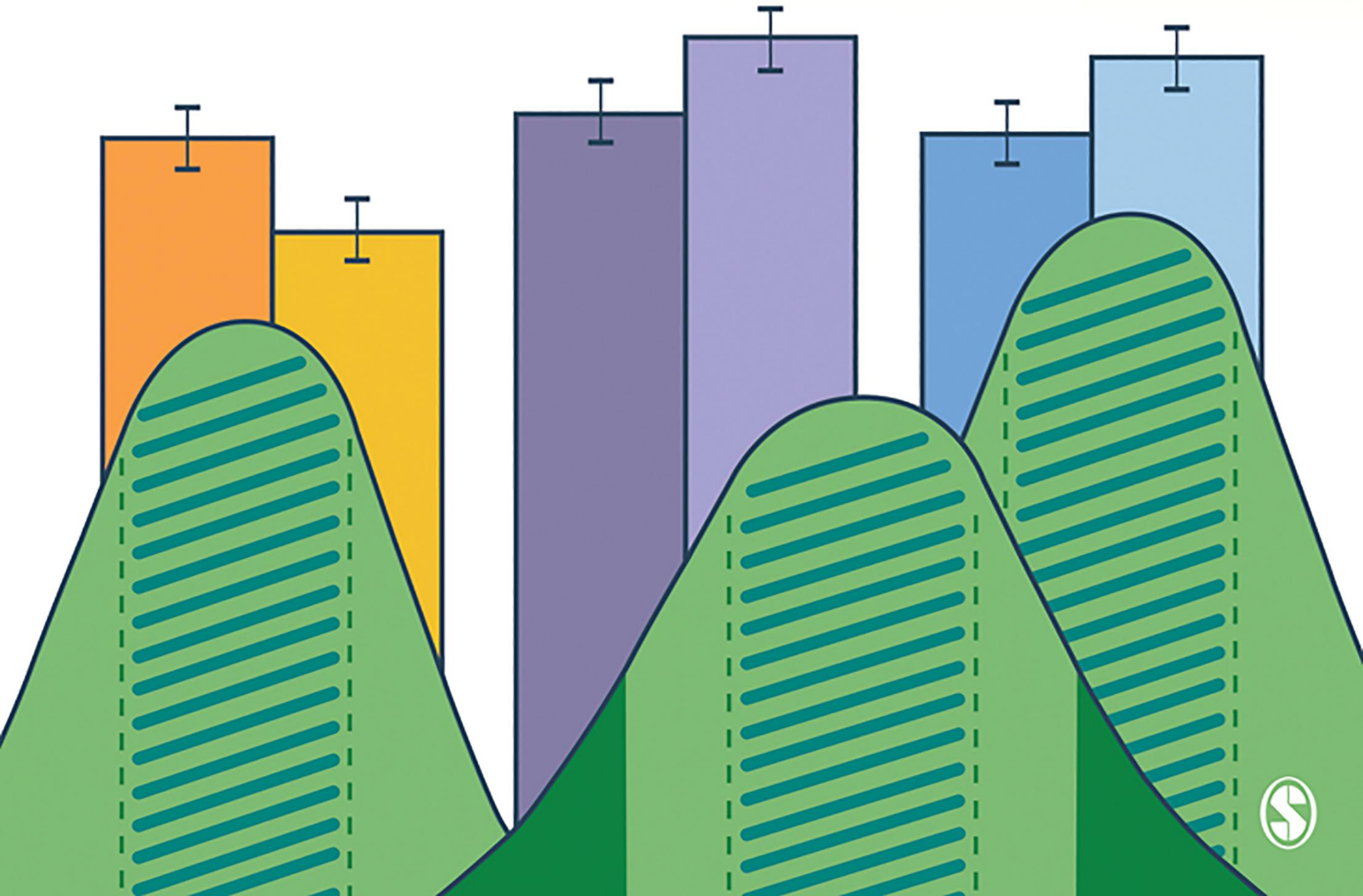
3e

Kieth A. Carlson · Jennifer R. Winquist

# AN INTRODUCTION TO STATISTICS

## An Active Learning Approach



# An Introduction to Statistics

## Third Edition

Sara Miller McCune founded SAGE Publishing in 1965 to support the dissemination of usable knowledge and educate a global community. SAGE publishes more than 1000 journals and over 800 new books each year, spanning a wide range of subject areas. Our growing selection of library products includes archives, data, case studies and video. SAGE remains majority owned by our founder and after her lifetime will become owned by a charitable trust that secures the company's continued independence.

Los Angeles | London | New Delhi | Singapore | Washington DC | Melbourne

# An Introduction to Statistics

## An Active Learning Approach

### Third Edition

Kieth A. Carlson

*Valparaiso University*

Jennifer R. Winquist

*Valparaiso University*

SAGE

Los Angeles | London | New Delhi
Singapore | Washington DC | Melbourne

SAGE

FOR INFORMATION:

SAGE Publications, Inc.
2455 Teller Road
Thousand Oaks, California 91320
E-mail: order@sagepub.com

SAGE Publications Ltd.
1 Oliver's Yard
55 City Road
London, EC1Y 1SP
United Kingdom

SAGE Publications India Pvt. Ltd.
B 1/I 1 Mohan Cooperative Industrial Area
Mathura Road, New Delhi 110 044
India

SAGE Publications Asia-Pacific Pte. Ltd.
18 Cross Street #10-10/11/12
China Square Central
Singapore 048423

Acquisitions Editor: Leah Fargotstein
Editorial Assistant: Elizabeth Cruz
Production Editor: Bennie Clark Allen
Copy Editor: Alison Hope
Typesetter: Hurix Digital
Proofreader: Scott Oney
Indexer: Integra
Cover Designer: Lysa Becker
Marketing Manager: Victoria Velasquez



Printed in the United States of America

*Library of Congress Cataloging-in-Publication Data*

Names: Carlson, Kieth A., author. | Winquist, Jennifer R., author.

Title: An introduction to statistics : an active learning approach / Kieth A. Carlson, Jennifer R. Winquist, Valparaiso University.

Description: Third edition. | Thousand Oaks, CA : SAGE, [2022] | Includes bibliographical references and index.

Identifiers: LCCN 2020036989 | ISBN 9781544375090 (paperback) | ISBN 9781544375106 (epub) | ISBN 9781544375113 (epub) | ISBN 9781544375120 (ebook)

Subjects: LCSH: Social sciences—Statistical methods. | Statistics.

Classification: LCC HA29 .C288 2022 | DDC 519.5—dc23

LC record available at https://lccn.loc.gov/2020036989

This book is printed on acid-free paper.

21 22 23 24 25 10 9 8 7 6 5 4 3 2 1

# BRIEF CONTENTS


**Preface** xix

**Acknowledgments** xxv

**About the Authors** xxvii

**PART 1 • DESCRIPTIVE STATISTICS AND SAMPLING ERROR** **1**

Chapter 1 • Introduction to Statistics and Frequency Distributions 3

Chapter 2 • Central Tendency and Variability 35

Chapter 3 • *z* Scores 77

Chapter 4 • Sampling Error and Confidence Intervals With *z* and *t* Distributions 97

**PART 2 • APPLYING THE FOUR PILLARS OF SCIENTIFIC REASONING TO MEAN DIFFERENCES** **135**

Chapter 5 • Single Sample *t*, Effect Sizes, and Confidence Intervals 137

Chapter 6 • Related Samples *t*, Effect Sizes, and Confidence Intervals 191

Chapter 7 • Independent Samples *t*, Effect Sizes, and Confidence Intervals 229

Chapter 8 • One-Way ANOVA, Effect Sizes, and Confidence Intervals 271

Chapter 9 • Two-Way ANOVA, Effect Eizes, and Confidence Intervals 315

PART 3 • APPLYING THE FOUR PILLARS OF SCIENTIFIC REASONING TO ASSOCIATIONS 373

Chapter 10 • Correlations, Effect Sizes, and Confidence Intervals 375

Chapter 11 • Chi Square and Effect Sizes 415

**Appendices** **445**

**References** **465**

**Index** **471**

# DETAILED CONTENTS


**Preface** xix

**Acknowledgments** xxv

**About the Authors** xxvii

## PART 1 • DESCRIPTIVE STATISTICS AND SAMPLING ERROR 1

### Chapter 1 • Introduction to Statistics and Frequency Distributions 3

How to Be Successful in This Course 3
Math Skills Required in This Course 5
Statistical Software Options 6
Why Do You Have to Take Statistics? 7
The Four Pillars of Scientific Reasoning 8
Populations and Samples 11
Independent and Dependent Variables 13
Identify How a Variable Is Measured 15
- **Scales of Measurement** 15
- **Discrete vs. Continuous Variables** 18

Graphing Data 18
Shapes of Distributions 21
Frequency Distribution Tables 23
**Activity 1.1 • How Obedient Were Milgram's Participants?** 25
**Activity 1.2 • How Do Psychologists Measure Aggression?** 31

### Chapter 2 • Central Tendency and Variability 35

Frequency Distribution Graphs and Tables 36
- **Bar Graphs** 36
- **Data Plots** 38
- **Boxplots** 38
- **Tables** 39

Central Tendency: Choosing Mean, Median, or Mode 41
Computing Measures of Central Tendency 44

Variability: Range or Standard Deviation 46
Steps in Computing a Population's Standard Deviation 47
**Steps 1–3: Computing the Sum of the Squared Deviation Scores (*SS*) 48**
**Step 4: Computing the Variance ($\sigma^2$) 50**
**Step 5: Computing the Standard Deviation ($\sigma$) 50**
Steps in Computing a Sample's Standard Deviation 52
**Steps 1–3 : Obtaining the *SS* 52**
**Step 4: Computing the Sample Variance ($SD^2$) 52**
**Step 5: Computing the Sample Standard Deviation (*SD*) 53**
Constructing a Scientific Conclusion 54
**Activity 2.1 • Computing Central Tendency and Standard Deviation 57**
**Activity 2.2 • Identifying Causes of Variability 62**
**Activity 2.3 • Does Neonatal Massage Increase Infants' Weight? 67**
**Activity 2.4 • Choosing Measure of Central Tendency and Variability 71**

## Chapter 3 • *z* Scores 77

Computing and Interpreting *z* for a Raw Score 77
**Computing a *z* for a Single Score 78**
**Interpreting the *z* for a Single Score 79**
Finding Raw Score "Cut Lines" 80
Finding the Probability of *z* Scores Using the Standard Normal Curve 81
Positive *z* Score Example 84
**Step 1: Find the *z* Score 84**
**Step 2: Sketch a Normal Distribution and Shade Target Area 84**
**Step 3: Use a Unit Normal Table to Find the Area of Targeted Curve Proportion 84**
Negative *z* Score Example 85
**Step 1: Find the *z* Score 85**
**Step 2: Sketch a Normal Distribution and Shade Target Area 85**
**Step 3: Use a Unit Normal Table to Find the Area of Targeted Curve Proportion 86**
Proportion Between Two *z* Scores Example 87
**Step 1: Find the *z* Scores 87**
**Step 2: Sketch a Normal Distribution and Shade Target Area 87**
**Step 3: Use a Unit Normal Table to Find the Area of Targeted Curve Proportion 87**
**Activity 3.1 • *z* Scores and Probabilities 88**
**Activity 3.2 • What Are Your Big Five *z* Scores? 91**

## Chapter 4 • Sampling Error and Confidence Intervals With *z* and *t* Distributions 97

Sampling and Sampling Error 97
The Central Limit Theorem and the Standard Error of the Mean (*SEM*) 101
Applying the *SEM* to Find Statistical Evidence 110
**The Probability of a *z* for Sample Mean 110**

*Step 1: Compute the Observed Deviation* 111
*Step 2: Compute the Deviation Expected by Sampling Error* 111
*Step 3: Compute the Ratio Between Observed and Expected Deviation (z for a Sample Mean)* 112
*Step 4: Locate the z Score in the Distribution* 112
*Step 5: Look Up the z Score Probability* 112
*Step 6: Interpret the z Score Probability* 112
**The 95% Confidence Interval** **113**
*Step 1: Determine the Point Estimate* 113
*Step 2: Determine the Margin of Error* (MOE) 114
*Step 3: Compute the Confidence Interval* (CI) 114
*Step 4: Interpret the Boundary Values* 115
*UB*s and *LB*s Are Estimates 115
Some Values Are More Plausible Than Others 115
Wide *CI*s Are Imprecise Estimates 115
Consider the Plausible Extremes 116
**The Probability of a Single Sample *t* Test** **116**
*Step 1: Compute the Observed Deviation* 118
*Step 2: Compute the Deviation Expected by Sampling Error* 118
*Step 3: Compute the Ratio Between Observed and Expected Deviation (t for a Sample Mean)* 118
*Step 4: Locate the* t *Score in the Distribution* 118
*Step 5: Find the* t *Score Probability* 118
*Step 6: Interpret the* t *Score Probability* 119
**The 95% Confidence Interval Using $SEM_s$** **119**
*Step 1: Determine the Point Estimate* 119
*Step 2: Determine the Margin of Error* (MOE) 119
*Step 3: Compute the Confidence Interval* (CI) 120
*Step 4: Interpret the Boundary Values* 120
**Exact Probabilities vs. Probability Estimates** **120**
**Activity 4.1 • Applying the CLT: What Is Your Expected Sampling Error?** **121**
**Activity 4.2 • Computing Probabilities and Estimating Parameters** **128**
**Activity 4.3 • Choosing Test Statistics** **131**

## PART 2 • APPLYING THE FOUR PILLARS OF SCIENTIFIC REASONING TO MEAN DIFFERENCES 135

### Chapter 5 • Single Sample *t*, Effect Sizes, and Confidence Intervals 137

Four Pillars of Scientific Reasoning 137
Apply the Four Pillars of Scientific Reasoning 139
**Designing Your Study** **140**
Determining Sample Size 141
**Pillar 1: Hypothesis Testing With a Continuous *p* Value** **142**
*Step 1: Choose the Correct Test Statistic* 142
*Step 2: Examine Statistical Assumptions* 143
*Step 3: Identify the Statistical Null Hypothesis* 144
Scientific vs. Statistical Null Hypothesis 144

*The One-Tailed vs. Two-Tailed Controversy* 145
*Step 4: Graph the Sample Data* 147
*Step 5: Compute the Test Statistic (Single Sample* t *Test) to Determine the* p *Value* 148
Step 5a: Compute the Deviation Between the Sample Mean and the Test Value (e.g., Population Mean) 148
Step 5b: Compute the Expected Sampling Error 148
Step 5c: Compute the Test Statistic (Single Sample *t* Test) 149
Step 5d: Find and Interpret the *p* Value 149
*Misconceptions About the* p *Value* 151
*Continuous* p *Value vs. Threshold/Cutoff Controversy* 153
Traditional Threshold/Cutoff Approach 153
Continuous *p* Value Approach 154
**Using Three-Valued Logic to Interpret Results** **155**
**Pillar 2: Practical Importance With Effect Size** **156**
*Step 6: Compute an Effect Size and Describe It* 156
Mean Difference Effect Size 156
Standardized Effect Size (*d*) 156
**Pillar 3: Population Estimation With Confidence Intervals** **157**
*Step 7: Compute Confidence Intervals* 157
Step 7a: Compute the Confidence Interval Around the Mean Difference 160
*UB*s and *LB*s Are Estimates 161
Some Values Are More Plausible Than Others 161
Wide *CI*s Are Imprecise Estimates 161
Consider the Plausible Extremes 161
Step 7b: Compute the Confidence Interval Around the Standardized Effect Size (*d*) 162
**Pillar 4: Research Methodology and Scientific Literature** **163**
*Step 8: Consider the Scientific Context* 163
Step 8a: Evaluate the Methodological Rigor of the Study 163
Step 8b: Determine How the Study Fits Into the Scientific Literature 164
Construct a Well-Supported Scientific Conclusion 164
*p* Value 165
Effect Size 165
Confidence Interval 166
Consider the Methodology and Scientific Literature 166
**Activity 5.1 • Are Students Overly Optimistic About Time Requirements?** **167**
**Activity 5.2 • Do Students Misestimate Studying Time?** **174**
**Activity 5.3 • Are People Overconfident About Their Driving Ability?** **179**
**Activity 5.4 • Is a Health at Every Size (HAES) Approach Associated With Weight Loss?** **184**
**Activity 5.5 • Cups vs. Spoons** **188**
**Activity 5.6 • Choosing Test Statistics** **189**

**Chapter 6 • Related Samples *t*, Effect Sizes, and Confidence Intervals** **191**

Related Samples *t* Test 191

Logic of the Single Sample and Related Samples *t* Tests 193

Apply the Four Pillars of Scientific Reasoning 195

**Designing Your Study** **196**

**Determining Sample Size** **196**

**Pillar 1: Hypothesis Testing With a Continuous *p* Value** **196**

*Step 1: Choose the Correct Test Statistic* *196*

*Step 2: Examine Statistical Assumptions* *196*

*Step 3: Identify the Statistical Null Hypothesis* *197*

*Step 4: Graph the Data for Each IV Condition* *198*

*Step 5: Compute the Test Statistic (Related Samples* t *Test) to Determine the* p *Value* *199*

Step 5a: Compute *D* for Each Participant/Matched Pair 199

Step 5b: Compute the Observed Mean Difference ($M_D$) 200

Step 5c: Compute the Typical Mean Difference Expected Due to Sampling Error 200

Step 5d: Compute the Test Statistic (Related Samples *t* Test) and Interpret the *p* Value 202

**Pillar 2: Practical Importance With Effect Size** **203**

*Step 6: Compute an Effect Size, Mean Difference or* d, *and Describe It* *203*

Mean Difference Effect Size 203

Standardized Effect Size (*d*) 204

**Pillar 3: Population Estimation With Confidence Intervals** **205**

*Step 7: Compute Confidence Intervals* *205*

Step 7a: Compute the Confidence Interval Around the Mean Difference 206

Step 7b: Compute the Confidence Interval Around the Standardized Effect Size ($d_z$) 206

**Pillar 4: Research Methodology and Scientific Literature** **207**

*Step 8: Consider the Scientific Context* *207*

Step 8a: Evaluate the Methodological Rigor of the Study 207

Step 8b: Determine How the Study Fits Into the Scientific Literature 208

Construct a Well-Supported Scientific Conclusion 208

*p* Value 208

Effect Size 208

*CI*s 208

**Activity 6.1 • Does an Intervention Reduce Intent to Use Cell Phones While Driving?** **210**

**Activity 6.2 • Does an Intervention Reduce Intent to Use Hands-Free Cell Phones While Driving?** **218**

**Activity 6.3 • Data Collection: When Does Multitasking Hinder Performance?** **223**

**Activity 6.4 • Data Collection: Does Switching Tasks Hinder Performance?** **225**
**Activity 6.5 • Choosing Test Statistics** **227**

**Chapter 7 • Independent Samples *t*, Effect Sizes, and Confidence Intervals** **229**

When to Use the Three *t* Tests 229
The *t* Test Logic and the Independent Samples *t* Formula 231
Apply the Four Pillars of Scientific Reasoning 232
**Designing Your Study** **232**
One- vs. Two-Tailed Hypothesis 232
Determining Sample Size 233
**Pillar 1: Hypothesis Testing With a Continuous *p* Value** **233**
*Step 1: Choose the Correct Test Statistic* *233*
*Step 2: Examine Statistical Assumptions* *234*
*Step 3: Identify the Statistical Null Hypotheses* *235*
*Step 4: Graph the Data for Each IV Condition* *236*
*Step 5: Compute the Test Statistic (Independent Samples* t *Test) to Determine the* p *Value* *236*
Step 5a: Compute the Deviation Between the Two Sample Means 236
Step 5b: Compute the Expected Sampling Error 236
Step 5c: Compute the Test Statistic (Independent *t* Test) 238
**Pillar 2: Practical Importance With Effect Size** **238**
*Step 6: Compute an Effect Size and Describe It* *238*
**Pillar 3: Population Estimation With Confidence Intervals** **239**
*Step 7: Compute CIs* *239*
Step 7a: Compute the Confidence Interval Around the Sample Mean Difference 239
Step 7b: Compute the Confidence Interval Around the Effect Size 240
**Pillar 4: Research Methodology and Scientific Literature** **240**
*Step 8: Consider the Scientific Context* *240*
Step 8a: Evaluate the Methodological Rigor of the Study 240
Step 8b: Determine How the Study Fits Into the Scientific Literature 240
Construct a Well-Supported Scientific Conclusion 241
*p* Value 241
Effect Size 241
*CI*s 241
Methodology and Scientific Literature 241
How to Interpret High *p* Values 242
**Group Report of Collected Data** **245**
**Activity 7.1 • Does Acetaminophen Reduce Social Pain?** **246**
**Activity 7.2 • Does Anchoring Influence Judgment?** **254**
**Activity 7.3 • Does Anchoring Influence Sales?** **258**
**Activity 7.4 • Data Collection and *t* Test Review: Does Body Position Influence Heart Rate?** **260**
**Activity 7.5 • Evaluating Research Summaries** **262**

**Activity 7.6 • Is Walking Beneficial? 265**
**Activity 7.7 • Choosing Test Statistics 266**

**Chapter 8 • One-Way ANOVA, Effect Sizes, and Confidence Intervals 271**

Independent Samples One-Way ANOVA 271
Logic of the ANOVA 272
Apply Four Pillars of Scientific Reasoning 276
**Designing Your Study 276**
**Determining Sample Size 276**
**Pillar 1: Hypothesis Testing With a Continuous *p* Value 276**
*Step 1: Choose the Correct Statistic 276*
*Step 2: Examine Statistical Assumptions 277*
*Step 3: Identify the Statistical Null Hypothesis 278*
*Step 4: Graph the Data for Each IV Condition 278*
*Step 5: Compute the Test Statistic (Independent ANOVA) to Determine the p Value 279*
Step 5a: Compute the Omnibus *F* 279
Step 5b: Find and Interpret the *p* Value 281
Omnibus *F* Controversy 281
Step 5c: If Necessary, Compute Post Hoc Tests 282
Post Hoc Controversy 284
**Pillar 2: Practical Importance With Effect Size 286**
*Step 6: Compute the Effect Sizes and Describe Them 286*
Step 6a: Effect Size for Overall ANOVA, $\eta_p^2$ or $\omega^2$ 286
Step 6b: Effect Size for Pairwise Comparisons, Mean Differences, or *d*s 287
Mean Difference Effect Size 288
Standardized Effect Size (*d*) 288
**Pillar 3: Population Estimation With Confidence Intervals 289**
*Step 7: Compute CIs 289*
**Pillar 4: Research Methodology and Scientific Literature 290**
*Step 8: Consider the Scientific Context 290*
Step 8a: Evaluate the Methodological Rigor of the Study 290
Step 8b: Determine How the Study Fits Into the Scientific Literature 291
**Construct a Well-Supported Scientific Conclusion 291**
*Overall ANOVA 291*
*p* Values 291
Effect Sizes 291
*Post Hoc Tests 291*
*p* Values 291
Effect Sizes 292
*CI*s 292
Methodology and Scientific Literature 292
**Activity 8.1 • Does Sleep Deprivation Lead to Social Media Use? 294**

**Activity 8.2 • Does Sleep Deprivation Lead to Increased Food Intake? 300**
**Activity 8.3 • Does Exercise Lead to More Sleep? 304**
**Activity 8.4 • Does Exercise Lead to Shorter Sleep Onset? 307**
**Activity 8.5 • Does Providing Books Increase Reading Time? 308**
**Activity 8.6 • How Do Study Methods Influence Test Performance (A Four-Group Example) 308**
**Activity 8.7 • Choosing Test Statistics 311**

**Chapter 9 • Two-Way Anova, Effect Sizes, and Confidence Intervals 315**

Purpose of Two-Way ANOVA 315
Logic of Two-Way ANOVA 318
Apply Four Pillars of Scientific Reasoning 321
**Designing Your Study 321**
Determining Sample Size 322
**Pillar 1: Hypothesis Testing With a Continuous *p* Value 322**
*Step 1: Choose the Correct Statistic 322*
*Step 2: Examine Statistical Assumptions 322*
*Step 3: Identify the Statistical Null Hypotheses 323*
Step 3a: Interaction Null Hypothesis 323
Step 3b: First Main Effect Null Hypothesis 325
Step 3c: Second Main Effect Null Hypothesis 326
*Step 4: Graph the Data for Each IV Condition 327*
Step 4a: Graph the Interaction Effect (i.e., Cell Means) 327
Step 4b: Graph the First Main Effect (i.e., First Pair of Marginal Means) 327
Step 4c: Graph the Second Main Effect (i.e., Second Pair of Marginal Means) 327
*Step 5: Compute the Test Statistics (Three F Tests) 328*
Step 5a: *F* Test and *p* Value for the Interaction 328
Step 5b: *F* Test for the First Main Effect (Study Method) 330
Step 5c: *F* Test for the Second Main Effect (Time Delay) 330
**Pillar 2: Assess Practical Importance With Effect Sizes 331**
*Step 6: Compute the Effect Sizes 331*
Step 6a: Compute Overall Effect Sizes (Interaction, Study Method, and Time Delay) 331
Step 6b: Compute Additional Effect Size Measures When Useful 334
*Mean Difference as Effect Size 334*
Main Effect of Studying Method 334
Main Effect of Time Delay 334
Study Method × Time Delay Interaction 334
**Pillar 3: Estimate Population Parameters With Confidence Intervals 335**
*Step 7: Population Estimation With 95% CIs 335*
Step 7a: Computing 95% *CI*s for Simple Effect Mean Differences 335
Step 7b: Computing 95% *CI*s for Mean Differences of Marginal Means 336

95% *CI* for the Main Effect for Study Method 336
95% *CI* for the Main Effect for Time Delay 336
**Pillar 4: Research Methodology and Scientific Literature 337**
*Step 8: Consider the Scientific Context 337*
Step 8a: Evaluate the Methodological Rigor of the Study 337
Step 8b: Determine How the Study Fits Into the Scientific Literature 337
Construct a Well-Supported Scientific Conclusion 337
**Interaction Effect (Studying Method × Time Delay) 338**
*p* Values 338
Effect Sizes 338
*CI*s 338
**First Main Effect (Time Delay) 338**
*p* Values 338
Effect Size 339
Mean Difference *CI* 339
**Second Main Effect (Studying Method) 339**
*p* Value 339
Effect Size 339
Mean Difference *CI* 339
Methodological Rigor 339
Existing Scientific Literature 339
General Format 340
APA-Style Summary 341
**Activity 9.1 • Do These Drugs Work Differently for Men vs. Women? (Part 1) 342**
**Activity 9.2 • Do These Drugs Work Differently for Men vs. Women? (Part 2) 349**
**Activity 9.3 • How Does Expertise Influence Memory? 355**
**Activity 9.4 • Are Cell Phones Worse Than Alcohol? 360**
**Activity 9.5 • Does Age Affect Treatment Effectiveness? 364**
**Activity 9.6 • Reading Research Article: Does Social Facilitation Theory Work on Cockroaches? 364**
**Activity 9.7 • Choosing Test Statistics 369**

**PART 3 • APPLYING THE FOUR PILLARS OF SCIENTIFIC REASONING TO ASSOCIATIONS 373**

**Chapter 10 • Correlations, Effect Sizes, and Confidence Intervals 375**
When to Use Correlations 375
The Logic of Correlation 376
Interpreting Correlation Coefficients 379
**Using Scatterplots to Identify Problems 381**
Outliers 381
Range Restriction 382
Heteroscedasticity 382
Nonlinear Relationships 383

Spearman's ($r_s$) Correlation 384
Correlation Does Not Equal Causation: True but Misleading 385
Apply the Four Pillars of Scientific Reasoning 385
**Designing Your Study** **386**
One- vs. Two-Tailed Correlations 386
Determining Sample Size 386
**Pillar 1: Hypothesis Testing With a Continuous *p* Value** **387**
*Step 1: Choose the Correct Test Statistic* *387*
*Step 2: Examine Statistical Assumptions* *387*
*Step 3: Identify the Null Hypothesis* *388*
*Step 4: Graph the Sample Data* *388*
*Step 5: Compute the Test Statistic (Pearson's r) to Determine the p Value* *388*
**Pillar 2: Practical Importance With Effect Size** **389**
*Step 6: Compute the Effect Size ($r^2$), and Describe It* *389*
**Pillar 3: Population Estimation With Confidence Intervals** **390**
*Step 7: Compute Confidence Interval (CI) Around r* *390*
**Pillar 4: Research Methodology and Scientific Literature** **391**
*Step 8: Consider the Scientific Context* *391*
Step 8a: Evaluate the Methodological Rigor of the Study 391
Step 8b: Determine How the Study Fits Into the Scientific Literature 392
Construct a Well-Supported Scientific Conclusion 392
*p* Value 392
Effect Size 392
*CI* 392
Methodological Rigor 393
Scientific Literature 393
**Activity 10.1 • Interpreting Scatterplots** **393**
**Activity 10.2 • Are Study Strategies Associated With Academic Performance? (Pilot Study)** **399**
**Activity 10.3 • Are Study Strategies Associated With Academic Performance?** **403**
**Activity 10.4 • Regression: How Well Do Reading Question Scores Predict Final Exam Scores?** **405**
**Activity 10.5 • Choosing Test Statistics** **410**

**Chapter 11 • Chi Square and Effect Sizes** **415**
When to Use $\chi^2$ Statistics 415
Logic of the $\chi^2$ Test 417
Apply the Pillars of Scientific Reasoning 419
**Goodness-of-Fit $\chi^2$** **419**
*Designing Your Study* *419*
Determining Sample Size 420
**Pillar 1: Hypothesis Testing With a Continuous *p* Value** **420**

*Step 1: Choose the Correct Test Statistic* 420
*Step 2: Examine Statistical Assumptions* 420
*Step 3: Identify the Statistical Null Hypotheses* 420
*Step 4: Graph the Sample Data* 421
*Step 5: Compute the Test Statistic (Goodness-of-Fit $\chi^2$)* 421
Step 5a: Determine the Expected Frequencies Defined by the Null Hypothesis 421
Step 5b: Compute the Test Statistic (Goodness-of-Fit $\chi^2$) 422
Step 5c: Find and Interpret the *p* Value 423
**Pillar 2: Practical Importance With Effect Size** **423**
*Step 6: Compute an Effect Size and Describe It* 423
**Pillar 3: Research Methodology and Scientific Literature** **423**
*Step 7: Consider the Scientific Context* 423
Step 7a: Evaluate the Methodological Rigor of the Study 423
Step 7b: Determine How the Study Fits Into the Scientific Literature 424
Construct a Well-Supported Scientific Conclusion 424
*p* Value 424
Methodological Rigor 424
Previous Literature 424
Apply the Pillars of Scientific Reasoning: $\chi^2$ for Independence 425
*Designing Your Study* 425
**Pillar 1: Hypothesis Testing With a Continuous *p* Value** **425**
*Step 1: Choose the Correct Test Statistic* 425
*Step 2: Examine Statistical Assumptions* 425
*Step 3: Identify the Null Hypothesis* 425
*Step 4: Graph the Sample Data* 426
*Step 5: Compute the Test Statistic ($\chi^2$ for Independence)* 427
Step 5a: Determine the Expected Frequencies Defined by the Null Hypothesis 427
Step 5b: Compute the Obtained Value 428
Step 5c: Find and Interpret the Overall *p* Value 428
Step 5d: Conduct Follow-Up Pairwise $\chi^2$ Analyses 428
**Pillar 2: Practical Importance With Effect Size** **429**
*Step 6: Compute the Effect Size and Describe It* 429
**Pillar 3: Research Methodology and Scientific Literature** **430**
*Step 7: Consider the Scientific Context* 430
Step 7a: Evaluate the Methodological Rigor of the Study 430
Step 7b: Determine How the Study Fits Into the Scientific Literature 431
*p* Value 431
Effect Size 431
Methodological Rigor and Scientific Context 431
**Construct a Well-Supported Scientific Conclusion** **431**
**Activity 11.1 • Goodness-of-Fit $\chi^2$: Do Voters Approve of Policy Changes?** **432**

Activity 11.2 • $\chi^2$ Goodness-of-Fit: Do Voters Approve of Tax Cuts? 435
Activity 11.3 • Independence $\chi^2$: Are Political Affiliation and Opinion Associated? 436
Activity 11.4 • Reading Research Article: Does Pre-registration Reduce Publication Bias in Clinical Trials? 437
Activity 11.5 • Are Political Affiliation and Age Associated With Support for Free College Tuition? 439
Activity 11.6 • Choosing Test Statistics 439

**Appendices** 445

**References** 465

**Index** 471

# PREFACE

## THE STORY OF THIS TEXT

We once attended a teaching workshop in which the speaker described a common experience in college classrooms and the pedagogical problems it frequently creates: Instructors carefully define basic concepts (e.g., population, sample) and gradually progress to applying those concepts to more-complex topics (e.g., sampling error) as the end of class approaches. Then students attempt homework assignments covering the more-complicated topics. All too frequently, students think they understand things while listening to us in class, but struggle when they attempt homework on their own. While some students can eventually figure things out, others become frustrated, and still others give up. The teaching workshop made us recognize, reluctantly, this happened to us (and our students) in our statistics classes. While we did our best to address this problem by refining our lectures, our students still struggled with homework assignments and we were disappointed with their exam performance. Students frequently said to us, "I understand it when you do it in class, but when I try it on my own it doesn't make sense." This common experience motivated us to change our statistics classes and, eventually, to write the first edition of this textbook.

We decided that we needed to change our course so (1) students came to class understanding basic concepts, (2) they used challenging concepts *during* class so we could answer their questions immediately, and (3) they learned to interpret and report statistical results as if they were researchers. We started our course revisions by emphasizing the importance of actually reading the text before class. Even though we were using excellent statistics textbooks, many students insisted that they needed lectures to help them understand the text. Eventually, we opted for creating our own readings that emphasize the basics (i.e., the "easy" stuff). We embedded relatively easy reading questions to help students *read with purpose,* so they came to class understanding the basic concepts. Next, over several years, we developed activities that reinforced the basics and also introduced more-challenging material (i.e., the "hard" stuff). Hundreds of students completed these challenging activities in our courses. After each semester we strove to improve every activity based on our students' feedback and exam performances.

Our statistics courses are dramatically different from what they were more than a decade ago when we first decided to change them. In our old classes, few students read prior to class and most class time was spent lecturing on the material in the book. In our current statistics courses, students answer online reading questions prior to class, we give very brief lectures at the beginning of class, and students complete activities (i.e., assignments) during class. We've compared our current students' attitudes about statistics to those taking our more-traditional statistics course (Carlson & Winquist, 2011) and found our current students to be more confident in their ability to perform statistics and to like statistics more than their peers. We've also learned that, after completing this revised statistics course, students score nearly half a standard deviation higher on a nationally standardized statistics test that they take during their senior year (approximately 20 months after taking the course) compared to students taking the more-traditional course (Winquist & Carlson, 2014).

Of course, not all our students master the course material. Student motivation still plays an important part in student learning. If students don't do the reading, or don't work on understanding the assignments in each chapter, they will still struggle. In our current courses, we encourage students to read and complete the assignments by giving points for completing them. We have found that, if students do these things, they do well in our courses. We have far fewer struggling students in our current courses than we had in our traditional course, even though our exams are more challenging.

## WHAT IS NEW IN THE THIRD EDITION?

*Fewer Chapters.* If you used the second edition of the text, the first thing you might notice is that the third edition has 11 chapters rather than 14, but we cover the same topics and in more depth. Specifically, we combined central tendency and variability into a single chapter, integrated confidence intervals into multiple chapters, and introduced hypothesis testing with single sample $t$ rather than $z$ for a sample mean. These changes resulted in three fewer chapters.

*Four Pillars of Scientific Reasoning.* Importantly, the third edition introduces Four Pillars of Scientific Reasoning (Hypothesis Testing With a Continuous $p$ Value, Practical Importance With Effect Size, Population Estimation With Confidence Intervals, and Methodology and Scientific Literature) to help students think about statistical results and create well-supported scientific conclusions. Consequently, this edition places more emphasis on interpreting statistics than the second edition did, and computations are used to help students understand how each statistic works. Perhaps more importantly, the third edition, following the advice of statisticians, encourages interpreting $p$ values continuously rather than dichotomously (i.e., using critical value cutoffs). A consequence of all these changes is that we've reorganized and rewritten nearly every chapter.

*More Software Options.* In previous editions, we focused exclusively on SPSS. We have started using JASP and jamovi in our courses because they are simple to use, provide effect sizes and confidence intervals, and are free. Instructors can use this textbook with any statistics program. Software instructions are posted on the textbook website.

*Decreased Emphasis on Hand Computations.* We continue to show hand computations for most analyses so that students can understand where the numbers come from. The activities are written so that instructors can decide if they want students to do the computations by hand and/or by using software.

*New Textbook Website.* Also new to the third edition is a new textbook website (learn-stats.com) containing many useful resources for students and instructors including

- data files for all activities in the book; and
- software instructions.

In addition, instructors have access to additional resources through the SAGE instructor website at **https://edge.sagepub.com/carlson3e** including

- answer keys for the activities;
- practice tests for each chapter;
- tests for each chapter; and
- lecture slides.

## HOW TO USE THIS BOOK

This text certainly could be used in a lecture-based course in which the activities function as detailed, conceptually rich homework assignments. We also are confident that there are creative instructors and students who will find ways to use this text that we never considered. However, it may be helpful to know how we use this text. In our courses, students read and answer online reading questions *twice* prior to class and receive the average of their two attempts. We begin classes with brief lectures (about 15 minutes) and then students work for the remaining 60 minutes to complete the assignments during class. We find three significant advantages in this approach. First, students come to class more prepared. Second, students do the easier work (i.e., answering foundational questions) outside of class and complete the more-difficult work in class when peers and an instructor can answer their questions. Third, students work at their own paces. We have used this approach for several years with positive results (Carlson & Winquist, 2011; Winquist & Carlson, 2014).

This approach encourages students to review and correct misunderstandings on the reading questions as well as the assignments. Mistakes are inevitable and even desirable. After all, each mistake is an opportunity to learn. In our view, students should first engage with the material without concern about evaluation. Therefore, we allow students to redo the activities at the end of each chapter as many times as they wish, awarding them their highest grade. This lenient approach to homework assignments encourages students to focus on checking their answers and then correcting their mistakes. We collect their answers (now online!) to confirm that they did the work for each chapter. Over the years, these assignments have constituted between 7% and 17% of students' course grades. We have tried making the chapter activities (i.e., homework assignments) optional so we did not have to grade them. We told students that completing the activities is the best way to prepare for exams, but we awarded no points for completing them. When we did this, we found greater variability in activity completion and exam performance.

Although we have not taught the class completely online (yet), we did recently teach it online for half of the semester and found that the materials worked very well in this online format. In fact, some students even preferred this format so that they could take their time and work at their own pace.

## UNIQUE FEATURES OF THIS TEXT

By now you probably recognize that this is not a typical statistics text. For ease of review we've listed and described the two most unique aspects of this text:

- *Embedded Reading Questions.* All 11 chapters contain embedded reading questions that focus students' attention on the key concepts *as they read* each paragraph/section of the text. Researchers studying reading comprehension report that similar embedded questions help students with lower reading abilities achieve levels of performance comparable to that of students with greater reading abilities (Callender & McDaniel, 2009).
- *Activities (Assignments).* All 11 chapters contain active learning assignments, called *Activities.* While the 11 chapters start by introducing foundational concepts, they are followed by activity sections in which students *test or*

*demonstrate their understanding of basic concepts while they read detailed explanations of more-complex statistical concepts.* When using most traditional textbooks, students perform statistical procedures *after* reading multiple pages. This text adopts a workbook approach in which students are actively performing tasks *while* they read explanations. Most of the activities are self-correcting so if students misunderstand a concept it is corrected early in the learning process. After completing these activities, students are far more likely to understand the material than if they simply read the material.

## Ancillaries

- *Instructors Manual.* Includes lecture outlines and answers to activities.
- *Blackboard Cartridges.* Includes reading questions, practice tests for each chapter, tests for each chapter's online activities, data files, and activity answers.
- *Test Bank Questions.* Exam questions are available to instructors of the course. Questions are formatted for online use.

## Appropriate Courses

This text is ideal for introductory statistics courses in psychology, sociology, and social work, as well as the health, exercise, or life sciences. The text would work well for any course intending to teach the statistical procedures of hypothesis testing, effect sizes, and confidence intervals that are commonly used in the behavioral sciences.

Finally, we are grateful to SAGE Publications for giving us the opportunity to share this text with others.

# ACKNOWLEDGMENTS

We greatly appreciate the invaluable feedback of our students, without whom this text would not have been possible. We also thank the reviewers listed here who helped us improve the text with their insightful critiques and comments.

Marie T. Balaban, Eastern Oregon University

Tyson S. Barrett, Utah State University

Christina L. Crosby, San Diego Mesa College

Amanda ElBassiouny, California Lutheran University

Franklin H. Foote, University of Miami

Indranil Goswami, State University of New York, Buffalo

Brooke Magnus, Marquette University

Amy R. Pearce, Arkansas State University, Jonesboro

Chrislyn Randell, Metropolitan State University of Denver

Joshua Williams, California Lutheran University

Leora Yetnikoff, College of Staten Island, City University of New York

# ABOUT THE AUTHORS

**Kieth A. Carlson** received his doctoral degree in experimental psychology with an emphasis in cognitive psychology from the University of Nebraska in 1996. He is currently professor of psychology at Valparaiso University. He has published research on visual attention, memory, student cynicism toward college, and active learning. He enjoys teaching a wide range of courses including statistics, research methods, sensation and perception, cognitive psychology, learning psychology, the philosophy of science, and the history of psychology. Dr. Carlson was twice honored with the Teaching Excellence Award from the United States Air Force Academy.

**Jennifer R. Winquist** is currently professor of psychology at Valparaiso University. Dr. Winquist received her doctoral degree in social psychology from the University of Illinois at Chicago and her bachelor's degree in psychology from Purdue University. She has published research on self-focused attention, group decision-making, distributive justice, and the scholarship of teaching and learning. Dr. Winquist regularly teaches courses in introductory and advanced statistics and research methods.

PART 1

# DESCRIPTIVE STATISTICS AND SAMPLING ERROR

1

# INTRODUCTION TO STATISTICS AND FREQUENCY DISTRIBUTIONS

## LEARNING OBJECTIVES

- Explain how you can be successful in this course.
- Explain why many academic majors require a statistics course.
- Describe the four pillars of scientific reasoning.
- Use common statistical terms correctly in a statistical context.
- Identify how a variable is measured.
- Use software to generate frequency distribution tables and graphs.
- Interpret frequency distribution tables, bar graphs, histograms, and line graphs.
- Explain when to use a bar graph, histogram, and line graph.

## HOW TO BE SUCCESSFUL IN THIS COURSE

Have you ever read a few pages of a textbook and realized you were not thinking about what you were reading? Your mind wandered to topics completely unrelated to the text, and you could not identify the point of the paragraph (or sentence) you just read. For most of us, this experience is not uncommon even when reading books that we've chosen to read for pleasure. Therefore, it is not surprising that our minds wander while reading textbooks. Although this lack of focus is understandable, it seriously hinders effective learning. Thus, one goal of this book is to discourage minds from wandering and to encourage reading with purpose. To some extent, you need to force yourself to read with purpose. As you read each paragraph, ask "What is the purpose of this paragraph?" or "What am I supposed to learn from this paragraph?" If you make asking these questions a habit your reading comprehension will improve.

This text is structured to make it easier for you to read with purpose. The short chapters have frequent reading questions embedded in the text that make it easier for you to remember key points. Resist the temptation to go immediately to the reading questions and search for answers in the preceding paragraphs. *Read first, and then answer the questions as you come to them.* Using this approach will increase your understanding of the material in this text.

**Reading Question**

1. Reading with purpose means
   a. thinking about other things while you are reading a textbook.
   b. actively trying to extract information from each paragraph as you read.

**Reading Question**

2. Is it better to read the paragraph and then answer the reading question or to read the reading question and then search for the answer? It's better to
   a. read the paragraph, then answer the reading question.
   b. read the reading question, then search for the question's answer.

The chapter readings introduce the material giving you a foundation in the basics before you attempt the activities at the end of the chapter. The activities give you the opportunity to assess your knowledge of the chapter material by working with research problems. Your emphasis when working on the activities should be on understanding your answers. If you generate a wrong answer, figure out your error. We often think of errors as things that should be avoided at all costs. However, quite the opposite is true. Making mistakes and fixing them is a great way to learn. If you find your errors and correct them, you will probably not repeat them. Resist the temptation to "get the right answer quickly." It is more important that you understand why every answer is correct.

**Reading Question**

3. When completing activities, your primary goal should be to get the correct answer quickly.
   a. True
   b. False

## MATH SKILLS REQUIRED IN THIS COURSE

The math skills required in this course are fairly basic. You need to be able to add, subtract, multiply, divide, square numbers, and take the square root of numbers using a calculator. You also need to be able to do some basic algebra. For example, you should be able to solve the following equation for $X$: $22 = \frac{X}{3}$. [The correct answer is $X = 66$.]

**Reading Question**

4. This course requires basic algebra.
   a. True
   b. False

**Reading Question**

5. Solve the following equation for $X$: $30 = \frac{X}{3}$.
   a. 10
   b. 90

You will also need to follow the correct order of mathematical operations. As a review, the correct order of operations is (1) the operations in parentheses, (2) exponents, (3) multiplication or division, and (4) addition or subtraction. Some of you may have learned the mnemonic, Please Excuse My Dear Aunt Sally, to help remember the correct order. For example, when solving the equation $(3 + 4)^2$, you would first add $(3 + 4)$ to get 7 and then square the 7 to get 49. Try to solve the next more complicated problem. The answer is 7.125. If you have trouble with this problem, talk with your instructor about how to review the necessary material for this course.

$$X = \frac{(6-1)3^2 + (4-1)2^2}{(6-1)+(4-1)}$$

**Reading Question**

6. Solve the following equation for $X$: $X = \frac{(3-1)4^2 + (5-1)3^2}{(3-1)+(5-1)}$
   a. 11.33
   b. 15.25

You will use statistical software or a calculator to perform computations in this course. You should be aware that order of operations is very important when using your calculator. Unless you are very comfortable with the parentheses buttons on your calculator, we recommend that you do one step at a time rather than trying to enter the entire equation into your calculator.

**Reading Question**

7. Order of operations is only important when doing computations by hand, not when using your calculator.
   a. True
   b. False

Although the math in this course should not be new, you will see new notation throughout the course. When you encounter this new notation, relax and realize that the notation is simply a shorthand way of giving instructions. While you will be learning how to interpret numbers in new ways, the actual mathematical skills in this course are no more complex than the order of operations. The primary goal of this course is teaching you to use numbers to make decisions. Occasionally, we will give you numbers solely to practice computation, but most of the time you will use the numbers you compute to make decisions within a specific, real-world context.

If you ever do research in a subsequent class or for a career after you graduate, you almost certainly will NOT conduct your statistical analyses by hand. Rather, you will use statistical software. It is reasonable to ask, "Why am I crunching numbers by hand if nobody does hand calculations anymore?" Many statistics instructors believe that some amount of number crunching is necessary to understand foundational concepts like variability and expected sampling error. These concepts are so fundamental to this course that computing them by hand helps you understand and subsequently interpret statistical results more accurately than if you don't understand the math behind these concepts. There is less agreement among statistics educators on how much number crunching is necessary. In this text, we think computing numerous forms of sampling error by hand is helpful, but after you understand these concepts it's probably more beneficial to focus on learning how to use statistical software.

## STATISTICAL SOFTWARE OPTIONS

There are many statistical software packages available, each with advantages and disadvantages. Two widely used packages are IBM SPSS Statistics and R (r-project.org). SPSS, which stands for Statistical Package for the Social Sciences, offers a point-and-click interface that makes running common statistical analyses relatively easy. However, SPSS is expensive and, unless students purchase or rent their own copy of the program, it can be difficult to access off campus. The R statistical package is completely free but is substantially more complicated to use. Rather than having a point-and-click interface, it requires typing code. If you learn how to write the R code (or syntax) you have tremendous flexibility in the statistical analyses you can run; learning the code requires a level of sophistication that presents challenges for introductory statistics students, however. Fortunately, there are other options that offer a point-and-click interface and that are also free. Both JASP (jasp-stats.org) and jamovi (jamovi.org) offer point-and-click interfaces that are more intuitive than SPSS's interface and both are completely free. Additionally, both JASP and jamovi perform important statistics that SPSS does not perform. For this reason, we recently switched from using SPSS to JASP in our introductory statistics courses. Table 1.1 compares the current capabilities of JASP, jamovi, SPSS, and R. You can find instructions for running common statistical procedures in all three programs online, although you can use any software to complete the activities in the book.

**TABLE 1.1 ⬢ Comparison of JASP, jamovi, SPSS, and R Statistical Programs**

| | JASP Version 0.13.1.0 | jamovi Version 1.1.9.0 | SPSS Version 25 | R |
|---|---|---|---|---|
| Easy point-and-click interface | ✓+ | ✓+ | ✓ | |
| Sample mean, *N*, *SD*, and *SEM* | ✓ | ✓ | ✓ | ✓ |
| Mean difference, *SD*, and *SEM* | ✓ | ✓ | ✓ | ✓ |
| *t* tests and correlations with *p* value | ✓ | ✓ | ✓ | ✓ |
| Effect size (*d*) | ✓ | ✓ | | ✓ |
| Confidence interval around mean | | | ✓ | ✓ |
| Confidence interval around mean difference | ✓ | ✓ | ✓ | ✓ |
| Confidence interval around effect sizes | ✓ | ✓ | | ✓ |
| Factorial ANOVA simple effects analysis without requiring syntax | ✓ | ✓ | | |
| Price | Free | Free | $$$ | Free |

**Reading Question**

8. This textbook can be used with a variety of different statistical programs, some of which are free.
   a. True
   b. False

## WHY DO YOU HAVE TO TAKE STATISTICS?

You are probably reading this book because you are required to take a statistics course to complete your degree. Students majoring in business, economics, nursing, political science, pre-medicine, psychology, social work, and sociology are often required to take at least one statistics course. There are many reasons why statistics is a mandatory course for students in these varied disciplines. But the primary reason is that in every one of these disciplines people make decisions that have the potential to improve people's lives, and these decisions should be informed by data. For example, a psychologist may conduct a study to determine if a new treatment reduces the symptoms of depression. Based on this study, the psychologist will need to decide if the treatment is or is not effective. If the wrong decision is made, an opportunity to help people with depression may be missed. Even more troubling, a wrong decision might harm people. While statistical methods will not eliminate wrong decisions, understanding statistical methods will allow you to reduce the number of wrong decisions you make. You are taking this course because the professionals in your discipline recognize that statistical methods can improve decision making. Statistics make us more professional.

**Reading Question**

9. Why do many disciplines require students to take a statistics course? Taking a statistics course
   a. is a way to employ statistics instructors, which is good for the economy.
   b. can help people make better decisions in their chosen professions.

## THE FOUR PILLARS OF SCIENTIFIC REASONING

People in many professions rely on statistical evidence to make scientifically supported decisions. For example, medical, sociological, economic, and psychological researchers all rely on statistical evidence when they recommend one course of action (e.g., a drug treatment, a social policy, a tax policy, or a learning strategy) rather than another. When the statistical evidence is evaluated thoughtfully, there is little doubt that these statistically justified decisions benefit our society with healthier people, better schools, and better government. However, all too frequently statistical evidence is not evaluated thoughtfully. Unfortunately, statistical evidence can be intentionally misrepresented or unintentionally misinterpreted. Either leads to less-than-optimal, even detrimental, outcomes (e.g., higher death rates, lower graduation rates, or less trust in government). Without being overly dramatic, thoughtfully interpreted statistical evidence can be a positive force for improvement, but misinterpreted statistical evidence hinders, even obstructs, progress. Clearly, our society would be better off if more of us could thoughtfully interpret statistical evidence.

Constructing accurate scientific conclusions based on statistical evidence can be tricky, but two simple recommendations can help. First, when constructing a scientific conclusion, use multiple types of evidence. You will learn that different statistical procedures have different advantages and potential problems. Therefore, by using multiple statistical procedures to inform the same scientific decision you gain the advantages of each; if used properly, these differing advantages can compensate for the potential weakness of any single statistical procedure. If constructing a sound scientific conclusion were like constructing a house, using multiple statistical procedures would be like using multiple different types of materials (e.g., wood, glass, and aluminum) to make a house rather than just one material. A house incorporating the advantages of wood, glass, and aluminum is generally more desirable than one constructed from just one type of material. Similarly, a scientific conclusion constructed from multiple statistical procedures possesses more desirable characteristics. Second, when constructing a scientific conclusion, evaluate the strength of your evidence and adjust your conclusion accordingly. Grand scientific conclusions, those made with great certainty, require extraordinary evidence. When your evidence is weak, your conclusions should be tentative. The best scientific conclusions are not overly certain but rather appropriately certain; they convey uncertainty when the evidence is uncertain.

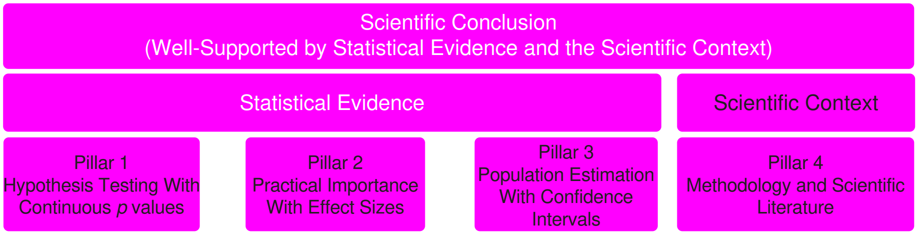
Within any scientific context, using multiple types of evidence and adjusting the certainty of one's conclusions based on the evidence's strength are sound recommendations. Within the medical and social sciences, following this advice frequently means using hypothesis testing, effect sizes, and confidence intervals and then interpreting these statistics based on the methodology used to generate them and their consistency with the scientific literature. The first three are different types of statistical evidence, each offering its own strengths and weaknesses. We will discuss each of these in depth in later chapters. Consistent with using multiple kinds of evidence, we advocate using these three statistical procedures jointly when answering scientific questions and then interpreting (i.e., contextualizing) them based on methodology and scientific literature. Collectively, as Figure 1.1 suggests, these four types of evidence form the foundation of a sound, appropriately certain, scientific conclusion. In fact, we view these four sources of evidence as so foundational to scientific decision-making that from this point forward we refer to them as the four pillars of scientific reasoning. For clarity, the four pillars of scientific reasoning are

1. hypothesis testing with a continuous *p* value,
2. practical importance with effect size,
3. population estimation with confidence intervals, and
4. research methodology and scientific literature.

Using these four pillars when making scientific decisions will likely lead to well-supported, scientific conclusions. We will be discussing each of these pillars in greater detail in subsequent chapters, but for now we offer a brief description of each.

The first pillar is hypothesis testing with a continuous *p* value, *a formal multiple-step procedure for evaluating a null hypothesis*. This statistical procedure is also called **significance testing** or **null hypothesis testing**. In later chapters, you will learn a variety of statistics that test different hypotheses. All of the hypothesis-testing procedures that you will learn are needed because of one fundamental problem that

**FIGURE 1.1 ⬢ Concept Map Depicting How the Four Pillars of Scientific Reasoning Support a Sound Scientific Conclusion**




plagues all researchers—the problem of sampling error. For example, researchers evaluating a new depression treatment want to know if it effectively lowers depression in all people with depression, called the population of people with depression. However, researchers cannot possibly study every depressed person in the world. Instead, researchers have to study a subset of this population, perhaps a sample of 100 people with depression. *The purpose of any sample is to represent the population from which it came.* In other words, if the 100 people with depression are a good sample, they will be similar to the population of people with depression. Thus, if the average score on a clinical assessment of depression in the population is 50, the average score of a good sample will also be 50. Likewise, if the ratio of women with depression to men with depression is 2:1 in the population, it will also be 2:1 in a good sample. Of course, you do not really expect a sample to be exactly like the population. *The differences between a sample and the population create* **sampling error**.

**Reading Question**

10. All hypothesis-testing procedures were created so that researchers could
    a. study entire populations rather than samples.
    b. deal with sampling error.

**Reading Question**

11. If a sample represents a population well, it will
    a. respond in a way that is similar to how the entire population would respond.
    b. generate a large amount of sampling error.

While hypothesis testing is extremely useful, it has limitations. Therefore, another primary purpose of this book is to describe these limitations and how researchers address them by using two additional statistical procedures. **Effect sizes** *describe the magnitude of a study's results,* helping researchers determine if a research result is large enough to be useful or if it is too small to be meaningful in real-world situations. In other words, effect sizes help researchers determine the *practical importance* of a study's results. Another valuable statistical procedure is a confidence interval. **Confidence intervals** *estimate the plausible values that might occur in the population based on the likely sampling error* *in the study.* Samples rarely represent the population perfectly; consequently, confidence intervals help researchers estimate how a study's results might generalize to the entire population. Each of these statistical procedures help researchers give meaning to the results of a hypothesis test. In fact, the American Psychological Association (APA) Publication Manual recommends that researchers use effect sizes and confidence intervals whenever hypothesis tests are used (American Psychological Association, 2020). These three statistical procedures are most beneficial when they are used side by side.

**Reading Question**

12. Effect sizes and confidence intervals help researchers
    a. interpret (i.e., give meaning to) the results of hypothesis tests.
    b. address the limitations of hypothesis tests.
    c. Do both of the above.

After computing the hypothesis test, effect size, and confidence intervals, you need to **contextualize these statistical results** by evaluating the methodological rigor of the study and comparing the results to the existing scientific literature. If a study's methodology had limitations, your confidence in the statistical results should be limited as well. A major methodological flaw might even make the statistical results meaningless. Accurately interpreting statistical results requires careful consideration of the methodology used to generate those results. Contextualizing results also requires comparing your results to those in the scientific literature. You should have more confidence in results that have been replicated many times than you have in those that contradict the existing scientific literature. While contradictory results are not necessarily wrong, it is wise to interpret them cautiously when they contradict well-established theory or a large body of empirical findings.

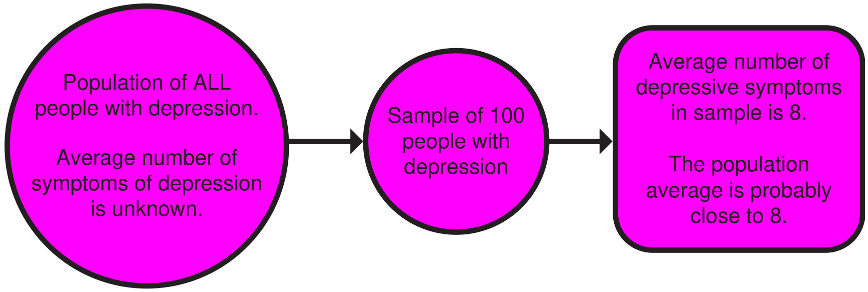
**Reading Question**

13. Before constructing a final scientific conclusion, you should consider
    a. the statistical evidence provided by hypothesis tests, effect sizes, and confidence intervals.
    b. the methodological strengths and weaknesses of the study that generated the data.
    c. the existing scientific literature.
    d. All of the above.

## POPULATIONS AND SAMPLES

Suppose that a researcher studying depression gave a new treatment to a sample of 100 people with depression. Figure 1.2 is a pictorial representation of this research scenario. The large circle on the left represents a **population**, *a group of all things that share a set of characteristics.* In this case, the "things" are people, and the characteristic they

**FIGURE 1.2** ⬢ A Pictorial Representation of Using a Sample to Estimate a Population Parameter (i.e., Inferential Statistics)




all share is depression. Researchers want to know what the mean depression score for the population would be if all people with depression were treated with the new depression treatment. In other words, researchers want to know the **population parameter**, *the value that would be obtained if the entire population were actually studied*. Of course, the researchers don't have the resources to study every person with depression in the world, so they must instead study a **sample**, *a subset of the population that is intended to represent the population*. In most cases, the best way to get a sample that accurately represents the population is by taking a **random sample** from the population. When taking a **random sample**, *each individual in the population has the same chance of being selected for the sample*. In other words, while researchers want to know a population parameter, their investigations usually produce a **sample statistic**, *the value obtained from the sample*. The researchers then use the sample statistic value as an estimate of the population parameter value. The researchers are making an *inference* that the sample statistic is a value similar to the population parameter value based on the premise that the characteristics of those in the sample are similar to the characteristics of those in the entire population. *When researchers use a sample statistic to infer the value of a population parameter* it is called **inferential statistics**. For example, a researcher studying depression wants to know how many depressive symptoms are exhibited by people in the general population. He can't survey everyone in the population and so he selects a random sample of people from the population and finds that the average number of symptoms in the sample is eight (see Figure 1.2). If he then inferred that the entire population of people would have an average of eight depressive symptoms, he would be basing his conclusion on inferential statistics. It should be clear to you that if the sample did not represent the population well (i.e., if there was a lot sampling error), the sample statistic would NOT be similar to the population parameter. In fact, **sampling error** is defined as *the difference between a sample statistic value and an actual population parameter value*.

Reading Question

14. The value obtained from a population is called a
    a. statistic.
    b. parameter.

Reading Question

15. Parameters are
    a. always exactly equal to sample statistics.
    b. often estimated or inferred from sample statistics.

Reading Question

16. When a statistic and parameter differ,
    a. it is called an inferential statistic.
    b. there is sampling error.

The researchers studying depression were using inferential statistics because they were using data from a sample to infer the value of a population parameter. The component of the process that makes it inferential is that researchers are using data they actually have to estimate (or infer) the value of data they don't actually have. In contrast, researchers use **descriptive statistics** *when their intent is to describe the data that they actually collected*. For example, if a clinical psychologist conducted a study in which she gave some of her clients a new depression treatment and she wanted to describe the average depression score of only those clients who got the treatment, she would be using descriptive statistics. Her intent is only to describe the results she observed in the clients who actually got the treatment. However, if she then wanted to estimate what the results would be if she were to give the same treatment to additional clients, she would then be performing inferential statistics.

Reading Question

17. Researchers are using descriptive statistics if they are using their results to
    a. estimate a population parameter.
    b. describe the data they actually collected.

## INDEPENDENT AND DEPENDENT VARIABLES

Researchers design experiments to test if one or more variables cause changes to another variable. For example, if a researcher thinks a new treatment reduces depressive symptoms, he could design an experiment to test this prediction. He might give a sample of people with depression the new treatment and a placebo treatment to another sample of people with depression. Later, if those who received the new treatment had lower levels of depression, he would have evidence that the new treatment reduces depression. In this experiment, the type of treatment each person received (i.e., new treatment vs. no treatment) is the **independent variable (IV)**. This experiment's IV has two **IV levels**: (1) the new treatment and (2) the placebo treatment. The main point of the study is to determine if the two different IV levels were differentially effective at reducing depressive symptoms. More generally, *the IV is a variable with two or more levels that are expected to have different impacts on another variable.* In this study, after both samples of people with depression were given their respective treatment levels, the amount of depression in each sample was compared by counting the number of depressive symptoms in each person. In this experiment, the number of depressive symptoms observed in each person is the **dependent variable (DV)**. Given that the researcher expects the new treatment to work and the placebo treatment not to work, he expects the new treatment DV scores to be lower than the placebo treatment DV scores. More generally, *the DV is the outcome variable that is used to compare the effects of the different IV levels.*

Reading Question

18. The IV (independent variable) in a study is the
    a. variable expected to change the outcome variable.
    b. outcome variable.

In true experiments, those in which researchers manipulate a variable so that some participants have one value and others have a different value, the manipulated variable is referred to as the IV. For example, if a researcher gives some participants a drug (Treatment A) and others a placebo (Treatment B), this manipulation defines the IV of Treatment as having two levels, namely, drug and placebo. However, in this text, we also use IV in a more general way. The IV is any variable predicted to influence another variable even when the IV was not manipulated. For example, if a researcher predicted that women would be more depressed than men, we will refer to gender as the IV because it is the variable that is expected to influence the DV (i.e., depression score). If you take a research methods course, you will learn an important distinction between *manipulated* IVs (e.g., Type of Treatment: Drug vs. Placebo) and *measured* IVs (e.g., Gender: Male vs. Female). Very briefly, the ultimate goal of science is to discover causal relationships, and manipulated IVs allow researchers to draw causal conclusions while measured IVs do not. You can learn more about this important distinction and its implications for drawing causal conclusions in a research methods course. It is important to keep in mind that statistics alone do not allow you to determine if an IV causes changes in a DV. An understanding of research methodology is just as important as an understanding of statistics in order to critically evaluate research claims.

Reading Question

19. All research studies allow you to determine if the IV causes changes in the DV.
    a. True
    b. False

# IDENTIFY HOW A VARIABLE IS MEASURED

## Scales of Measurement

All research is based on measurement. For example, if researchers are studying depression, they will need to devise a way to measure depression accurately and reliably. The way a variable is measured has a direct impact on the types of statistical procedures that can be used to analyze that variable. Generally speaking, researchers want to devise measurement procedures that are as precise as possible because more-precise measurements enable more-sophisticated statistical procedures. Researchers recognize four different **scales of measurement** that vary in their degree of measurement precision: (1) nominal, (2) ordinal, (3) interval, and (4) ratio (Stevens, 1946). Each of these scales of measurement is increasingly more precise than its predecessor, and, therefore, each succeeding scale of measurement allows more-sophisticated statistical analyses than its predecessor.

**Reading Question**

20. The way a variable is measured
    a. determines the kinds of statistical procedures that can be used on that variable.
    b. has very little impact on how researchers conduct their statistical analyses.

For example, researchers could describe depression using a nominal scale by categorizing people with different kinds of major depressive disorders into groups, including those with melancholic depression, atypical depression, catatonic depression, seasonal affective disorder, or postpartum depression. **Nominal scales** of measurement *categorize things into groups that are qualitatively different from other groups*. Because nominal scales of measurement involve categorizing individuals into qualitatively distinct categories, they yield **qualitative** data. In this case, clinical researchers would interview each person and then decide which type of major depressive disorder each person has. With nominal scales of measurement, it is important to note that the categories are not in any particular order. A diagnosis of melancholic depression is not considered to be "more depressed" than a diagnosis of atypical depression. With all other scales of measurement, the categories are ordered. For example, researchers could also measure depression on an ordinal scale by ranking individual people in terms of the severity of their depression. **Ordinal scales** of measurement also categorize people into different groups but on ordinal scales these groups are rank ordered. In this case, researchers might interview people and diagnose them with a "mild depressive disorder," "moderate depressive disorder," or "severe depressive disorder." An ordinal scale clearly indicates that people *differ in the amount of something they possess*. Thus, someone who was diagnosed with mild depressive disorder would be less depressed than someone diagnosed with moderate depressive disorder. Although ordinal scales rank diagnoses by severity, they do not quantify how much more depressed a moderately depressed person is relative to a mildly depressed person. To make statements about how much more depressed one person is than another, an interval or ratio measurement scale is required. Researchers could measure depression on an interval scale by having people complete a multiple-choice questionnaire that is designed to yield a score reflecting the amount of depression each person has. **Interval scales** of measurement *quantify how much of something people have*. While the ordinal scale indicates that some people have more or less of something than other people, the interval scale is more precise, and indicates *how much* of something someone has. Another way to think about this is that for interval scales the intervals between categories are equivalent whereas for ordinal scales the intervals are not equivalent. For example, on an ordinal scale, the interval (or distance) between a mild depressive disorder and a moderate depressive disorder may not be the same as the interval between a moderate depressive disorder and a severe depressive disorder. However, on an interval scale, the distances between values are equivalent. For example, if people completed a well-designed survey instrument that yielded a score between 1 and 50, the difference in the amount of depression between scores 21 and 22 would be the same as the difference in the amount of depression between scores 41 and 42. Most questionnaires used for research purposes yield scores that are measured on an interval scale of measurement. **Ratio scales** of measurement also *involve quantifying how much of something people have but a score of zero on a ratio scale indicates that the person has none of the thing being measured*. For example, if people are asked how much money they earned last year, the income variable would be measured on a ratio scale: not only are the intervals between values equivalent, but also there is an absolute zero point. A value of zero means the complete absence of income last year. Because they involve quantifying how much of something an individual has, interval and ratio scales yield **quantitative** data. Interval and ratio scales are similar in that they both determine how much of something someone has but some interval scales can yield a negative number while the lowest score possible on a ratio scale is zero. Within the behavioral sciences, the distinction between interval and ratio scales of measurement is not usually very important. Researchers typically use the same statistical procedures to analyze variables measured on interval and ratio scales of measurement.

Although most variables can be easily classified as nominal, ordinal, or interval/ratio, some data are more difficult to classify. Researchers often obtain data by asking participants to answer questions on a survey. These survey responses are then combined into a single measure of the construct.

For example, participants may answer a series of questions related to depression using a Likert scale with response items 1 = *strongly disagree*, 2 = *somewhat disagree*, 3 = *neutral*, 4 = *somewhat agree*, 5 = *strongly agree*. A participant's response to one item on this type of scale is thought of as ordinal but when researchers combine a participant's responses to all of the questions into a single depression score it is typically thought of as an interval/ratio scale of measurement. Another example may help. Suppose you took a test and each item was either right or wrong. Each question would be on an ordinal scale of measurement. If, however, you computed the percent of items you got correct, that total score would be on an interval/ratio scale. The same basic idea applies to Likert scales. The entire scale is treated differently than any single item. Although there is not complete agreement among statisticians on this issue, most researchers classify questionnaire and survey scores as interval (e.g., Carifio & Perla, 2007). Thus, in this book scores on surveys will be considered interval/ratio data.

**Reading Question**

21. Researchers typically treat questionnaire/survey scores as which scale of measurement?
    a. Nominal scale of measurement.
    b. Ordinal scale of measurement.
    c. Interval scale of measurement.

When trying to identify the scale of measurement of a variable, it can also be helpful to think about what each scale of measurement allows you to do. For example, if you can only count the number of things in a given category, you know that you have a nominal scale. Table 1.2 summarizes what you can do with each type of scale and provides examples of each scale of measurement.

**TABLE 1.2 ⬢ The Four Scales of Measurement, What They Allow, and Examples**

| Scale of Measurement | What the Scale Allows You to Do | Examples |
|---|---|---|
| Nominal | COUNT the number of things within different categories. | Favorite pet: 5 dogs, 12 cats, 7 fish, 2 hamsters. |
| | | Marital status: 12 married, 10 divorced, 2 separated. |
| Ordinal | COUNT AND RANK some things as having more of something than others (but DO NOT QUANTIFY how much of it they have). | Annual income: above average, average, or below average. |
| | | Speed (measured by place of finish in a race): 1st, 2nd, 3rd, etc. |
| Interval | COUNT, RANK, AND QUANTIFY how much of something there is but a score of zero does not mean the absence of the thing being measured. | Temperature: −2° F, 98° F, 57° F; 0° F is not the absence of heat. |
| Ratio | COUNT, RANK, AND QUANTIFY how much of something there is and a score of zero means the absence of the thing being measured. | Annual income: $25,048, $48,802, $157,435, etc. |
| | | Number of text messages sent in a day: 0, 3, 351, 15, etc. |

Reading Question

22. The scale of measurement that quantifies the thing being measured (i.e., indicates *how much* of it there is) is _______________ scale(s) of measurement.
    a. the nominal
    b. the ordinal
    c. both the interval and ratio

Reading Question

23. The scale of measurement that categorizes objects into different kinds of things is _______________ scale(s) of measurement.
    a. the nominal
    b. the ordinal
    c. both the interval and ratio

Reading Question

24. The scale of measurement that indicates that some objects have more of something than other objects but not how much more is _______________ scale(s) of measurement.
    a. the nominal
    b. the ordinal
    c. both the interval and ratio

### Discrete vs. Continuous Variables

Variables can also be categorized as discrete or continuous. A **discrete variable** *is measured in whole units rather than fractions of units*. For example, the variable "number of siblings" is a discrete variable because someone can only have a whole number of siblings (e.g., no one can have 2.7 siblings). A **continuous variable** *is measured in fractions of units*. For example, the variable "time to complete a test" is a continuous variable because someone can take a fraction of minutes to complete a test (e.g., 27.39 minutes). Nominal and ordinal variables are always discrete variables. Interval and ratio variables can be either discrete or continuous.

Reading Question

25. If a variable can be measured in fractions of units, it is a _______________ variable.
    a. discrete
    b. continuous

## GRAPHING DATA

The first step in all statistical analyses is graphing because doing so helps you understand your data. For example, if you were looking at the number of siblings that college students have, you could begin by looking at a graph to determine how many siblings most students have. Inspection of the graph also allows you to find out if there is anything odd in the data file that requires further examination. For example, if you graphed the data and found that most people reported having between 0 and 4 siblings but one person reported having 20 siblings, you should probably investigate to determine if that 20 was an error.

Three basic types of graphs are (1) **bar graphs**, (2) **histograms**, and (3) **line graphs**. The names of the first two are a bit misleading because both are created using bars. The only difference between a bar graph and a histogram is that in a bar graph the bars do not touch while in a histogram they do touch. In general, use bar graphs when the data are discrete or qualitative. The spaces between the bars of a bar graph emphasize that there are no possible values between any two categories (i.e., bars). For example, when graphing the number of children in a family, a bar graph is appropriate because there is no possible value between any two categories (e.g., you cannot have 1.5 children). When the data are continuous, use a histogram. For example, if you are graphing the variable "time to complete a test," and you were creating a bar for each minute category, the bars would touch to indicate that the variable we are graphing is continuous (i.e., 27.46 minutes is possible).

**Reading Question**

26. What type of graph is used for discrete data or qualitative data?
    a. A bar graph.
    b. A histogram.

**Reading Question**

27. What type of graph is used for continuous data?
    a. A bar graph.
    b. A histogram.

**Reading Question**

28. In bar graphs, the bars
    a. touch.
    b. don't touch.

**Reading Question**

29. In histograms, the bars
    a. touch.
    b. don't touch.

To create either a bar graph or a histogram, you should put categories on the *x*-axis and the number of scores in a particular category (i.e., the frequency) on the *y*-axis. For example, suppose we asked 19 students how many siblings they have and obtained the following responses:

0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 3, 4, 4, 6

**FIGURE 1.3** ● Bar Graph of Variable, Number of Siblings, Collected From a Sample of 19 Students

To graph these responses, you would list the range of responses to the question "How many siblings do you have?" on the *x*-axis (i.e., in this case 0 through 6). The *y*-axis is the frequency within each category. For each response category, you will draw a bar with a height equal to the number of times that response was given. For example, in the bar graph (Figure 1.3), 4 people said they had 0 siblings and so the bar above the 0 has a height of 4.

**Reading Question**

30. Use the graph to determine how many people said they had 1 sibling.
    a. 4
    b. 5
    c. 6

**FIGURE 1.4** ● Frequency Histogram of Variable, Height in Inches, Collected From a Sample of 13 Volleyball Players

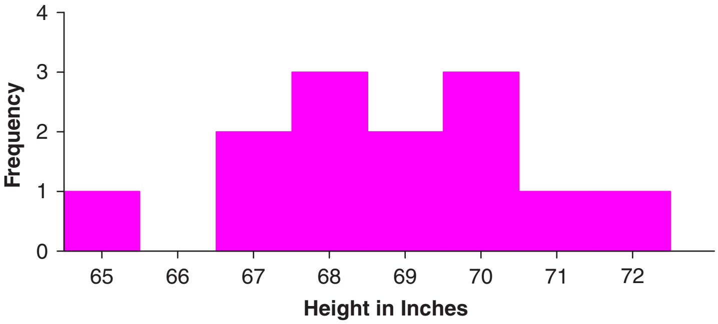
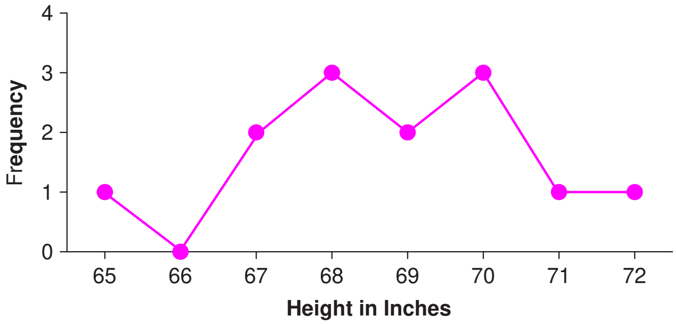



**FIGURE 1.5** ● Frequency Line Graph of Variable, Height in Inches, Collected From a Sample of 13 Volleyball Players




The procedure for creating a histogram is similar to that for creating a bar graph. The only difference is that the bars should touch. For example, suppose that you recorded the height of players on a volleyball team and obtained the following heights rounded to the nearest inch:

65, 67, 67, 68, 68, 68, 69, 69, 70, 70, 70, 71, 72

Height in inches is continuous because there are an infinite number of possible values between any two categories (e.g., between 68 and 69 inches). The data are continuous so we create a histogram (i.e., the bars touch) as shown in Figure 1.4.

Whenever a histogram is appropriate, you may also use a **line graph** in its place. To create a line graph, you use dots to indicate frequencies and connect adjacent dots with lines (Figure 1.5).

Whether the data are discrete or continuous should determine how the data are graphed. You should use a bar graph for discrete data and a histogram or a line graph for continuous data. Nominal data should be graphed with a bar graph. Throughout the text we will use these guidelines, but you should be aware of the fact that histograms and bar graphs are often used interchangeably outside of statistics classes.

**Reading Question**

31. Line graphs can be used whenever a _______________ is appropriate.
    a. histogram
    b. bar graph

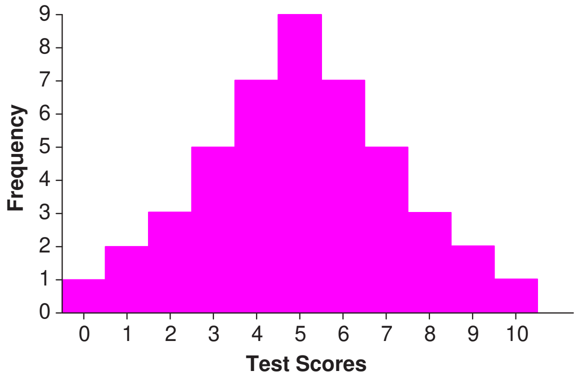
**Reading Question**

32. What type of graph should be used if the data are measured on a nominal scale?
    a. A histogram.
    b. A bar graph.

## SHAPES OF DISTRIBUTIONS

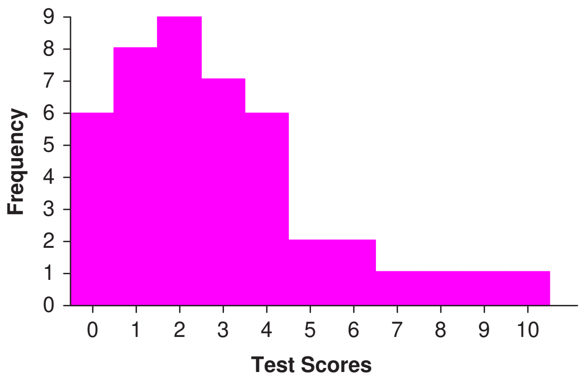
A **distribution** is *a group of scores*. If a distribution is graphed, the resulting bar graph or histogram can have any "shape," but certain shapes occur so frequently that they have specific names. The most common shape you will see is a bell curve (see Figure 1.6). These bell-shaped distributions are also called *normal distributions* or *Gaussian distributions*. One important characteristic of normal distributions is that the most frequent scores pile up in the middle and as you move farther from the middle, the frequency of the scores gets less. Additionally, normal distributions are symmetrical in that the right and left sides of the graph are identical.

FIGURE 1.6 ⬢ Frequency Histogram of Test Scores That Form a Normal Curve



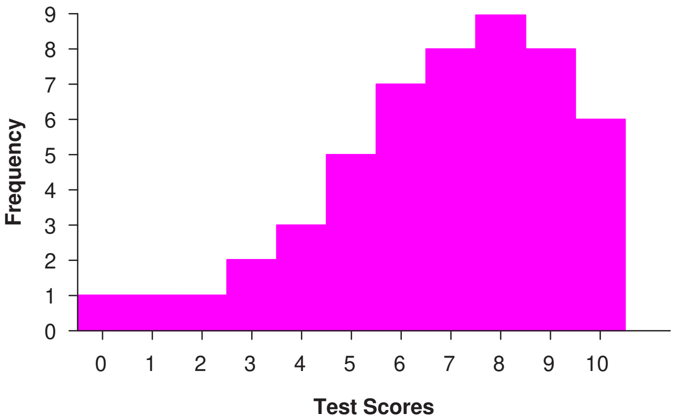

For the purposes of this class, you do not need to know the exact mathematical properties that define the normal curve. However, you should know that a normal curve looks bell shaped and symmetrical. You will use the normal curve frequently in this class.

FIGURE 1.7 ⬢ Positively Skewed Distribution of Scores




The normal curve is important because many variables, when graphed, have a normal shape; this fact will be very important in later chapters. While normal curves are common, there are specific ways for graphs to deviate from a bell shape. Some of these deviations have specific names.

**FIGURE 1.8 ⬢ Negatively Skewed Distribution of Scores**




For example, graphs can deviate from bell shaped because of **skew**. A skewed distribution is asymmetrical, meaning that the right and left sides are not identical. Instead, the scores are shifted such that most of them occur on one side of the scale with fewer scores on the other side of the scale. For example, the distributions in Figures 1.7 and 1.8 are both skewed, but in different ways. The positively skewed distribution (Figure 1.7) has the majority of the scores on the low end of the distribution with just a few scores on the higher end. The negatively skewed distribution is the opposite. Distinguishing between positive and negative skew is as easy as noticing which side of the distribution has the longer "tail" (i.e., which side takes longer to descend from the peak to zero frequency). In positively skewed distributions, the longer tail points toward the right, or the positive side of the *x*-axis. In negatively skewed distributions (Figure 1.8), the longer tail points toward the left, or the negative side of the *x*-axis. There are statistics that you can compute to quantify exactly how skewed a distribution is (see Field, 2017, for an excellent discussion), but we will just eyeball the graphs to determine if they deviate from normal.

**Reading Question**

33. The scores on an exam are distributed such that most scores are low (between 30% and 50%), but a couple of people had very high scores (i.e., above 95%). How is this distribution skewed?
    a. Positively skewed.
    b. Negatively skewed.

Distributions also vary in **kurtosis**, which is the extent to which they have an exaggerated peak versus a flatter appearance. Distributions that have a higher, more-exaggerated peak than a normal curve are called leptokurtic while those that have a flatter peak are called platykurtic. Figures 1.9 and 1.10 display a leptokurtic and platykurtic distribution, respectively. As with skew, there are ways to quantify kurtosis in a distribution (again, see Field, 2017), but we will just eyeball it in this book.

**Reading Question**

34. Distributions that are flatter than a normal distribution are called
    a. platykurtic.
    b. leptokurtic.

FIGURE 1.9 ⬢ Example of a Leptokurtic Distribution

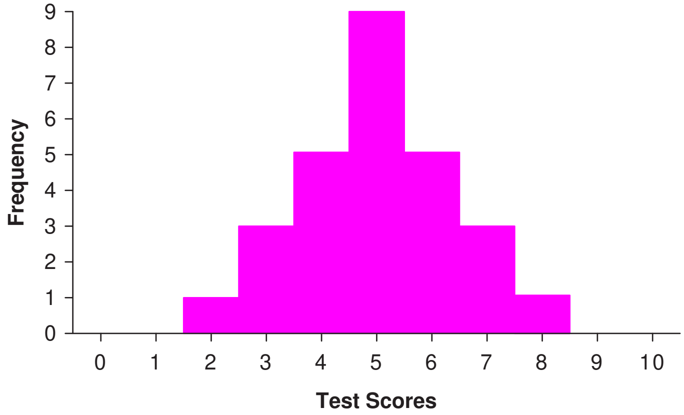
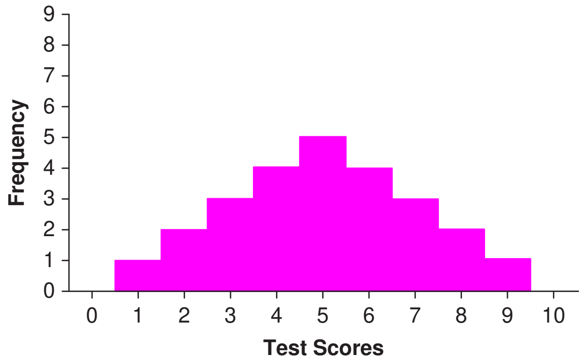



FIGURE 1.10 ⬢ Example of a Platykurtic Distribution




## FREQUENCY DISTRIBUTION TABLES

Graphing data is typically the best way to see patterns in the data (e.g., normal, leptokurtic, or platykurtic). However, some precision is often lost with graphs. Therefore, it is sometimes useful to look at the raw data in a **frequency distribution table**. To create a frequency distribution table, you need to know the measurement categories as well as the number of responses within a given measurement category. For example, suppose that a market researcher asked cell phone users to respond to the following statement: "I am very happy with my cell phone service provider." People were asked to respond with 1 = *strongly agree*, 2 = *agree*, 3 = *neither agree nor disagree*, 4 = *disagree*, 5 = *strongly disagree*. The responses are listed below:

1, 1, 2, 2, 2, 2, 3, 3, 3, 3, 3, 3, 3, 4, 4, 4, 4, 4, 4, 5, 5, 5, 5

It is probably obvious that a string of numbers like that one is not a particularly useful way to present data. A frequency distribution table organizes the data, so it is easier to interpret; one is shown in Table 1.3.

**TABLE 1.3 ⬢ Frequency Distribution Table of the Variable "I Am Very Happy With My Cell Phone Service Provider"**

| | *X* | *f* | *Percent* | *Cumulative Percent* |
|---|---|---|---|---|
| Strongly agree | 1 | 2 | 8.70 | 8.70 |
| Agree | 2 | 4 | 17.39 | 26.09 |
| Neither agree nor disagree | 3 | 7 | 30.43 | 56.52 |
| Disagree | 4 | 6 | 26.09 | 82.61 |
| Strongly disagree | 5 | 4 | 17.39 | 100.00 |

The first column ($X$) represents the possible response categories. People *could* respond with any number between 1 and 5, therefore the $X$ column (i.e., the measurement categories) must include all of the *possible* response values, namely 1 through 5. In this case, we chose to put the categories in ascending order from 1 to 5, but they could also be listed in descending order from 5 to 1.

The next column ($f$) is where you record the frequency of each response. For example, 4 people gave responses of 5 (*strongly disagree*) and so a 4 is written in the "$f$" column across from the response category of 5 (*strongly disagree*).

The next column (percent) is simply the number of responses in each category (i.e., the frequency $f$) divided by the total number of responses (i.e., $N$ = 23). For example, 4 people responded with "strongly agree" so 4/23 = 17.39%. The percent column is a useful way to compare the relative number of responses in each category.

The final column is the cumulative percent or percentile. A percentile score is the percent of respondents with a score equal to or less than a given score. For example, in Table 1.3 26.09% of the scores are equal to or lower than the score of 2 "Agree." So, the score of 2 is at the 26.09 percentile. In this example, this means that 26.09% of the sample agree or strongly agree with the statement, "I am happy with my cell phone provider."

**Reading Question**

35. The value for "$f$" represents the
    a. number of measurement categories.
    b. number of responses within a given measurement category.

**Reading Question**

36. In Table 1.3 how many people responded with an answer of 3?
    a. 2
    b. 4
    c. 7

Reading Question

37. What score is at the 56.52 percentile?
   a. 2
   b. 3
   c. 4

Reading Question

38. What percent of the sample disagreed or strongly disagreed with the statement "I am happy with my cell phone provider"?
   a. 82.61%
   b. 26.09%
   c. 43.48%

## How Obedient Were Milgram's Participants?

You are probably somewhat familiar with Stanley Milgram's (1974) famous studies on obedience to authority. If you have never seen a video of one of his studies, it would be helpful to watch one that demonstrates the procedures. There is a link to a video on the textbook website, but you can find videos online by searching for "Milgram Experiment." What you may not know is that Milgram conducted many studies using the same basic procedures but varying different elements of the study. In the first study, men between the ages of 20 and 50 were recruited to participate in a one-hour study of memory at Yale. When each participant arrived at the laboratory, he met an Experimenter as well as another participant. In fact, both the Experimenter and the other participant were confederates hired by Milgram to play specific roles in the study. The Experimenter said the study's purpose was testing a theory of learning suggesting that people learn information more effectively when they receive punishment for their memory error. To lend credence to this cover story, the Experimenter showed the participants a book that presumably contained this theory and he stated that the theory suggested that parents could help children learn to behave by spanking them. He continued the cover story by "explaining" that it is unclear based on current scientific studies how much punishment is necessary for the most effective learning. The Experimenter said one participant would serve as the Teacher in the study and one participant would serve as the Learner. He asked the participants if they preferred being the Learner or the Teacher. They were both allowed to express their preferences, but ultimately the Experimenter decided that the only fair way was to draw from a hat. Of course, the draw was rigged so that the real participant was always the Teacher and the confederate was always the Learner. Once the roles were assigned, the Learner/confederate was strapped into a chair and an electrode was placed on his wrist. The participant was then led to another room and the learning task was explained to him. In this task, the Teacher read a set of word pairs. Later there was a testing session where the Learner indicated which words were paired together by

## ACTIVITY FIGURE 1.1 ⬢ Shock Labels in Milgram's Studies

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 15 VOLTS | 30 | 45 | 60 | 75 VOLTS | 90 | 105 | 120 | 135 VOLTS | 150 | 165 | 180 | 195 VOLTS | 210 | 225 |
| SLIGHT SHOCK | .... | .... | .... | MODERATE SHOCK | .... | .... | .... | STRONG SHOCK | .... | .... | .... | VERY STRONG SHOCK | .... | .... |

| 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 240 | 255 VOLTS | 270 | 285 | 300 | 315 VOLTS | 330 | 345 | 360 | 375 VOLTS | 390 | 405 | 420 | 435 VOLTS | 450 VOLTS |
| .... | INTENSE SHOCK | .... | .... | .... | EXTEME INTENSITY SHOCK | .... | .... | .... | DANGER SEVERE SHOCK | .... | .... | .... | X X X | |