

GRAVETTER ■ WALLNAU ■ FORZANO

**ESSENTIALS** OF  
**STATISTICS** FOR THE  
**BEHAVIORAL**  
**SCIENCES** 9<sup>e</sup>

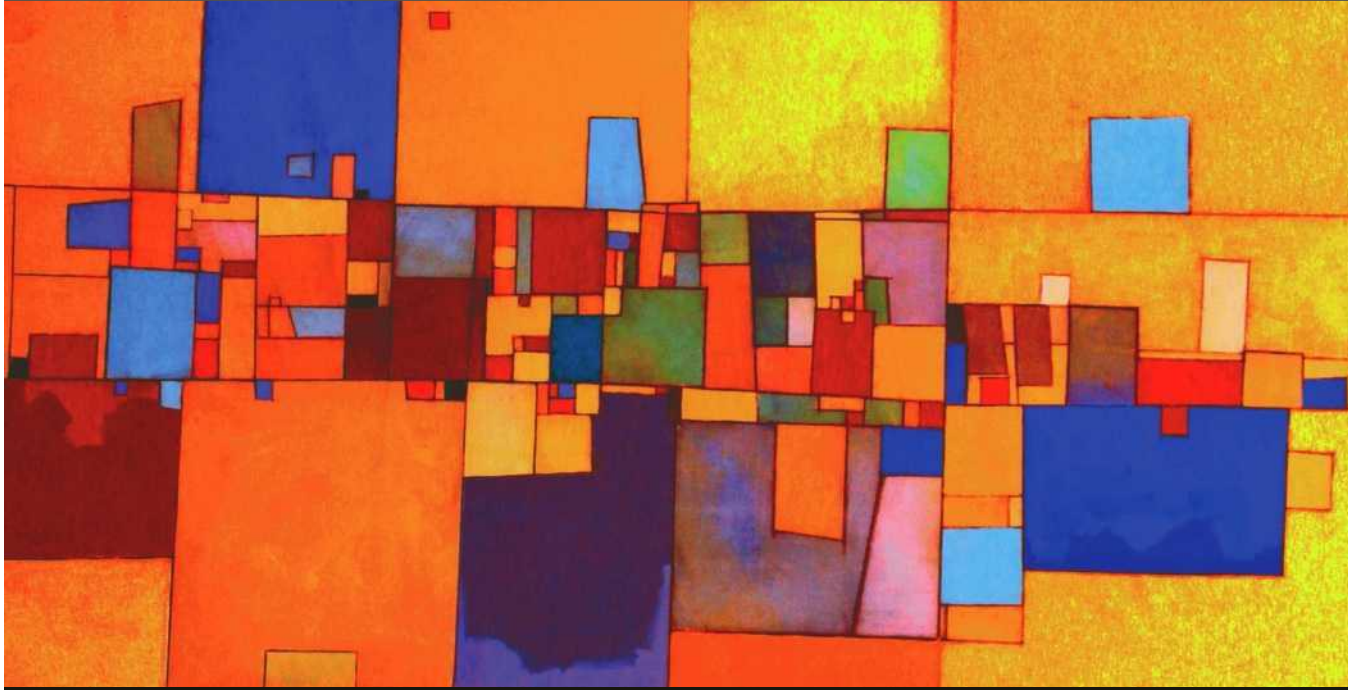




EDITION

9

# Essentials of Statistics FOR THE Behavioral Sciences



**FREDERICK J GRAVETTER**

The College at Brockport, State University of New York

**LARRY B. WALLNAU**

The College at Brockport, State University of New York

**LORI-ANN B. FORZANO**

The College at Brockport, State University of New York



Australia • Brazil • Mexico • Singapore • United Kingdom • United States

**Essentials of Statistics for The Behavioral Sciences, Ninth Edition****Frederick J Gravetter, Larry B. Wallnau,  
Lori-Ann B. Forzano**

Product Director: Marta Lee-Perriard

Product Manager: Carly McJunkin

Content Development Manager:  
Jasmin Tokatlian

Content Developer: Linda Man

Product Assistant: Kimiya Hojjat

Marketing Manager: James Finlay

Content Project Manager: Carol Samet

Art Director: Vernon Boes

Manufacturing Planner: Karen Hunt

Production and Composition Service:  
MPS Limited

Text and Cover Designer: Lisa Henry

Cover Image: Deborah Batt

© 2018, 2014 Cengage Learning

ALL RIGHTS RESERVED. No part of this work covered by the copyright herein may be reproduced or distributed in any form or by any means, except as permitted by U.S. copyright law, without the prior written permission of the copyright owner.

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: 2016941378

Student Edition:

ISBN: 978-1-337-09812-0

Loose-leaf Edition:

ISBN: 978-1-337-27331-2

**Cengage Learning**

20 Channel Center Street

Boston, MA 02210

USA

Cengage Learning is a leading provider of customized learning solutions with employees residing in nearly 40 different countries and sales in more than 125 countries around the world. Find your local representative at **www.cengage.com**

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

To learn more about Cengage Learning Solutions, visit

**www.cengage.com**

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

Printed in Canada

Print Number: 01

Print Year: 2016

# BRIEF CONTENTS

CHAPTER	<b>1</b>	Introduction to Statistics	1
CHAPTER	<b>2</b>	Frequency Distributions	35
CHAPTER	<b>3</b>	Central Tendency	57
CHAPTER	<b>4</b>	Variability	87
CHAPTER	<b>5</b>	z-Scores: Location of Scores and Standardized Distributions	119
CHAPTER	<b>6</b>	Probability	143
CHAPTER	<b>7</b>	Probability and Samples: The Distribution of Sample Means	169
CHAPTER	<b>8</b>	Introduction to Hypothesis Testing	197
CHAPTER	<b>9</b>	Introduction to the $t$ Statistic	237
CHAPTER	<b>10</b>	The $t$ Test for Two Independent Samples	267
CHAPTER	<b>11</b>	The $t$ Test for Two Related Samples	301
CHAPTER	<b>12</b>	Introduction to Analysis of Variance	329
CHAPTER	<b>13</b>	Repeated-Measures and Two-Factor Analysis of Variance	371
CHAPTER	<b>14</b>	Correlation and Regression	421
CHAPTER	<b>15</b>	The Chi-Square Statistic: Tests for Goodness of Fit and Independence	473



## CHAPTER 1 Introduction to Statistics 1

---



- 1-1 Statistics, Science, and Observations 2
- 1-2 Variables and Measurement 10
- 1-3 Three Data Structures, Research Methods, and Statistics 17
- 1-4 Statistical Notation 25
- Summary 29
- Focus on Problem Solving 31
- Demonstration 1.1 31
- Problems 32

## CHAPTER 2 Frequency Distributions 35

---



- 2-1 Frequency Distributions and Frequency Distribution Tables 36
- 2-2 Grouped Frequency Distribution Tables 39
- 2-3 Frequency Distribution Graphs 43
- Summary 50
- Focus on Problem Solving 52
- Demonstration 2.1 52
- Problems 53

## CHAPTER 3 Central Tendency 57

---



- 3-1 Overview 58
- 3-2 The Mean 59
- 3-3 The Median 67
- 3-4 The Mode 71
- 3-5 Central Tendency and the Shape of the Distribution 74
- 3-6 Selecting a Measure of Central Tendency 76
- Summary 82
- Focus on Problem Solving 83
- Demonstration 3.1 84
- Problems 84

**CHAPTER 4** **Variability** **87**

---



- 4-1 Introduction to Variability 88
- 4-2 Defining Variance and Standard Deviation 91
- 4-3 Measuring Variance and Standard Deviation for a Population 95
- 4-4 Measuring Variance and Standard Deviation for a Sample 99
- 4-5 Sample Variance as an Unbiased Statistic 104
- 4-6 More about Variance and Standard Deviation 107
- Summary 113
- Focus on Problem Solving 115
- Demonstration 4.1 115
- Problems 116

**CHAPTER 5** **z-Scores: Location of Scores and Standardized Distributions** **119**

---



- 5-1 Introduction 120
- 5-2 z-Scores and Locations in a Distribution 121
- 5-3 Other Relationships between z, X, the Mean, and the Standard Deviation 125
- 5-4 Using z-Scores to Standardize a Distribution 128
- 5-5 Other Standardized Distributions Based on z-Scores 133
- 5-6 Looking Ahead to Inferential Statistics 135
- Summary 138
- Focus on Problem Solving 139
- Demonstration 5.1 140
- Demonstration 5.2 140
- Problems 140

**CHAPTER 6** **Probability** **143**

---



- 6-1 Introduction to Probability 144
- 6-2 Probability and the Normal Distribution 149
- 6-3 Probabilities and Proportions for Scores from a Normal Distribution 156

6-4 Looking Ahead to Inferential Statistics 163

Summary 165

Focus on Problem Solving 166

Demonstration 6.1 166

Problems 167

## CHAPTER 7

### Probability and Samples: The Distribution of Sample Means

169



7-1 Samples, Populations, and the Distribution of Sample Means 170

7-2 Shape, Central Tendency, and Variability for the Distribution of Sample Means 175

7-3 z-Scores and Probability for Sample Means 181

7-4 More About Standard Error 185

7-5 Looking Ahead to Inferential Statistics 190

Summary 193

Focus on Problem Solving 194

Demonstration 7.1 194

Problems 195

## CHAPTER 8

### Introduction to Hypothesis Testing

197



8-1 The Logic of Hypothesis Testing 198

8-2 Uncertainty and Errors in Hypothesis Testing 209

8-3 More about Hypothesis Tests 213

8-4 Directional (One-Tailed) Hypothesis Tests 218

8-5 Concerns about Hypothesis Testing: Measuring Effect Size 222

8-6 Statistical Power 226

Summary 231

Focus on Problem Solving 232

Demonstration 8.1 233

Demonstration 8.2 234

Problems 234



**CHAPTER 9 Introduction to the  $t$  Statistic 237**

---



- 9-1 The  $t$  Statistic: An Alternative to  $z$  238
- 9-2 Hypothesis Tests with the  $t$  Statistic 244
- 9-3 Measuring Effect Size for the  $t$  Statistic 248
- 9-4 Directional Hypotheses and One-Tailed Tests 257
- Summary 260
- Focus on Problem Solving 262
- Demonstration 9.1 262
- Demonstration 9.2 263
- Problems 263

**CHAPTER 10 The  $t$  Test for Two Independent Samples 267**

---



- 10-1 Introduction to the Independent-Measures Design 268
- 10-2 The Hypotheses and the Independent-Measures  $t$  Statistic 270
- 10-3 Hypothesis Tests with the Independent-Measures  $t$  Statistic 277
- 10-4 Effect Size and Confidence Intervals for the Independent-Measures  $t$  284
- 10-5 The Role of Sample Variance and Sample Size in the Independent-Measures  $t$  Test 288
- Summary 291
- Focus on Problem Solving 293
- Demonstration 10.1 294
- Demonstration 10.2 295
- Problems 295

**CHAPTER 11 The  $t$  Test for Two Related Samples 301**

---



- 11-1 Introduction to Repeated-Measures Designs 302
- 11-2 The  $t$  Statistic for a Repeated-Measures Research Design 303
- 11-3 Hypothesis Tests for the Repeated-Measures Design 307
- 11-4 Effect Size, Confidence Intervals, and the Role of Sample Size and Sample Variance for the Repeated-Measures  $t$  310
- 11-5 Comparing Repeated- and Independent-Measures Designs 316
- Summary 320
- Focus on Problem Solving 322
- Demonstration 11.1 323

Demonstration 11.2 324

Problems 324

## CHAPTER 12 Introduction to Analysis of Variance 329



12-1 Introduction (An Overview of Analysis of Variance) 330

12-2 The Logic of Analysis of Variance 334

12-3 ANOVA Notation and Formulas 338

12-4 Examples of Hypothesis Testing and Effect Size with ANOVA 346

12-5 Post Hoc Tests 353

12-6 More about ANOVA 357

Summary 362

Focus on Problem Solving 364

Demonstration 12.1 365

Demonstration 12.2 366

Problems 367

## CHAPTER 13 Repeated-Measures and Two-Factor Analysis of Variance 371



13-1 Introduction to the Repeated-Measures ANOVA 372

13-2 Hypothesis Testing and Effect Size with the Repeated-Measures ANOVA 375

13-3 More about the Repeated-Measures Design 384

13-4 An Overview of the Two-Factor, Independent-Measures ANOVA 388

13-5 An Example of the Two-Factor ANOVA and Effect Size 396

Summary 406

Focus on Problem Solving 410

Demonstration 13.1 411

Demonstration 13.2 413

Problems 415

## CHAPTER 14 Correlation and Regression 421



14-1 Introduction 422

14-2 The Pearson Correlation 425

14-3 Using and Interpreting the Pearson Correlation 430

14-4 Hypothesis Tests with the Pearson Correlation 437

14-5 Alternatives to the Pearson Correlation 441  
 14-6 Introduction to Linear Equations and Regression 450  
 Summary 463  
 Focus on Problem Solving 468  
 Demonstration 14.1 468  
 Problems 470

CHAPTER **15**

**The Chi-Square Statistic: Tests for Goodness of Fit and Independence**

**473**



15-1 Introduction to Chi-Square: The Test for Goodness of Fit 474  
 15-2 An Example of the Chi-Square Test for Goodness of Fit 479  
 15-3 The Chi-Square Test for Independence 485  
 15-4 Effect Size and Assumptions for the Chi-Square Tests 493  
 15-5 The Relationship between Chi-Square and Other Statistical Procedures 498  
 Summary 500  
 Focus on Problem Solving 502  
 Demonstration 15.1 504  
 Demonstration 15.2 505  
 Problems 506

**APPENDIXES**

**A** Basic Mathematics Review 511  
     A-1 Symbols and Notation 513  
     A-2 Proportions: Fractions, Decimals, and Percentages 515  
     A-3 Negative Numbers 521  
     A-4 Basic Algebra: Solving Equations 523  
     A-5 Exponents and Square Roots 526  
**B** Statistical Tables 533  
**C** Solution for Odd-Numbered Problems in the Text 545  
**D** General Instructions for Using SPSS 559  
 Statistics Organizer: Finding the Right Statistics for Your Data 563  
 References 575  
 Name Index 579  
 Subject Index 581

**M**any 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, 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.

*Essentials of Statistics for the Behavioral Sciences*, ninth edition, is intended for an undergraduate statistics course in psychology or any of the behavioral sciences. The overall learning objectives of this book include the following, which correspond to some of the learning goals identified by the American Psychological Association (Noland and the Society for the Teaching of Psychology Statistical Literacy Taskforce, 2012):

1. Calculate and interpret the meaning of basic measures of central tendency and variability.
2. Distinguish between causal and correlational relationships.
3. Interpret data displayed as statistics, graphs, and tables.
4. Select and implement an appropriate statistical analysis for a given research design, problem, or hypothesis.
5. Identify the correct strategy for data analysis and interpretation when testing hypotheses.
6. Select, apply, and interpret appropriate descriptive and inferential statistics.
7. Produce and interpret reports of statistical analyses using APA style.
8. Distinguish between statistically significant and chance findings in data.
9. Calculate and interpret the meaning of basic tests of statistical significance.
10. Calculate and interpret the meaning of confidence intervals.
11. Calculate and interpret the meaning of basic measures of effect size statistics.
12. Recognize when a statistically significant result may also have practical significance.

The book chapters are organized in the sequence that we use for our own statistics courses. We begin with descriptive statistics (Chapters 1–4), next lay the foundation for inferential statistics (Chapters 5–8), and then examine a variety of statistical procedures focused on sample means and variance (Chapters 9–13), before moving on to correlational methods and nonparametric statistics (Chapters 14 and 15). 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), learning objectives and learning checks for each section, a summary and list of key terms, instructions for using SPSS, detailed problem-solving tips and demonstrations, and a set of end-of-chapter problems.

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 “To the Instructor” section. Students who are using this edition should read the section of the preface entitled “To the Student.” In revising this text, our students have been foremost in our minds. Over the years, they have provided honest and useful feedback, and their hard work and perseverance has made our writing and teaching most rewarding. We sincerely thank them.

## 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 ninth edition. Throughout the book, research examples have been updated, real-world examples of particular interest to students have been added, and the end-of-chapter problems have been extensively revised.

Major revisions for this edition include:

1. Each section of every chapter begins with a list of Learning Objectives for that specific section.
2. Each section ends with a Learning Check consisting of multiple-choice questions and answers, with at least one question for each Learning Objective.
3. Do-it-yourself examples have been added to each chapter. These present numerical examples and ask student to perform specific statistical calculations to test their understanding of topics presented in the chapter.

Other examples of specific and noteworthy revisions include:

**Chapter 1** The section on data structures and research methods has been edited to parallel the *Statistics Organizer: Finding the Right Statistics for Your Data* in the appendix. The chapter has been reorganized to simplify the sequence of topics.

**Chapter 2** The chapter has undergone relatively minor editing to clarify and simplify.

**Chapter 3** The sequence of topics within the chapter has been reorganized to facilitate the flow of concepts. The median discussion has been refined and clarified and includes both samples and populations.

**Chapter 4** The opening paragraphs have been edited to relate the concept of variability to the more familiar concepts of diversity and consistency. The sections on standard deviation and variance have been edited to increase emphasis on concepts rather than calculations.

**Chapter 5** The section discussing  $z$ -scores for samples has been incorporated into the other sections of the chapter so that populations and samples are consistently discussed together.

**Chapter 6** The section “Looking Ahead to Inferential Statistics” has been shortened and simplified.

**Chapter 7** The box feature explaining the difference between standard deviation and standard error has been deleted, with the content being incorporated into Section 7.4. Furthermore, the content was edited to emphasize that the standard error is the primary new element introduced in the chapter. The final section, “Looking Ahead to Inferential Statistics,” was simplified and shortened to be consistent with the changes in Chapter 6.

**Chapter 8** A redundant example was deleted, which shortened and streamlined the remaining material so that most of the chapter is focused on the same research example.

**Chapter 9** The section introducing confidence intervals was edited to clarify the origin of the confidence interval equation and to emphasize that the interval is constructed at the sample mean.

**Chapter 10** The section presenting the estimated standard error of  $(M_1 - M_2)$  has been simplified and shortened.

**Chapter 11** The discussion of matched-subjects designs has been moved to the end of the chapter as part of the strengths and weaknesses of independent- versus repeated-measures designs. The section discussing hypothesis testing has been separated from the section on effect size and confidence intervals to be consistent with the other two chapters on  $t$  tests. The section comparing independent- and repeated-measures designs has been expanded.

**Chapter 12** One redundant example of an ANOVA has been eliminated to simplify and shorten the chapter. Sections of the chapter have been reorganized to allow the discussion to flow directly from hypothesis tests and effect size to post tests.

**Chapter 13** The section discussing factors that influence the outcome of a repeated-measures hypothesis test and associated measures of effect size has been substantially expanded. The same research example is used to introduce and to demonstrate the two-factor ANOVA, which simplifies that portion of the chapter.

**Chapter 14** The section on partial correlations was deemed not essential to the topic of correlation and has been eliminated, which substantially simplifies and shortens the chapter.

**Chapter 15** A new section introduces Cohen’s  $w$  as a measure of effect size for both chi-square tests.

### ■ Matching the Text to Your Syllabus

The book chapters are organized in the sequence that we use for our own statistics courses. However, different instructors may prefer different organizations and probably will choose to omit or deemphasize specific topics. 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 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 memorizing 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 “Study Hints” that follow, 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.

### ■ 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. Distributed practice is best for learning. For example, it is far more effective to study and do problems for half an hour every night than to have a single three-and-a-half-hour study session once a week. We cannot even work on *writing* this book without frequent rest breaks.
- Do some work before class. Stay a little bit 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. As you are reading the textbook, stop and do the examples. Also, stop and do the Learning Checks at the end of each section. After each lecture, work some of the end-of- chapter problems. 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 observe their instructor solving 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.

### ■ Contact Us

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, Professor Emeritus Larry Wallnau, or Professor Lori-Ann Forzano at the Department of Psychology, The College at Brockport, SUNY, 350 New Campus Drive, Brockport, New York 14420. You can also contact Professor Emeritus Gravetter directly at [fgravett@brockport.edu](mailto:fgravett@brockport.edu).

## Ancillaries

---

Ancillaries for this edition include the following.

- **MindTap® Psychology** *MindTap® Psychology for Gravetter/Wallnau/Forzano’s Essentials of Statistics for the Behavioral Sciences, Ninth Edition* is the digital learning solution that helps instructors engage and transform today’s students into critical thinkers. Through paths of dynamic assignments and applications that you can personalize, real-time course analytics, and an accessible reader, MindTap helps you turn cookie cutter into cutting edge, apathy into engagement, and memorizers into higher-level thinkers. As an instructor using MindTap you have at your fingertips the right content and unique set of tools curated specifically for your course, such as video tutorials that walk students through various concepts and interactive problem tutorials that provide students opportunities to practice what they have learned, all in an interface designed to improve workflow and save time when planning lessons and course structure. The control to build and personalize your course is all yours, focusing on the most relevant material while also lowering costs for your students. Stay connected and informed in your course through real time student tracking that



provides the opportunity to adjust the course as needed based on analytics of interactivity in the course.

- **Aplia** 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.
- **Online Instructor’s Manual** The manual includes learning objectives, key terms, a detailed chapter outline, a chapter summary, lesson plans, discussion topics, student activities, “What If” scenarios, media tools, a sample syllabus, and an expanded test bank. The learning objectives are correlated with the discussion topics, student activities, and media tools.
- **Online PowerPoints** Helping you make your lectures more engaging while effectively reaching your visually oriented students, these handy Microsoft PowerPoint® slides outline the chapters of the main text in a classroom-ready presentation. The PowerPoint® slides are updated to reflect the content and organization of the new edition of the text.
- **Cengage Learning Testing, powered by Cognero®** Cengage Learning Testing, powered by Cognero®, is a flexible online system that allows you to author, edit, and manage test bank content. You can create multiple test versions in an instant and deliver tests from your LMS in your classroom.

## Acknowledgments

---

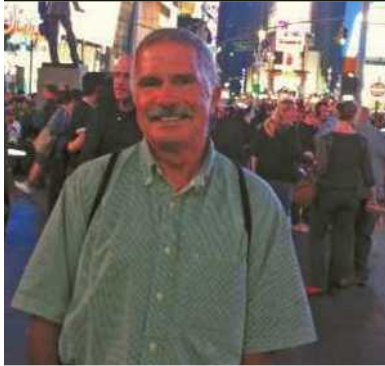
It takes a lot of good, hard-working people to produce a book. Our friends at Cengage have made enormous contributions to this textbook. We thank: Marta Lee-Perriard, Product Director; Carly McJunkin, Product Manager; Kimiya Hojjat, Product Assistant; Carol Samet, Senior Content Project Manager; and Vernon Boes, Art Director. Special thanks go to Linda Man, our Content Developer, and to Lynn Lustberg who led us through production at MPS.

Reviewers play an important role in the development of a manuscript. Accordingly, we offer our appreciation to the following colleagues for their assistance: 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.

Finally, our heartfelt thanks to our spouses and children: Debbie Gravetter; Charlie, Ryan, and Alex Forzano; and Naomi and Nico Wallnau. This book could not have been written without their patience and support.

*Frederick J Gravetter  
Larry B. Wallnau  
Lori-Ann B. Forzano*

# ABOUT THE AUTHORS



**FREDERICK J. GRAVETTER** is Professor Emeritus of Psychology at The College at Brockport, State University of New York. 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 coauthored *Research Methods for the Behavioral Sciences* and *Statistics for the Behavioral Sciences*.



**LARRY B. WALLNAU** is Professor Emeritus of Psychology at The College at Brockport, State University of New York. While teaching at Brockport, he published numerous research articles in biopsychology and has provided editorial consultation for journals. With Dr. Gravetter, he coauthored *Statistics for the Behavioral Sciences*. Dr. Wallnau received his bachelor's degree from the University of New Haven and his Ph.D. in psychology from the State University of New York at Albany. He has taken up running and has competed in races in New York and New England. He takes great pleasure in adopting neglected and rescued dogs.



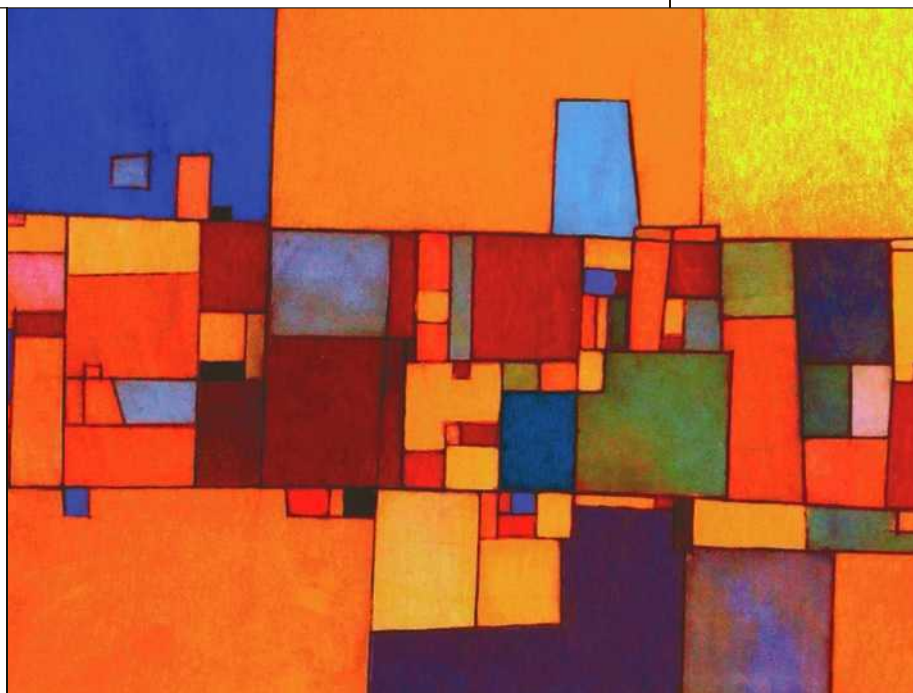
**LORI-ANN B. FORZANO** is Professor of Psychology at The College at Brockport, State University of New York, where she regularly teaches undergraduate and graduate courses in research methods, learning, the psychology of eating, and statistics. She earned a Ph.D. in experimental psychology from the State University of New York at Stony Brook, where she also received her B.S. in psychology. Dr. Forzano's research examines impulsivity and self-control in adults and children. Her research has been published in the *Journal of the Experimental Analysis of Behavior, Learning and Motivation*, and *The Psychological Record*. Dr. Forzano has also coauthored *Research Methods for the Behavioral Sciences*.



# Introduction to Statistics

CHAPTER

1



**1-1** Statistics, Science, and Observations

**1-2** Variables and Measurement

**1-3** Three Data Structures, Research Methods, and Statistics

**1-4** Statistical Notation

Summary

Focus on Problem Solving

Demonstration 1.1

Problems



## 1-1 Statistics, Science, and Observations

### LEARNING OBJECTIVES

1. Define the terms population, sample, parameter, and statistic; describe the relationships between them; and identify examples of each.
2. Define the two general categories of statistics, descriptive and inferential, and describe how they are used in a typical research study.
3. Describe the concept of sampling error and explain how sampling error creates the fundamental problem that inferential statistics must address.

### ■ Introduction

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? In part, we simply wanted to lighten the mood with a bit of humor—starting a statistics course is typically not viewed as one of life’s joyous moments. In addition, the paragraph is an excellent counter-example for the purpose of this book. Specifically, our goal is to do everything possible to prevent you from stumbling around in the dark by providing lots of help and illumination as you journey through the world of statistics. To accomplish this, we begin each section of the book with clearly stated learning objectives and end each section with a brief quiz to test your mastery of the new material. We also introduce each new statistical procedure by explaining the purpose it is intended to serve. If you understand why a new procedure is needed, you will find it much easier to learn.

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

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

### ■ Definitions of Statistics

By one definition, *statistics* consist of facts and figures such as the average annual snowfall in Denver or Derek Jeter’s lifetime batting average. These statistics are usually informative

\*Candappa, R. (2000). *The Little Book of Wrong Shui*. Kansas City, MO: Andrews McMeel Publishing. Reprinted by permission.

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 general field of mathematics. 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 to summarize and evaluate research results in the behavioral sciences.

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 test scores, personality scores, and opinions. 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.

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

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

### ■ 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*.

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

As you can well imagine, a population can be quite large—for example, the entire set of women on the planet Earth. A researcher might be more specific, limiting the population for study to women who are registered voters in the United States. Perhaps the investigator would like to study the population consisting of women who are U.S. senators. Populations can obviously vary in size from extremely large to very small, depending on how the

investigator 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 an investigator 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.

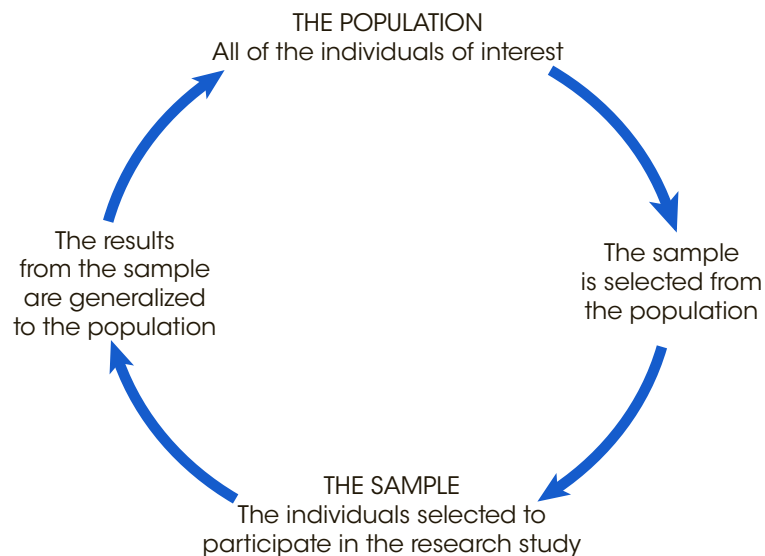
A **sample** is a set of individuals selected from a population, usually intended to represent the population in a research study.

Just as we saw with populations, samples can vary in size. For example, one study might examine a sample of only 10 students in a graduate program and another study might use a sample of more than 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 (or in the sample), or they are interested in the factors that may influence individuals



**FIGURE 1.1**  
The relationship between a population and a sample.

or their behaviors. For example, a researcher may be interested in the influence of television commercials on people's fast-food preferences. As new commercials appear on TV, do people's food choices also change? Something that can change or have different values is called a *variable*.

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 weight, gender, personality, or fast-food preferences. Also, variables can be environmental conditions that change such as temperature, time of day, or television commercials.

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*.

**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 cholesterol patients. 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 specify 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.

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 (Figure 1.1). 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.

**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.

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

Because populations are typically very large, it usually is not possible to measure everyone in the population. Therefore, a sample is selected to represent the population. By analyzing the results from the sample, we hope to make general statements about the population. Typically, researchers use sample statistics as the basis for drawing conclusions about population parameters. One problem with using samples, however, is that a sample provides only limited information about the population. Although samples are generally *representative* of their populations, a sample is not expected to give a perfectly accurate picture of the whole population. Thus, there usually is some discrepancy between a sample statistic and the corresponding population parameter. This discrepancy is called *sampling error*, and it creates the fundamental problem that inferential statistics must address.

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

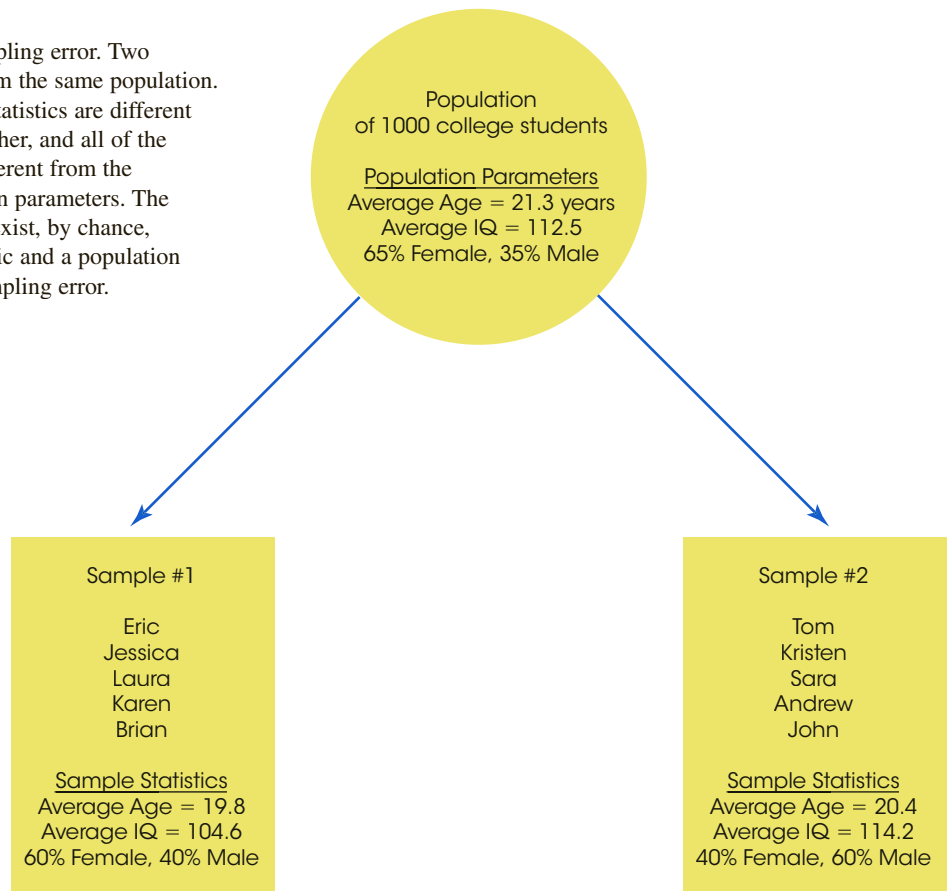
The concept of sampling error is illustrated in Figure 1.2. The figure shows a population of 1,000 college students and two samples, each with five students who were selected from the population. Notice that each sample contains different individuals who have different characteristics. Because the characteristics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

It is also very unlikely that the statistics obtained for a sample will be 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.



**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.



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.

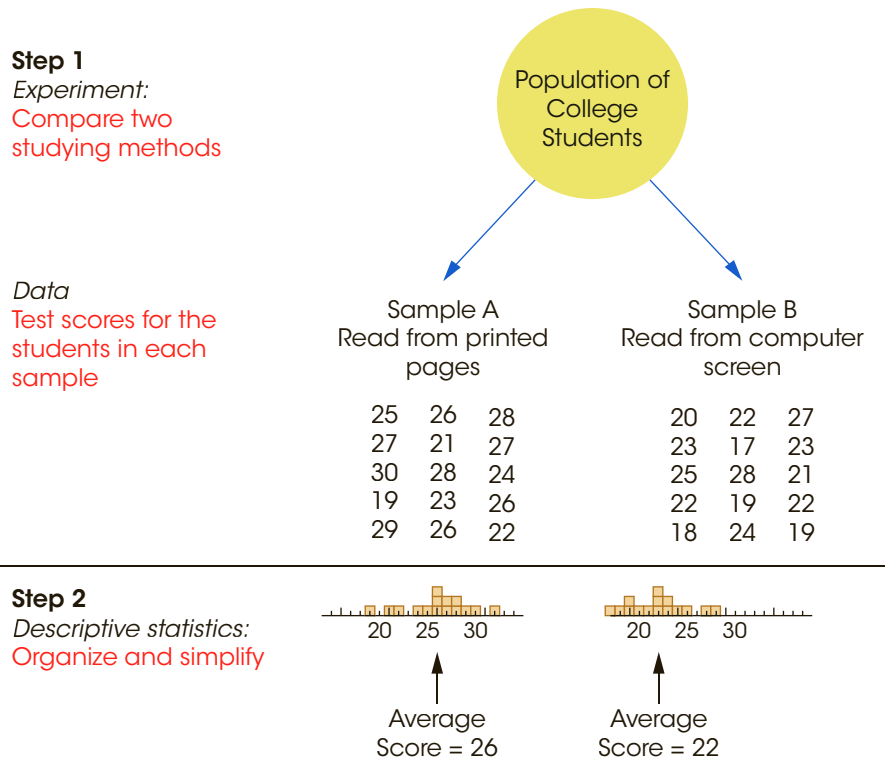
The “margin of error” is the sampling error. In this case, the percentages that are reported were obtained from a sample and are being generalized to the whole population. As always, you do not expect the statistics from a sample to be perfect. There always will be some “margin of error” when sample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

### ■ Statistics in the Context of Research

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

**EXAMPLE 1.1**

Figure 1.3 shows an overview of a general research situation and demonstrates the roles that descriptive and inferential statistics play. The purpose of the research study is to address a question that we posed earlier: Do college students learn better by studying text on printed pages or on a computer screen? Two samples are selected from the population of college students. The students in sample A are given printed pages of text to study for 30 minutes and the students in sample B study the same text on a computer screen. Next, all of the students are given a multiple-choice test to evaluate their knowledge of the material. At this point, the researcher has two sets of data: the scores for sample A and the scores for sample B (see the figure). Now is the time to begin using statistics.



**Step 1**  
*Experiment:*  
 Compare two studying methods

*Data*  
 Test scores for the students in each sample

**Step 2**  
*Descriptive statistics:*  
 Organize and simplify

**Step 3**  
*Inferential statistics:*  
 Interpret results

The sample data show a 4-point difference between the two methods of studying. However, there are two ways to interpret the results.

1. There actually is no difference between the two studying methods, and the sample difference is due to chance (sampling error).
2. There really is a difference between the two methods, and the sample data accurately reflect this difference.

The goal of inferential statistics is to help researchers decide between the two interpretations.

**FIGURE 1.3**  
 The role of statistics in research.

First, descriptive statistics are used to simplify the pages of data. For example, the researcher could draw a graph showing the scores for each sample or compute the average score for each sample. Note that descriptive methods provide a simplified, organized description of the scores. In this example, the students who studied printed pages had an average score of 26 on the test, and the students who studied text on the computer averaged 22.

Once the researcher has described the results, the next step is to interpret the outcome. This is the role of inferential statistics. In this example, the researcher has found a difference of 4 points between the two samples (sample A averaged 26 and sample B averaged 22). The problem for inferential statistics is to differentiate between the following two interpretations:

1. There is no real difference between the printed page and a computer screen, and the 4-point difference between the samples is just an example of sampling error (like the samples in Figure 1.2).
2. There really is a difference between the printed page and a computer screen, and the 4-point difference between the samples was caused by the different methods of studying.

In simple English, does the 4-point difference between samples provide convincing evidence of a difference between the two studying methods, or is the 4-point difference just chance? The purpose of inferential statistics is to answer this question. ■

### LEARNING CHECK

- LO1 1.** A researcher is interested in the fast-food eating habits of American college students. A group of 50 students is interviewed and the researcher finds that these students eat an average of 2.3 commercially prepared meals per week. For this study, the average of 2.3 meals is an example of a \_\_\_\_\_.
- a. parameter
  - b. statistic
  - c. population
  - d. sample
- LO1 2.** A researcher is curious about the average distance traveled by Canada geese during peak fall migration in the state of New York. The entire group of Canada geese in the state is an example of a \_\_\_\_\_.
- a. sample
  - b. statistic
  - c. population
  - d. parameter
- LO2 3.** What term is used for the statistical techniques that use sample data to draw conclusions about the population from which the sample was obtained?
- a. population statistics
  - b. sample statistics
  - c. descriptive statistics
  - d. inferential statistics