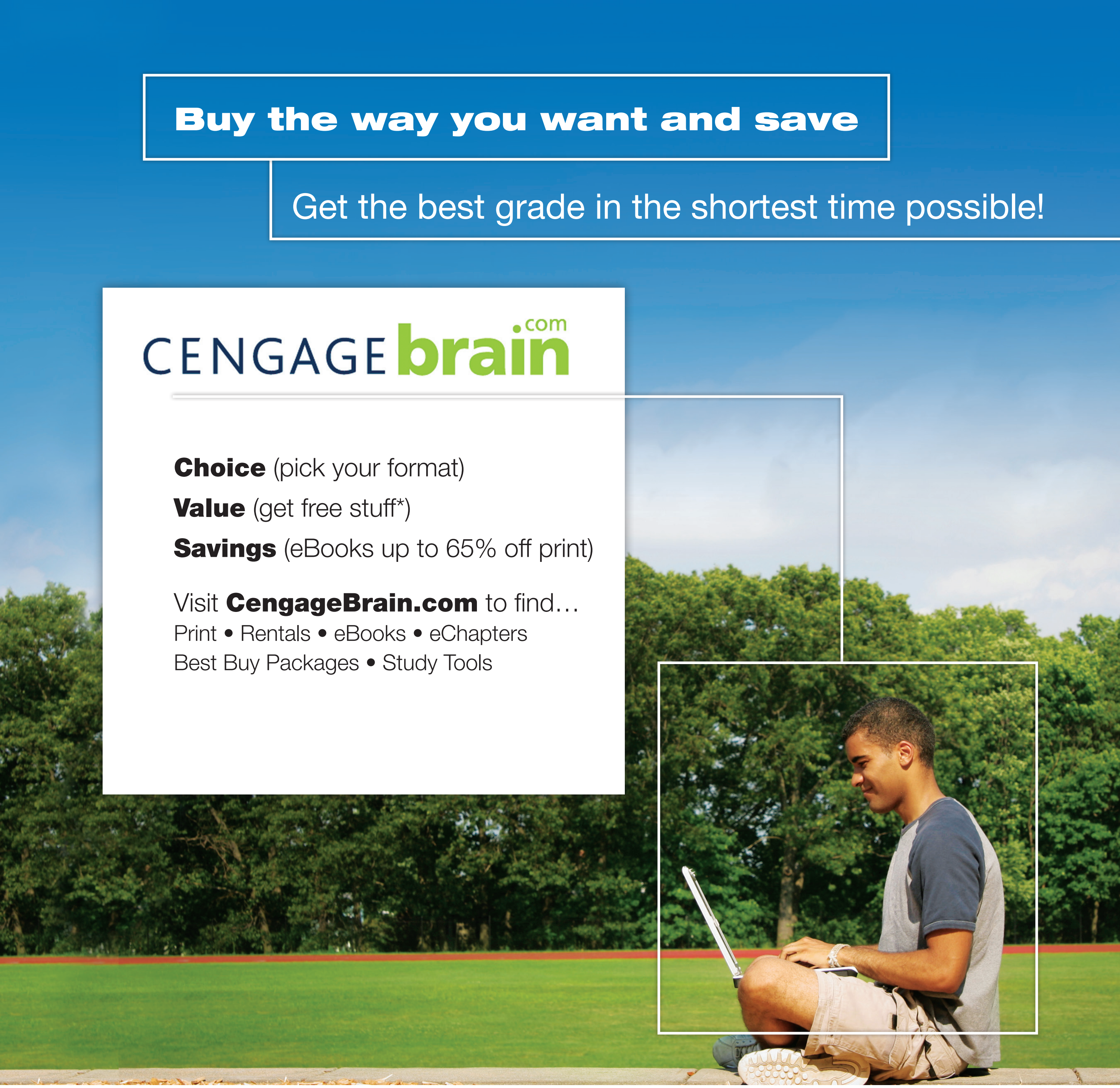
Eighth Edition

ESSENTIALS *of* STATISTICS *for the* BEHAVIORAL SCIENCES

Frederick J Gravetter
Larry B. Wallnau



Essentials of Statistics for the Behavioral Sciences

8th edition

Frederick J Gravetter
State University of New York, Brockport

Larry B. Wallnau
State University of New York, Brockport

WADSWORTH
CENGAGE Learning

Australia • Brazil • Japan • Korea • Mexico • Singapore • Spain • United Kingdom • United States

WADSWORTH
CENGAGE Learning

***Essentials of Statistics for the Behavioral Sciences*, 8th Edition**
Frederick J Gravetter and Larry B. Wallnau

Publisher: Jon-David Hague
Acquisitions Editor: Tim Matray
Developmental Editor: Arwen Petty
Developmental Editor: Liana Sarkisian
Assistant Editor: Paige Leeds
Editorial Assistant: Nicole Richards
Media Editor: Jasmin Tokatlian
Senior Brand Manager: Elisabeth Rhoden
Market Development Manager: Christine Sosa
Content Project Manager: Charlene M. Carpentier
Art Director: Pam Galbreath
Manufacturing Planner: Karen Hunt
Rights Acquisitions Specialist: Don Schlotman
Production and Composition: Graphic World Inc.
Text Researcher: Q2A/Bill Smith
Copy Editor: Graphic World Inc.
Text/Cover Designer: Lisa Delgado
Cover Image: David S. Rose/Shutterstock.com

For product information and technology assistance, contact us at
Cengage Learning Customer & Sales Support, 1-800-354-9706.
For permission to use material from this text or product,
submit all requests online at **www.cengage.com/permissions.**
Further permissions questions can be e-mailed to
permissionrequest@cengage.com.

Library of Congress Control Number: 2012944154

Student Edition:
ISBN-13: 978-1-133-95657-0
ISBN-10: 1-133-95657-2

Loose-leaf Edition:
ISBN-13: 978-1-285-05634-0
ISBN-10: 1-285-05634-5

Wadsworth
20 Davis Drive
Belmont, CA 94002-3098
USA

Cengage Learning is a leading provider of customized learning solutions with office locations around the globe, including Singapore, the United Kingdom, Australia, Mexico, Brazil, and Japan. Locate your local office at **www.cengage.com/global.**

Cengage Learning products are represented in Canada by Nelson Education, Ltd.

To learn more about Wadsworth, visit **www.cengage.com/Wadsworth**

Purchase any of our products at your local college store or at our preferred online store **www.CengageBrain.com.**

Printed in the United States of America
1 2 3 4 5 6 7 16 15 14 13 12

Contents in Brief

PART V Correlations and Nonparametric Tests

Contents

PART V Correlations and Nonparametric Tests

Preface

Many students in the behavioral sciences view the required statistics course as an intimidating obstacle that has been placed in the middle of an otherwise interesting curriculum. They want to learn about human behavior—not about math and science. As a result, the statistics course is seen as irrelevant to their education and career goals. However, as long as the behavioral sciences are founded in science, a knowledge of statistics will be necessary. Statistical procedures provide researchers with objective and systematic methods for describing and interpreting their research results. Scientific research is the system that we use to gather information, and statistics are the tools that we use to distill the information into sensible and justified conclusions. The goal of this book is not only to teach the methods of statistics but also to convey the basic principles of objectivity and logic that are essential for science and valuable for decision making in everyday life.

Those of you who are familiar with previous editions of *Essentials of Statistics for the Behavioral Sciences* will notice that some changes have been made. These changes are summarized in the section titled "To the Instructor." In revising this text, our students have been foremost in our minds. Over the years, they have provided honest and useful feedback. Their hard work and perseverance has made our writing and teaching most rewarding. We sincerely thank them. Students who are using this edition should please read the section of the preface titled "To the Student."

The book chapters are organized in the sequence that we use for our own statistics courses. We begin with descriptive statistics, and then examine a variety of statistical procedures focused on sample means and variance before moving on to correlational methods and nonparametric statistics. Information about modifying this sequence is presented in the "To the Instructor" section for individuals who prefer a different organization. Each chapter contains numerous examples—many based on actual research studies—along with learning checks, a summary and list of key terms, and a set of 20 to 30 problems.

TO THE INSTRUCTOR

Those of you familiar with the previous edition of *Essentials of Statistics for the Behavioral Sciences* will notice a number of changes in the eighth edition. Throughout the book, research examples have been updated, real-world examples have been added, and the end-of-chapter problems have been extensively revised. The book has been separated into five sections to emphasize the similarities among groups of statistical methods. Each section contains two to four chapters and begins with an introduction and concludes with a review, including review exercises. Major revisions for this edition include:

- The former Chapter 12 on estimation has been eliminated. In its place, sections on confidence intervals have been added to the three chapters presenting *t* statistics.
- A new appendix titled *Statistics Organizer: Finding the Right Statistics for Your Data,* discusses the process of selecting the correct statistics to be used with different categories of data and replaces the Statistics Organizer that appeared as an appendix in earlier editions.

Other specific and noteworthy revisions include:

Chapter 1 A separate section explains how statistical methods can be classified using the same categories that are used to group data structures and research methods.

Chapter 2 The discussion of histograms has been modified to differentiate discrete and continuous variables.

Chapter 3 A modified definition of the median acknowledges that this value is not algebraically defined and that determining the median, especially for discrete variables, can be somewhat subjective.

Chapter 4 Relatively minor editing for clarity. The section on variance and inferential statistics has been simplified.

Chapter 5 Relatively minor editing for clarity.

Chapter 6 The concepts of random sample and independent random sample have been clarified with separate definitions. A new figure helps demonstrate the process of using the unit normal table to find proportions for negative *z*-scores.

Chapter 7 Relatively minor editing for clarity.

Chapter 8 The chapter has been shortened by substantial editing that eliminated several pages of unnecessary text, particularly in the sections on errors (Type I and II) and power.

Chapter 9 The section describing how sample size and sample variance influence the outcome of a hypothesis test has been moved so that it appears immediately after the hypothesis test example. A new section introduces confidence intervals in the context of describing effect size, describes how confidence intervals are reported in the literature, and discusses factors affecting the width of a confidence interval.

Chapter 10 An expanded section discusses how sample variance and sample size influence the outcome of an independent-measures hypothesis test and measures of effect size. A new section introduces confidence intervals as an alternative for describing effect size. The relationship between a confidence interval and a hypothesis test is also discussed.

Chapter 11 The description of repeated-measures and matched-subjects designs has been clarified and we increased emphasis on the concept that all calculations for the related-samples test are done with the difference scores. A new section introduces confidence intervals as an alternative for describing effect size and discusses the relationship between a confidence interval and a hypothesis test.

The former Chapter 12 has been deleted. The content from this chapter discussing confidence intervals has been added to Chapters 9, 10, and 11.

Chapter 12 (former Chapter 13, introducing ANOVA) The discussion of testwise alpha levels versus experimentwise alpha levels has been moved from a box into the text, and definitions of the two terms have been added. To emphasize the concepts of ANOVA rather than the formulas, $SS_{\text{between treatments}}$ is routinely found by subtraction instead of being computed directly. Two alternative equations for $SS_{\text{between treatments}}$ have been moved from the text into a box.

Chapter 13 (former Chapter 14, introducing repeated-measures and two-factor ANOVA) A new section demonstrates the relationship between ANOVA and the *t* test when a repeated-measures study is comparing only two treatments. Extensive editing has shortened the chapter and simplified the presentation.

Chapter 14 (formerly Chapter 15, introducing correlation and regression) New sections present the *t* statistic for testing hypotheses about the Pearson correlation and demonstrate how the *t* test for significance of a correlation is equivalent to the *F*-ratio used for analysis of regression.

Chapter 15 (formerly Chapter 16, introducing chi-square tests) Relatively minor editing has shortened and clarified the chapter.

Matching the Text to Your Syllabus We have tried to make separate chapters, and even sections of chapters, completely self-contained so that they can be deleted or reorganized to fit the syllabus for nearly any instructor. Some common examples are as follows:

- It is common for instructors to choose between emphasizing analysis of variance (Chapters 12 and 13) or emphasizing correlation/regression (Chapter 14). It is rare for a one-semester course to provide complete coverage of both topics.
- Although we choose to complete all the hypothesis tests for means and mean differences before introducing correlation (Chapter 14), many instructors prefer to place correlation much earlier in the sequence of course topics. To accommodate this, sections 14.1, 14.2, and 14.3 present the calculation and interpretation of the Pearson correlation and can be introduced immediately following Chapter 4 (variability). Other sections of Chapter 14 refer to hypothesis testing and should be delayed until the process of hypothesis testing (Chapter 8) has been introduced.
- It is also possible for instructors to present the chi-square tests (Chapter 15) much earlier in the sequence of course topics. Chapter 15, which presents hypothesis tests for proportions, can be presented immediately after Chapter 8, which introduces the process of hypothesis testing. If this is done, we also recommend that the Pearson correlation (Sections 14.1, 14.2, and 14.3) be presented early to provide a foundation for the chi-square test for independence.

TO THE STUDENT

A primary goal of this book is to make the task of learning statistics as easy and painless as possible. Among other things, you will notice that the book provides you with a number of opportunities to practice the techniques you will be learning in the form of learning checks, examples, demonstrations, and end-of-chapter problems. We encourage you to take advantage of these opportunities. Read the text rather than just memorize the formulas. We have taken care to present each statistical procedure in a conceptual context that explains why the procedure was developed and when it should be used. If you read this material and gain an understanding of the basic concepts underlying a statistical formula, you will find that learning the formula and how to use it will be much easier. In the following section, “Study Hints,” we provide advice that we give our own students. Ask your instructor for advice as well; we are sure that other instructors will have ideas of their own.

Over the years, the students in our classes and other students using our book have given us valuable feedback. If you have any suggestions or comments about this book, you can write to either Professor Emeritus Frederick Gravetter or Professor Emeritus

Larry Wallnau at the Department of Psychology, SUNY College at Brockport, 350 New Campus Drive, Brockport, New York 14420. You can also contact Professor Emeritus Gravetter directly at fgravett@brockport.edu.

Study Hints You may find some of these tips helpful, as our own students have reported.

- The key to success in a statistics course is to keep up with the material. Each new topic builds on previous topics. If you have learned the previous material, then the new topic is just one small step forward. Without the proper background, however, the new topic can be a complete mystery. If you find that you are falling behind, get help immediately.
- You will learn (and remember) much more if you study for short periods several times per week rather than try to condense all of your studying into one long session. For example, it is far more effective to study half an hour every night than to have a single $3\frac{1}{2}$-hour study session once a week. We cannot even work on *writing* this book without frequent rest breaks.
- Do some work before class. Keep a little ahead of the instructor by reading the appropriate sections before they are presented in class. Although you may not fully understand what you read, you will have a general idea of the topic, which will make the lecture easier to follow. Also, you can identify material that is particularly confusing and then be sure the topic is clarified in class.
- Pay attention and think during class. Although this advice seems obvious, often it is not practiced. Many students spend so much time trying to write down every example presented or every word spoken by the instructor that they do not actually understand and process what is being said. Check with your instructor—there may not be a need to copy every example presented in class, especially if there are many examples like it in the text. Sometimes, we tell our students to put their pens and pencils down for a moment and just listen.
- Test yourself regularly. Do not wait until the end of the chapter or the end of the week to check your knowledge. After each lecture, work some of the end-of-chapter problems and do the Learning Checks. Review the Demonstration Problems, and be sure you can define the Key Terms. If you are having trouble, get your questions answered *immediately* (reread the section, go to your instructor, or ask questions in class). By doing so, you will be able to move ahead to new material.
- Do not kid yourself! Avoid denial. Many students watch their instructor solve problems in class and think to themselves, "This looks easy—I understand it." Do you really understand it? Can you really do the problem on your own without having to leaf through the pages of a chapter? Although there is nothing wrong with using examples in the text as models for solving problems, you should try working a problem with your book closed to test your level of mastery.
- We realize that many students are embarrassed to ask for help. It is our biggest challenge as instructors. You must find a way to overcome this aversion. Perhaps contacting the instructor directly would be a good starting point, if asking questions in class is too anxiety-provoking. You could be pleasantly surprised to find that your instructor does not yell, scold, or bite! Also, your instructor might know of another student who can offer assistance. Peer tutoring can be very helpful.

ANCILLARIES

Ancillaries for this edition include the following:

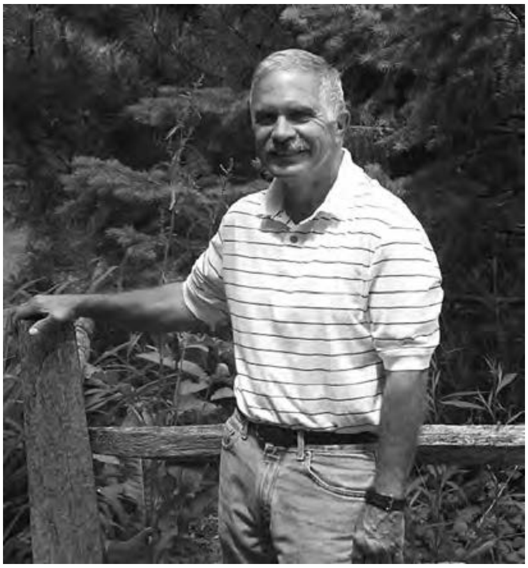
- *Aplia Statistics for Psychology and the Behavioral Sciences:* An online interactive learning solution that ensures students stay involved with their coursework and master the basic tools and concepts of statistical analysis. Created by a research psychologist to help students excel, Aplia's content engages students with questions based on real-world scenarios that help students understand how statistics applies to everyday life. At the same time, all chapter assignments are automatically graded and provide students with detailed explanations, making sure they learn from and improve with every question.
- *Instructor's Manual with Test Bank:* Contains chapter outlines, annotated learning objectives, lecture suggestions, test items, and solutions to all end-of-chapter problems in the text. Test items are also available as a Word download or for ExamView computerized test bank software.
- *PowerLecture with ExamView®:* This CD includes the instructor's manual, test bank, lecture slides with book figures, and more. Featuring automatic grading, ExamView, also available within PowerLecture, allows you to create, deliver, and customize tests and study guides (both print and online) in minutes. Assessments appear onscreen exactly as they will print or display online; you can build tests of up to 250 questions using up to 12 question types, and you can enter an unlimited number of new questions or edit existing questions.
- *WebTutor on Blackboard and WebCT:* Jump-start your course with customizable, text-specific content within your Course Management System.
- *Psychology CourseMate:* Cengage Learning's Psychology CourseMate brings course concepts to life with interactive learning, study, and exam preparation tools that support the printed textbook. Go to www.cengagebrain.com. Psychology CourseMate includes:
- An interactive eBook;
- Interactive teaching and learning tools, including
    - quizzes,
    - flashcards,
    - videos,
    - and more; plus
- The Engagement Tracker, a first-of-its-kind tool that monitors student engagement in the course.

ACKNOWLEDGMENTS

It takes a lot of good, hardworking people to produce a book. Our friends at Wadsworth/Cengage Learning have made enormous contributions to this textbook. We thank: Jon-David Hague, Publisher; Tim Matray, Acquisitions Editor; Paige Leeds, Assistant Editor; Nicole Richards, Editorial Assistant; Charlene M. Carpentier, Content Project Manager; Jasmin Tokatlian, Media Editor; and Pam Galbreath, Art Director. Special thanks go to Liana Sarkisian and Arwen Petty, our Developmental Editors, and to Mike Ederer, who led us through production at Graphic World.

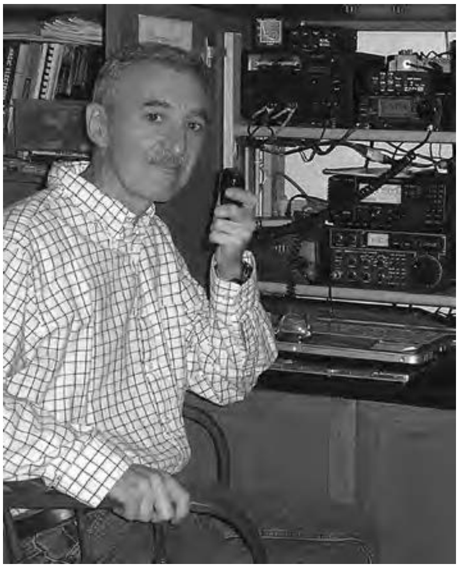
Reviewers play a very important role in the development of a manuscript. Accordingly, we offer our appreciation to the following colleagues for their assistance with the eighth edition: Patricia Tomich, Kent State University; Robert E. Wickham, University of Houston; Jessica Urschel, Western Michigan University; Wilson Chu, California State University, Long Beach; Melissa Platt, University of Oregon; Brian Detweiler-Bedell, Lewis and Clark College.

About the Authors



Frederick J Gravetter is Professor Emeritus of Psychology at the State University of New York College at Brockport. While teaching at Brockport, Dr. Gravetter specialized in statistics, experimental design, and cognitive psychology. He received his bachelor's degree in mathematics from M.I.T. and his Ph.D. in psychology from Duke University. In addition to publishing this textbook and several research articles, Dr. Gravetter co-authored *Research Methods for the Behavioral Sciences* and *Statistics for the Behavioral Sciences.*

Larry B. Wallnau is Professor Emeritus of Psychology at the State University of New York College at Brockport. At Brockport he taught courses relating to the biological basis of behavior and published numerous research articles, primarily in the field of biopsychology. With Dr. Gravetter, he co-authored *Statistics for the Behavioral Sciences.* He also has provided editorial consulting for a number of publishers and journals. He is an FCC-licensed amateur radio operator, and in his spare time he is seeking worldwide contacts with other radio enthusiasts.



PART

I

Introduction and Descriptive Statistics

We have divided this book into five sections, each covering a general topic area of statistics. The first section, consisting of Chapters 1 to 4, provides a broad overview of statistical methods and a more focused presentation of those methods that are classified as *descriptive statistics.*

By the time you finish the four chapters in this part, you should have a good understanding of the general goals of statistics and you should be familiar with the basic terminology and notation used in statistics. In addition, you should be familiar with the techniques of descriptive statistics that help researchers organize and summarize the results they obtain from their research. Specifically, you should be able to take a set of scores and present them in a table or in a graph that provides an overall picture of the complete set. Also, you should be able to summarize a set of scores by calculating one or two values (such as the average) that describe the entire set.

At the end of this section, there is a brief summary and a set of review problems that should help to integrate the elements from the separate chapters.

CHAPTER

1

Introduction to Statistics

aplia

Aplia for *Essentials of Statistics for the Behavioral Sciences*
After reading, go to "Resources" at the end of this chapter for an introduction on how to use Aplia's homework and learning resources.

1.1 STATISTICS, SCIENCE, AND OBSERVATIONS

OVERVIEW

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

The Journey to Enlightenment

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

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

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

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

DEFINITIONS OF STATISTICS

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

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

Research in the behavioral sciences (and other fields) involves gathering information. To determine, for example, whether college students learn better by reading material on printed pages or on a computer screen, you would need to gather information about students' study habits and their academic performance. When researchers finish the task of gathering information, they typically find themselves with pages and pages of measurements such as IQ scores, personality scores, exam scores, and so on. In this book, we present the statistics that researchers use to analyze and interpret the information that they gather. Specifically, statistics serve two general purposes:

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

DEFINITION

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

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

1.2 POPULATIONS AND SAMPLES

Research in the behavioral sciences typically begins with a general question about a specific group (or groups) of individuals. For example, a researcher may want to know what factors are associated with academic dishonesty among college students. Or a researcher may want to examine the amount of time spent in the bathroom for men compared to women. In the first example, the researcher is interested in the group of *college students.* In the second example, the researcher wants to compare the group of *men* with the group of *women.* In statistical terminology, the entire group that a researcher wishes to study is called a *population.*

DEFINITION

A **population** is the entire set of the individuals of interest for a particular research question.

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

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

DEFINITION

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

Just as we saw with populations, samples can vary in size. For example, one study might examine a sample of only 10 autistic children, and another study might use a sample of more than 10,000 people who take a specific cholesterol medication.

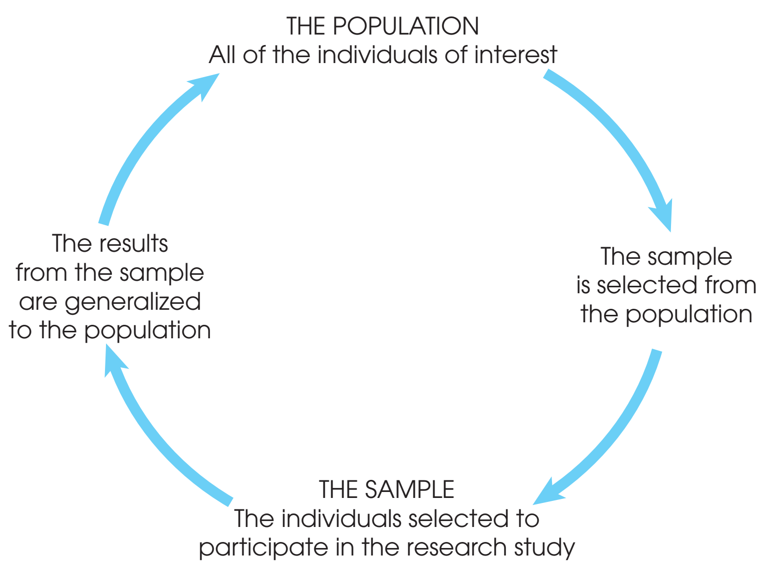
So far we have talked about a sample being selected from a population. However, this is actually only half of the full relationship between a sample and its population. Specifically, when a researcher finishes examining the sample, the goal is to generalize the results back to the entire population.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

VARIABLES AND DATA

Typically, researchers are interested in specific characteristics of the individuals in the population (and in the sample), or they are interested in outside factors that may influence the individuals. For example, a researcher may be interested in the influence

FIGURE 1.1

The relationship between a population and a sample.



of the weather on people's moods. As the weather changes, do people's moods also change? Something that can change or have different values is called a *variable.*

DEFINITION

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

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

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

DEFINITIONS

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

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

PARAMETERS AND STATISTICS

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

DEFINITIONS

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

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

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

DESCRIPTIVE AND INFERENTIAL STATISTICAL METHODS

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

DEFINITION

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

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

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

DEFINITION

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

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

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

DEFINITION

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

The concept of sampling error is illustrated in Figure 1.2. The figure shows a population of 1,000 college students and two samples, each with 5 students, who have been selected from the population. Notice that each sample contains different individuals who have different characteristics. Because the characteristics of each sample depend on the specific people in the sample, statistics vary from one

BOX 1.1 THE MARGIN OF ERROR BETWEEN STATISTICS AND PARAMETERS

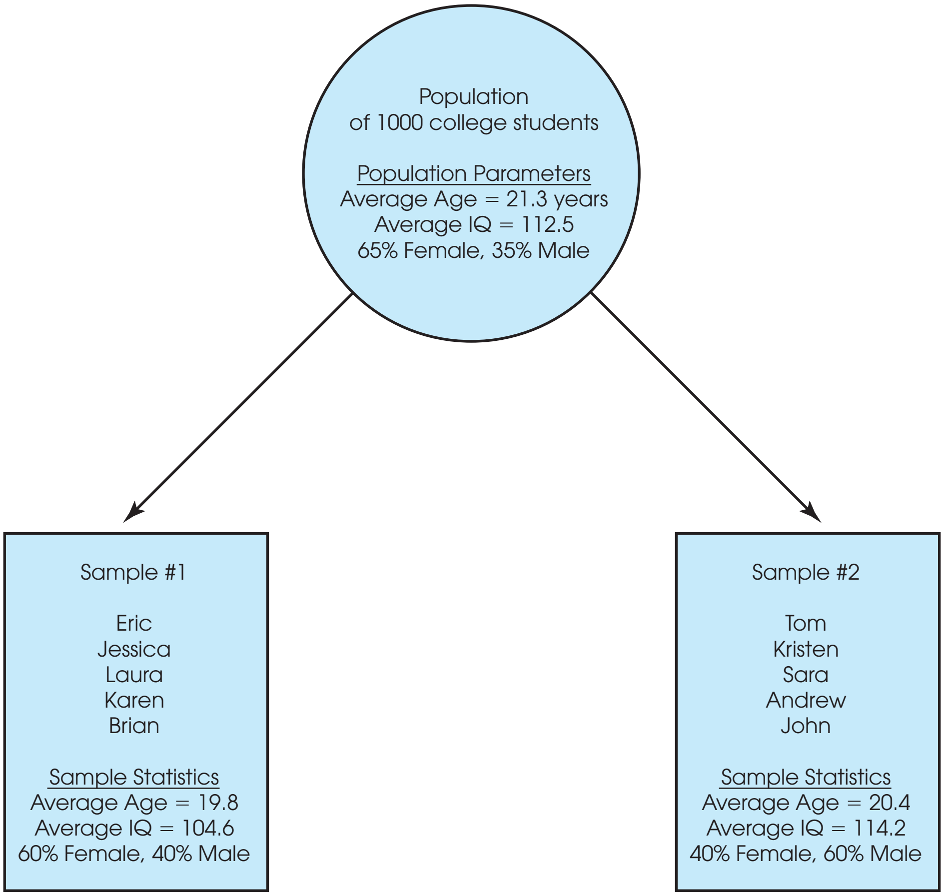
One common example of sampling error is the error associated with a sample proportion. For example, in newspaper articles reporting results from political polls, you frequently find statements such as this:

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

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

FIGURE 1.2

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



sample to another. For example, the five students in sample 1 have an average age of 19.8 years and the students in sample 2 have an average age of 20.4 years.

Also note that the statistics obtained for a sample are not identical to the parameters for the entire population. In Figure 1.2, for example, neither sample has statistics that are exactly the same as the population parameters. You should also realize that Figure 1.2 shows only two of the hundreds of possible samples. Each sample would contain different individuals and would produce different statistics. This is the basic concept of sampling error: sample statistics vary from one sample to another and typically are different from the corresponding population parameters.

As a further demonstration of sampling error, imagine that your statistics class is separated into two groups by drawing a line from front to back through the middle of the room. Now imagine that you compute the average age (or height, or IQ) for each group. Will the two groups have exactly the same average? Almost certainly they will not. No matter what you chose to measure, you will probably find some difference between the two groups. However, the difference you obtain does not necessarily mean that there is a systematic difference between the two groups. For example, if the average age for students on the right-hand side of the room is higher than the average for students on the left, it is unlikely that some mysterious force has caused the older people to gravitate to the right side of the room. Instead, the difference is probably the result of random factors such as chance. The unpredictable, unsystematic differences that exist from one sample to another are an example of sampling error.