



Fundamental Statistics for the Behavioral Sciences

EIGHTH EDITION



David C. Howell

List of Applications

Subject-Matter Area	Reference	Pages
Morphine Tolerance and Context	Siegel (1975)	4
Dextromethorphan and heroin	Mann-Jones, Ettinger, Baisden, & Baisdin (2003)	5
Social Desirability and Eating	Pliner & Chaiken (1990)	33
Moon Illusion	Kaufman & Rock (1962)	33, 97
Cigarette Consumption and Health	Landwehr & Watkins (1987)	35
Perception of Rotated Images	Krantz (unpublished)	37, 96
Video Game Playing	Gentile (2009)	46
Test-Taking Skills	Katz, Lautenschlager, Blackburn, & Harris (1990)	56, 68, 74, 222, 317
Immune Response	Cohen et al. (1992)	58
Race and Retribution	Rogers & Prentice-Dunn (1981)	58
Down Syndrome	Moran (1974)	59
Birth Month and Gender	Fombonne (1989)	60
Grades and Attendance	Howell (unpublished)	76
Optimism and Performance	Seligman et al. (1990)	76
Perceived Attractiveness	Langlois & Roggman (1990)	80
Attention Deficit Disorder	Howell & Huessy (1985)	57, 149, 517
Treatment of Anorexia	Everitt (1994)	105, 326, 329, 338, 351
Behavior Problems	Achenbach (1991)	110
M&M's	Unpublished	131
Death Penalty	Dieter (1998)	138, 142
Maternal Age at First Birth	CDC (2003)	144
Don't Litter	Geller et al. (1976)	146, 154, 511
Alcohol, Drugs, and Driving	Fell (1995)	149, 526
Height and Status	Duguid & Goncalo (2012)	152, 177
Sunk-Cost Fallacy	Strough et al. (2008)	154
Finger Tapping	Christianson & Leathem (2004)	164
Multiple Estimation	Vul & Pashler (2008)	182, 332
Reward Delayed	Magen, Dweck, & Gross (2008)	182
Infant Mortality and Physicians	St. Leger et al. (1978)	188
Health Care Expenditures	Cochrane et al. (1978)	188, 204
Breast Cancer and Sunshine	<i>Newsweek</i> (1991)	189
Red Wine and Heart Disease	Wong (2008)	192
Pace of Life	Levine (1990)	192
Weight, Height, and Gender	Ryan et al. (1985)	200
Smoking and Alcohol Consumption	British Government	201, 243
Educational Expenditures	Guber (1999)	208, 267
Course Quality and Grades	Unpublished	211, 216
Births in Sub-Saharan Africa	Guttmacher Institute (2002)	220
Down Syndrome and Age	Geyer (1991)	222
Stress and Mental Health	Wagner et al. (1988)	227
The Me Generation—Not	Trzesniewski et al. (2009)	233, 258
The Me Generation	Twenge (2006)	233
Regression to the Mean	Galton (1886)	236

Guns Don't Reduce Murder Rate	Grambsch (2009)	237
School Funding	Guber (1999)	208, 267
Maternal Confidence	Leerkes & Crockenberg (1999)	282
Psychology and Cancer Patients	Malcarne et al. (1995)	285
Optimism and Religion	Sethi & Seligman (1993)	290
Family Structure and Vulnerability	Mireault (1990)	292
Sexual harassment	Brooks & Perot (1991)	292
Stress and Behavior Problems	Williamson (2008)	297, 303
Children and the "Lie Scale"	Compas et al. (1994)	307
The Crowd Within	Vul & Pashler (2008)	182, 332
Marital Satisfaction with Sex	Hout et al. (1987)	341
Beta-Endorphins and Stress	Hoaglin et al. (1983)	342
Sources of Homophobia	Adams et al. (1996)	355
Lucky Charms	Damisch, Stoberock, & Mussweiler (2010)	361
TAT and Schizophrenia	Werner, Stabenau, & Pollin (1970)	369
Age and Memory	Eysenck (1974)	435, 444, 482
Gender, Anger, and Perception	Briscoll & Ullman (2008)	370
Stereotype Threat	Aronson et al. (1998)	375
Alcohol and Aggression	Giancola & Corman (2007)	397
Low-Birth-Weight Infants	Nurcombe et al. (1984)	427, 467
Marijuana and Behavior	Conti & Musty (1984)	429
Therapy for Rape Victims	Foa et al. (1991)	436
Insecurity in a Crowd	Darley and Latané	436, 500
Smoking and Performance	Spilich et al. (1992)	436
Attractiveness and Facial Features	Langlois & Roggman (1990)	438
Masculine Overcompensation Thesis	Willer (2005)	458
Sexual Identification	Liddle (1997)	469
Earthquakes and Depression	Nolen-Hoeksema & Morrow (1991)	471
Migraines	Blanchard et al. (1978)	485
Therapeutic Touch	Rosa et al. (1996)	490
Anorexia and Prozac	Walsh et al. (2006)	496
Helping Behavior and Gender	Latané & Dabbs (1975)	502
Fertility and Smoking	Weinberg & Gladen (1986)	516
Race and Desired Weight	Gross (1985)	517
Health and Inescapable Shock	Visintainer et al. (1982)	518
Race and Racial Identification	Clark & Clark (1947); Hraba & Grant (1970)	516
Testosterone and Antisocial Behavior	Dabbs & Morris (1990)	518
Race and Death Penalty Again	Peterson (2001)	519
Schizophrenia and Subcortical Structures	Suddath et al. (1990)	531
He, She, They	Foertsch & Gernsbacher (1997)	538
Learning Diagnostics	Nurcombe & Fitzhenry-Coor (1979)	544
Intervention for Depression	Horowitz & Garber (2006)	553
Nicotine Gum and Smoking	Stead et al. (2008)	559
Behavioral Activation	Mazzucchelli, Kane, & Rees (2010)	563
Attention Deficit Disorder	Bloch et al. (2009)	563
Lithium for Depression	Bauer & Döpfmer (1999)	564
Myeloma Treatment	Kapoor et al. (2010)	564
Cognitive Behavior Therapy and PTSD	Bisson & Andrew (2007)	565

Fundamental Statistics for the Behavioral Sciences

This is an electronic version of the print textbook. Due to electronic rights restrictions, some third party content may be suppressed. Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. The publisher reserves the right to remove content from this title at any time if subsequent rights restrictions require it. For valuable information on pricing, previous editions, changes to current editions, and alternate formats, please visit www.cengage.com/highered to search by ISBN#, author, title, or keyword for materials in your areas of interest.

Fundamental Statistics **for the Behavioral Sciences**

EIGHTH EDITION



David C. Howell

University of Vermont



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

Fundamental Statistics for the Behavioral Sciences, Eighth Edition**David C. Howell**

Publisher: Jon-David Hague

Acquisitions Editor: Tim Matray

Assistant Editor: Paige Leeds

Editorial Assistant: Nicole Richards

Associate Media Editor: Jasmin Tokatlian

Senior Brand Manager: Elisabeth Rhoden

Market Development Manager: Christine Sosa

Manufacturing Planner: Karen Hunt

Rights Acquisitions Specialist: Roberta Broyer

Design, Production Services, and
Composition: PreMediaGlobalCover Image: Creative Wood Background
© Imageman/shutterstock

© 2014, 2011 Wadsworth, Cengage Learning

ALL RIGHTS RESERVED. No part of this work covered by the copyright herein may be reproduced, transmitted, stored, or used in any form or by any means graphic, electronic, or mechanical, including but not limited to photocopying, recording, scanning, digitizing, taping, Web distribution, information networks, or information storage and retrieval systems, except as permitted under Section 107 or 108 of the 1976 United States Copyright Act, without the prior written permission of the publisher.

For product information and technology assistance, contact us at
Cengage Learning Customer & Sales Support, 1-800-354-9706

For permission to use material from this text or product,
submit all requests online at **www.cengage.com/permissions**.

Further permissions questions can be e-mailed to
permissionrequest@cengage.com

Library of Congress Control Number: 2012950801

Student Edition:

ISBN-13: 978-1-285-07691-1

ISBN-10: 1-285-07691-5

Loose-leaf Edition:

ISBN-13: 978-1-285-08180-9

ISBN-10: 1-285-08180-3

Wadsworth20 Davis Drive
Belmont, CA 94002-3098
USA

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

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

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

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

Printed in the United States of America
1 2 3 4 5 6 7 17 16 15 14 13

*Dedication: To my wife, Donna, who has tolerated
“I can't do that now, I am working on my book”
for far too long.*

Brief Contents

Preface xi

Chapter 1

Introduction 1

Chapter 2

Basic Concepts 18

Chapter 3

Displaying Data 36

Chapter 4

Measures of Central Tendency 62

Chapter 5

Measures of Variability 78

Chapter 6

The Normal Distribution 107

Chapter 7

Basic Concepts of Probability 129

Chapter 8

Sampling Distributions and Hypothesis Testing 151

Chapter 9

Correlation 184

Chapter 10

Regression 225

Chapter 11

Multiple Regression 262

Chapter 12

Hypothesis Tests Applied to Means:
One Sample 295

Chapter 13

Hypothesis Tests Applied to Means:
Two Related Samples 327

Chapter 14

Hypothesis Tests Applied to Means:
Two Independent Samples 344

Chapter 15

Power 372

Chapter 16

One-Way Analysis of Variance 396

Chapter 17

Factorial Analysis of Variance 440

Chapter 18

Repeated-Measures Analysis of Variance 470

Chapter 19

Chi-Square 488

Chapter 20

Nonparametric and Distribution-Free
Statistical Tests 520

Chapter 21

Meta-Analysis 547

Appendix A

Arithmetic Review 566

Appendix B

Symbols and Notation 573

Appendix C

Basic Statistical Formulae 576

Appendix D

Data Set 580

Appendix E

Statistical Tables 584

Glossary 602

References 608

Answers to Exercises 615

Index 637

Contents

Preface xi

Chapter 1

Introduction 1

- 1.1 A Changing Field 3
- 1.2 The Importance of Context 4
- 1.3 Basic Terminology 6
- 1.4 Selection among Statistical Procedures 10
- 1.5 Using Computers 12
- 1.6 Summary 14
- 1.7 A Quick Review 15
- 1.8 Exercises 16

Chapter 2

Basic Concepts 18

- 2.1 Scales of Measurement 19
- 2.2 Variables 25
- 2.3 Random Sampling 26
- 2.4 Notation 28
- 2.5 Summary 30
- 2.6 A Quick Review 31
- 2.7 Exercises 32

Chapter 3

Displaying Data 36

- 3.1 Plotting Data 38
- 3.2 Stem-and-Leaf Displays 41
- 3.3 Reading Graphs 46
- 3.4 Alternative Methods of Plotting Data 48
- 3.5 Describing Distributions 51
- 3.6 Using Computer Programs to Display Data 53
- 3.7 Summary 54
- 3.8 A Quick Review 55
- 3.9 Exercises 56

Chapter 4

Measures of Central Tendency 62

- 4.1 The Mode 63
- 4.2 The Median 63

- 4.3 The Mean 64
- 4.4 Relative Advantages and Disadvantages of the Mode, the Median, and the Mean 65
- 4.5 Obtaining Measures of Central Tendency Using SPSS 68
- 4.6 A Simple Demonstration—Seeing Statistics 70
- 4.7 Summary 73
- 4.8 A Quick Review 74
- 4.9 Exercises 74

Chapter 5

Measures of Variability 78

- 5.1 Range 81
- 5.2 Interquartile Range and Other Range Statistics 82
- 5.3 The Average Deviation 83
- 5.4 The Variance 84
- 5.5 The Standard Deviation 86
- 5.6 Computational Formulae for the Variance and the Standard Deviation 87
- 5.7 The Mean and the Variance as Estimators 88
- 5.8 Boxplots: Graphical Representations of Dispersion and Extreme Scores 90
- 5.9 A Return to Trimming 94
- 5.10 Obtaining Measures of Dispersion Using SPSS 96
- 5.11 The Moon Illusion 97
- 5.12 Seeing Statistics 100
- 5.13 Summary 101
- 5.14 A Quick Review 103
- 5.15 Exercises 103

Chapter 6

The Normal Distribution 107

- 6.1 The Normal Distribution 110
- 6.2 The Standard Normal Distribution 114
- 6.3 Setting Probable Limits on an Observation 120
- 6.4 Measures Related to z 122
- 6.5 Seeing Statistics 123
- 6.6 Summary 124
- 6.7 A Quick Review 125
- 6.8 Exercises 125

Chapter 7

Basic Concepts of Probability 129

- 7.1 Probability 130
- 7.2 Basic Terminology and Rules 133
- 7.3 The Application of Probability to Controversial Issues 138
- 7.4 Writing Up the Results 141
- 7.5 Discrete Versus Continuous Variables 142
- 7.6 Probability Distributions for Discrete Variables 143
- 7.7 Probability Distributions for Continuous Variables 144
- 7.8 Summary 146
- 7.9 A Quick Review 148
- 7.10 Exercises 148

Chapter 8

Sampling Distributions and Hypothesis Testing 151

- 8.1 Sampling Distributions and the Standard Error 152
- 8.2 Two More Examples Involving Course Evaluations and Human Decision Making 154
- 8.3 Hypothesis Testing 157
- 8.4 The Null Hypothesis 160
- 8.5 Test Statistics and Their Sampling Distributions 162
- 8.6 Using the Normal Distribution to Test Hypotheses 163
- 8.7 Type I and Type II Errors 168
- 8.8 One- and Two-Tailed Tests 172
- 8.9 Seeing Statistics 176
- 8.10 A Final Example 177
- 8.11 Back to Course Evaluations and Decision Making 179
- 8.12 Summary 179
- 8.13 A Quick Review 180
- 8.14 Exercises 181

Chapter 9

Correlation 184

- 9.1 Scatter Diagrams 185
- 9.2 An Example: The Relationship Between the Pace of Life and Heart Disease 192
- 9.3 The Covariance 193
- 9.4 The Pearson Product-Moment Correlation Coefficient (r) 194
- 9.5 Correlations with Ranked Data 196
- 9.6 Factors That Affect the Correlation 198

- 9.7 Beware Extreme Observations 201
- 9.8 Correlation and Causation 203
- 9.9 If Something Looks Too Good to Be True, Perhaps It Is 204
- 9.10 Testing the Significance of a Correlation Coefficient 205
- 9.11 Intercorrelation Matrices 208
- 9.12 Other Correlation Coefficients 210
- 9.13 Using SPSS to Obtain Correlation Coefficients 213
- 9.14 Seeing Statistics 213
- 9.15 Does Rated Course Quality Relate to Expected Grade? 216
- 9.16 Summary 219
- 9.17 A Quick Review 220
- 9.18 Exercises 221

Chapter 10

Regression 225

- 10.1 The Relationship Between Stress and Health 227
- 10.2 The Basic Data 229
- 10.3 The Regression Line 229
- 10.4 The Accuracy of Prediction 238
- 10.5 The Influence of Extreme Values 243
- 10.6 Hypothesis Testing in Regression 244
- 10.7 Computer Solution Using SPSS 246
- 10.8 Seeing Statistics 248
- 10.9 Course Ratings as a Function of Anticipated Grade 253
- 10.10 Regression Versus Correlation 254
- 10.11 Summary 255
- 10.12 A Quick Review 256
- 10.13 Exercises 257

Chapter 11

Multiple Regression 262

- 11.1 Overview 264
- 11.2 Funding Our Schools 267
- 11.3 Residuals 278
- 11.4 Hypothesis Testing 279
- 11.5 Refining the Regression Equation 281
- 11.6 A Second Example: What Makes a Confident Mother? 282
- 11.7 A Third Example: Psychological Symptoms in Cancer Patients 285
- 11.8 Summary 288
- 11.9 A Quick Review 289
- 11.10 Exercises 290

Chapter 12

Hypothesis Tests Applied to Means: One Sample 295

- 12.1 Sampling Distribution of the Mean 297
- 12.2 Testing Hypotheses about Means When σ Is Known 300
- 12.3 Testing a Sample Mean When σ Is Unknown 304
- 12.4 Factors That Affect the Magnitude of t and the Decision about H_0 310
- 12.5 A Second Example: The Moon Illusion 310
- 12.6 How Large Is Our Effect? 311
- 12.7 Confidence Limits on the Mean 312
- 12.8 Using SPSS to Run One-Sample t Tests 316
- 12.9 A Good Guess Is Better than Leaving It Blank 317
- 12.10 Seeing Statistics 320
- 12.11 Summary 323
- 12.12 A Quick Review 324
- 12.13 Exercises 325

Chapter 13

Hypothesis Tests Applied to Means: Two Related Samples 327

- 13.1 Related Samples 328
- 13.2 Student's t Applied to Difference Scores 329
- 13.3 The Crowd Within Is Like the Crowd Without 332
- 13.4 Advantages and Disadvantages of Using Related Samples 334
- 13.5 How Large an Effect Have We Found? 335
- 13.6 Confidence Limits on Changes 337
- 13.7 Using SPSS for t Tests on Related Samples 338
- 13.8 Writing Up the Results 338
- 13.9 Summary 339
- 13.10 A Quick Review 340
- 13.11 Exercises 341

Chapter 14

Hypothesis Tests Applied to Means: Two Independent Samples 344

- 14.1 Distribution of Differences Between Means 345
- 14.2 Heterogeneity of Variance 353
- 14.3 Nonnormality of Distributions 355
- 14.4 A Second Example with Two Independent Samples 355
- 14.5 Effect Size Again 357
- 14.6 Confidence Limits on $\mu_1 - \mu_2$ 358
- 14.7 Plotting the Results 359
- 14.8 Writing Up the Results 360

- 14.9 Use of Computer Programs for Analysis of Two Independent Sample Means 361

- 14.10 Do Lucky Charms Work? 361

- 14.11 Seeing Statistics 366

- 14.12 Summary 367

- 14.13 A Quick Review 368

- 14.14 Exercises 369

Chapter 15

Power 372

- 15.1 The Basic Concept of Power 375

- 15.2 Factors Affecting the Power of a Test 377

- 15.3 Calculating Power the Traditional Way 380

- 15.4 Power Calculations for the One-Sample t Test 382

- 15.5 Power Calculations for Differences Between Two Independent Means 385

- 15.6 Power Calculations for the t Test for Related Samples 388

- 15.7 Power Considerations in Terms of Sample Size 389

- 15.8 You Don't Have to Do It by Hand 390

- 15.9 Post-hoc (Retrospective) Power 391

- 15.10 Summary 392

- 15.11 A Quick Review 393

- 15.12 Exercises 393

Chapter 16

One-Way Analysis of Variance 396

- 16.1 The General Approach 397

- 16.2 The Logic of the Analysis of Variance 401

- 16.3 Calculations for the Analysis of Variance 406

- 16.4 Unequal Sample Sizes 413

- 16.5 Multiple Comparison Procedures 415

- 16.6 Violations of Assumptions 424

- 16.7 The Size of the Effects 424

- 16.8 Writing Up the Results 427

- 16.9 The Use of SPSS for a One-Way Analysis of Variance 428

- 16.10 A Final Worked Example 429

- 16.11 Seeing Statistics 432

- 16.12 Summary 433

- 16.13 A Quick Review 434

- 16.14 Exercises 435

Chapter 17

Factorial Analysis of Variance 440

- 17.1 Factorial Designs 441

- 17.2 The Eysenck Study 444

- 17.3 Interactions 449

- 17.4 Simple Effects 451

- 17.5 Measures of Association and Effect Size 453
- 17.6 Reporting the Results 456
- 17.7 Unequal Sample Sizes 457
- 17.8 Masculine Overcompensation Thesis: It's a Male Thing 458
- 17.9 Using SPSS for Factorial Analysis of Variance 461
- 17.10 Seeing Statistics 462
- 17.11 Summary 463
- 17.12 A Quick Review 464
- 17.13 Exercises 465

Chapter 18

Repeated-Measures Analysis of Variance 470

- 18.1 An Example: Depression as a Response to an Earthquake 471
- 18.2 Multiple Comparisons 474
- 18.3 Effect Size 476
- 18.4 Assumptions Involved in Repeated-Measures Designs 477
- 18.5 Advantages and Disadvantages of Repeated-Measures Designs 477
- 18.6 Using SPSS to Analyze Data in a Repeated-Measures Design 478
- 18.7 Writing Up the Results 481
- 18.8 A Final Worked Example 482
- 18.9 Summary 484
- 18.10 A Quick Review 485
- 18.11 Exercises 485

Chapter 19

Chi-Square 488

- 19.1 One Classification Variable: The Chi-Square Goodness-of-Fit Test 490
- 19.2 Two Classification Variables: Analysis of Contingency Tables 496
- 19.3 Possible Improvements on Standard Chi-Square 498
- 19.4 Chi-Square for Larger Contingency Tables 500
- 19.5 The Problem of Small Expected Frequencies 501
- 19.6 The Use of Chi-Square as a Test on Proportions 502
- 19.7 SPSS Analysis of Contingency Tables 504
- 19.8 Measures of Effect Size 506
- 19.9 A Final Worked Example 511
- 19.10 A Second Example of Writing Up Results 513
- 19.11 Seeing Statistics 513
- 19.12 Summary 514
- 19.13 A Quick Review 515
- 19.14 Exercises 516

Chapter 20

Nonparametric and Distribution-Free Statistical Tests 520

- 20.1 The Mann–Whitney Test 524
- 20.2 Wilcoxon's Matched-Pairs Signed-Ranks Test 531
- 20.3 Kruskal–Wallis One-Way Analysis of Variance 536
- 20.4 Friedman's Rank Test for k Correlated Samples 538
- 20.5 Measures of Effect Size 540
- 20.6 Writing Up the Results 540
- 20.7 Summary 541
- 20.8 A Quick Review 542
- 20.9 Exercises 543

Chapter 21

Meta-Analysis 547

- Meta-Analysis 548
- 21.1 A Brief Review of Effect Size Measures 549
- 21.2 An Example—Child and Adolescent Depression 553
- 21.3 A Second Example—Nicotine Gum and Smoking Cessation 559
- 21.4 A Quick Review 562
- 21.5 Exercises 563

Appendix A

Arithmetic Review 566

Appendix B

Symbols and Notation 573

Appendix C

Basic Statistical Formulae 576

Appendix D

Data Set 580

Appendix E

Statistical Tables 584

Glossary 602

References 608

Answers to Exercises 615

Index 637

Preface

Why Statistics?

Those of us who teach in this area hate to admit it, but statistics is seldom listed as the most sought-after course on campus. A high percentage of students enroll because their department has made this a required course. Under these conditions students have a right to ask “Why?” and there are at least two good answers to that question. The traditional answer is that we want our students to learn a specific set of skills about data analysis (including formulae and procedures) so that they can understand the experimental literature and conduct analyses on their own data. The broader answer, and one that applies to perhaps a larger number of students, is that some more general facility with numbers and data in general is an important skill that has lifelong and career-related value. Most of us, and not only those who do experimental work, frequently come across numerical data as part of our jobs, and some broad understanding of how to deal with those data is an important and marketable skill. It is my experience that students who have taken a course in statistics, even if they think that they have forgotten every technique they ever learned, have an understanding of numerical data that puts them ahead of their colleagues. And in a world increasingly dominated by quantitative data, that skill is more and more in demand. Former students have told me that they were assigned an important task because they were the only one in their office who wasn’t afraid of data. In fact, that was originally why I ended up teaching the statistics courses when I was first hired.

Statistics is not really about numbers; it is about understanding our world. Certainly an important activity for statisticians is to answer such questions as whether cocaine taken in a novel context has more of an effect than cocaine taken in a familiar context. But let’s not forget that what we are talking about here is drug addiction or the effect of the environment on learning and memory. The results of our experiment have a life beyond the somewhat limited world of the cognitive or social scientist. And let’s also remember that the numbers that most people see do not relate to tightly controlled experiments, but to the implications of a traffic study for the development of a shopping center, the density of residential housing and its impact on the local school budget, and a marketing survey for a new product. All of these examples involve many of the basic statistical concepts covered in this book.

Why This Text?

Enough preaching on the value of a course in statistics. Presumably the instructor was convinced before he or she started reading, and I hope that students have become at least a bit more open minded. But the question remains, why should you use this book instead of another of the many available texts? Part of the answer comes down to the matter of style. I have deliberately set out to make this book both interesting and useful for students and instructors. It is written in an informal style; every example is put in the context of an investigation that one might reasonably conduct; and almost all of the examples are taken from the published literature. It does not make much sense to ask people to learn a series of statistical procedures without supplying examples of situations in which those techniques would actually be applied.

This text is designed for an introductory statistics course in psychology, education, and other behavioral sciences. It does not presuppose a background in mathematics beyond high school algebra, and it emphasizes the *logic* of statistical procedures rather than their derivation.

Over the past 20 years the world of data analysis has changed dramatically. Whereas we once sat down with a calculator and entered data by hand to solve equations, we are now much more likely to use a statistical package running on a desktop computer. In fact, for some purposes we are likely to be using an online program written in Java or some similar language that we download free from the Internet. (I sometimes use an app downloaded to my iPhone.) As the mechanics of doing statistics have changed, so too must our approach to teaching statistical procedures. While we cannot, and should not, forego all reference to formulae and computations, it is time that we relaxed our emphasis on them. And by relaxing the emphasis on computation, we free up the time to increase the emphasis on interpretation. That is what this book tries to do. It moves away from simply declaring group differences to be significant or not significant toward an explanation of what such differences mean relative to the purpose behind the experiment. I like to think of it as moving toward an analysis of *data* and away from an analysis of numbers. It becomes less important to concentrate on whether there is a difference between two groups than to understand what that difference means.

In the process of moving away from a calculator toward a computer, I have altered my approach to formulae. In the past I often gave a definitional formula, but then immediately jumped to a computational one. But if I have to worry less about computation, and more about understanding, then I am able to revert to the use of definitional formulae. It is my hope that this will make students' lives a bit easier.

Unique Features

Several features of this book set it apart from other books written for the same audience. One of these was just noted: the use of examples from the research literature. I have attempted to choose studies that address problems of interest to students. Examples include the effect of context on heroin overdose, the relationship between daily stress and psychological symptoms, variables influencing

course evaluations, the effect of early parental death on children's feelings of vulnerability, and variables controlling memory changes as a function of age. I want students to have some involvement in the questions being asked, and I want to illustrate that statistical analyses involve more than just applying a few equations.

In most chapters a section is devoted to an example using SPSS. Readers have suggested that I concentrate most on SPSS, which is probably the most widely used package, and I have done so in this edition. My purpose is to familiarize students with the form of computer printouts and the kinds of information they contain. I am not trying to teach students how to use a particular statistical package, but I want them to get a feel for what is available. But if students are going to be using SPSS, I would hate to have them buy an SPSS manual just to do their work. I have two SPSS manuals on the Web and encourage students to go to them. They are not as complete as a printed book would be, but they are more than sufficient to allow students to work with SPSS. I recommend the shorter manual, but the longer one is there if additional information is needed.

In this edition I have gone a step further and suggested to students that they download and use either of two free statistical packages that are nearly as powerful as SPSS and related packages, while being somewhat easier to use when working through the text. The changes in available software over the past few years have been huge, and we should take advantage of them.

Data files for all of the examples and exercises used in the text are available on a website that I maintain for this book. The basic URL for that site is www.uvm.edu/~dhowell/fundamentals8/index.html. A link at that site will take you to the data. These files are formatted in ASCII, so that they can be read by virtually any statistical program. The variable names appear on the first line and can be directly imported to your software. The data can be saved to your computer simply by selecting your browser's Save option. The availability of these files makes it easy for students and instructors to incorporate any statistical package with the text.

A Student Handbook is also available at the website referred to above. It provides complete solutions for half the exercises. I have included answers only to the odd numbered questions because many instructors prefer to assign problems (or exam questions) on material that does not have an answer at the back of the book or the *Student Solution Handbook*. (I am very much aware that this does annoy students, from whom I sometimes receive unhappy mail messages, but it is a balance between the needs of students and the desires of the instructors.) In the past I have made available to instructors the answers to all of the questions (and I still do). Those answers frequently come with comments such as "In class you might point out . . ." or "The reason why I asked this question is to get at . . ." As I read through them in creating this edition, I realized that many, though not all, of those comments would also be useful to students. So I have included all of them in the Student Handbook as well. Some of them may appear unhelpful or out of context, but I think most of them are worth reading.

On my Web pages I have also included many links to other sites, where you can find good examples, small programs to demonstrate statistical techniques, a more extensive glossary, and so on. People have devoted a great deal of time to making material available over the Internet, and it is very worthwhile to use that material.

Why a New Edition?

When an author comes out with a new edition, I think that it is fair to ask what was wrong with the old one, other than the fact that it is widely available in the used book market. Normally I design a new edition to incorporate changes that are going on in the field and to remove things that are no longer needed. And, despite what many people think, there is a lot of new work going on. But in this edition and the previous one I have taken a different approach. While I have added some new material, the major effort has been to read the book as a new student would, and try to find ways to clarify and repeat concepts. For example, I know that the Y axis is the vertical one, but most people don't, and telling them once is not enough. So I often write something like "On the Y (vertical) axis . . ." And when you start looking at a book that way, you find many places for clarification—especially if you have a wife who has spent most of her life in secondary education and knows more about pedagogy than I do. (She actually read every chapter and made many fruitful suggestions.) I have also begun each chapter with a list of concepts that will be important in the chapter, in hopes that if you aren't sure what they are you will review them. Where necessary I have inserted important comments in boxes to pull several points together, to highlight material that you really need to understand, or to clarify difficult concepts. I have also inserted short biographies of important statisticians. Especially in the first half of the 20th century there were many interesting (and cantankerous) people in the field and they are worth meeting. Next, I have removed the very brief and weak chapter summaries and replaced them with much more complete ones. My goal was to condense the chapter into a few paragraphs, and you will do well to spend some time on them. A while back I was reading a programming text on Java and came across an author who inserted very simple questions with answers at the end of each chapter. I discovered that I learned a lot from those simple questions, so I have followed his lead in this edition. The questions are intended to focus your attention on many of the important points in the chapter. I hope that they are useful.

An important feature of this book is the continued increase in emphasis on measures of effect size. This is in line with trends in the field, but it is also important because it causes the student, and the researcher, to think carefully about what a result means. In presenting effect size measures I have tried to convey the idea that the writer is trying to tell the reader what the study found, and there are different ways of accomplishing that goal. In some situations it is sufficient to simply talk about the difference between means or proportions. In other situations a standardized measure, such as Cohen's d , is helpful. I have stayed away from correlation-based measures as much as I reasonably can because I don't think that they tell the reader much of what he or she wants to know.

I have maintained from earlier editions a section labeled "Seeing Statistics." These sections are built around a set of Java applets, written by Gary McClelland at the University of Colorado. These allow the students to illustrate for themselves many of the concepts that are discussed in the book. The student can open these applets, change parameters, and see what happens to the result. A nice illustration of this is the applet illustrating the influence of heterogeneous subsamples in a correlation problem (see Chapter 9, p. 200). These applets are available directly from my website, referred to earlier.

An important addition to this edition is the inclusion of a chapter on meta-analysis. Meta-analysis is an analysis of multiple studies at the same time. There have been many research studies on the treatment of depression, for example. A meta-analytic study of depression would bring all of those studies together and attempt to draw conclusions on the basis of their similar or differing findings. The current emphasis on evidence-based medicine is an excellent example. If I am to be treated for cancer, for example, I want that treatment to be based on more than the most recent small study that came out last week or on my oncologist's favorite study. What we really have here is the extension of the behavioral science's emphasis on effect sizes along with statistical significance.

In addition to the features already described, the website linked to this book through the publisher's pages (there is a link on my pages) contains a number of other elements that should be helpful to students. These include a Statistical Tutor, which is a set of multiple-choice questions covering the major topics in the chapter. Whenever a student selects an incorrect answer, a box appears explaining the material and helping the student to see what the correct answer should be. I did not write those questions, but I think that they are very well done. There are also links to additional resources, a review of basic arithmetic, and links to other examples and additional material.

Organization and Coverage

This section is meant primarily for instructors, because frequent reference is made to terms that students cannot yet be expected to know. Students may wish to skip to the next section.

- The first seven chapters of the book are devoted to standard descriptive statistics, including ways of displaying data, measures of central tendency and variability, the normal distribution, and those aspects of probability that are directly applicable to what follows.
- Chapter 8 on hypothesis testing and sampling distributions serves as a nontechnical introduction to inferential statistics. The chapter was specifically designed to allow students to examine the underlying logic of hypothesis testing without simultaneously being concerned with learning a set of formulae and the intricacies of a statistical test.
- Chapters 9, 10, and 11 deal with correlation and regression, including multiple regression.
- Chapters 12 through 14 are devoted to tests on means, primarily t tests.
- Chapter 15 is concerned with power and its calculation and serves as an easily understood and practical approach to that topic.
- Chapters 16 through 18 are concerned with the analysis of variance. I have included material on simple repeated-measures designs, but have stopped short of covering mixed designs. These chapters include consideration of basic multiple comparison procedures by way of Fisher's protected t , which not only is an easily understood statistic but has also been shown to be well behaved, under limited conditions,

with respect to both power and error rates. At the request of several users of the earlier editions, I have included treatment of the Bonferroni test, which does a very commendable job of controlling error rates, while not sacrificing much in the way of power when used judiciously. Also included are measures of magnitude of effect and effect size, a fairly extensive coverage of interactions, and procedures for testing simple effects. The effect size material, in particular, is considerably expanded from earlier editions.

- Chapter 19 deals with the chi-square test, although that material could very easily be covered at an earlier point if desired.
- Chapter 20 covers the most prominent distribution-free tests.
- Chapter 21 is a completely new chapter dealing with meta-analysis. Along with an increased emphasis on effect sizes for individual studies, meta-analysis takes us in the direction of combining many similar studies through the use of those effect sizes. This field is becoming much more important, and follows in the footsteps of those in medicine who espouse what is called Evidence Based Medicine. If you are going to be treated for Parkinson's disease, wouldn't you like that treatment to be based on a solid analysis of all of the literature dealing with that disorder? The same is true for our interests in the behavioral sciences.

Not every course would be expected to cover all these chapters, and several (most notably multiple regression, power, and distribution-free statistical methods) can be omitted or reordered without disrupting the flow of the material. (I cover chi-square early in my courses, but it was put late in the text on the advice of reviewers.)

Acknowledgments

Many people have played an important role in the development of this book. My editor, Timothy Matray, was extremely supportive of this revision, offering numerous suggestions for my consideration. The copyeditor did an excellent job of editing of the manuscript and was always supportive on those few occasions when I insisted that quaint spellings and my positioning of prepositions were better than the ones preferred by style manuals. Kristin Ruscetta did a great job of shepherding the manuscript through the publication process and keeping us on track and on time. My daughter, Lynda, compiled the index and did extensive work on aligning and formatting the Instructor and Student manuals and spotting the occasional error.

A number of reviewers made many helpful suggestions in earlier editions, especially Jeffrey Andre (James Madison University), Dr. Kevin J. Apple (Ohio University), Eryl Bassett (University of Kent at Canterbury), Drake Bradley (Bates College), Keith Burt (University of Vermont), Deborah M. Clauson (Catholic University of America), Jose M. Cortina (Michigan State University), Gary B. Forbach (Washburn University), Edward Johnson (University of North Carolina), Dennis Jowaisas (Oklahoma City University), Jeffrey Kahn (Illinois State University), Barbara J. Lehman (Western Washington University), David J. Mostofsky (Boston University), Maureen Powers (Vanderbilt University), David R. Owen (Brooklyn College CUNY), Dennis Roberts (Pennsylvania State University), Steven Rogelberg (Bowling Green State University), Deborah J. Runsey (Kansas State University), Robert Schutz (University of British Columbia), Paul Shirey (University of California at Irvine), N. Clayton Silver (University of Nevada), Patrick A. Vitale (University of South Dakota),

Bruce H. Wade (Spelman College), Todd Wiebers (Henderson State University), Robert Williams (Gallaudet University), Eleanor Willemssen (Santa Clara University), Pamela Zappardino (University of Rhode Island), Dominic Zerbolio (University of Missouri–St. Louis). For years Dr. Karl Wuensch (East Carolina University) has filled pages with suggestions, disagreements, and valuable advice. He deserves special recognition, as does Dr. Kathleen Bloom (University of Waterloo) and Joan Foster (Simon Fraser University). Gary McClelland, at the University of Colorado, graciously allowed me to use some of his Java applets, and was willing to modify them when necessary to meet my needs.

The reviewers for this edition include Yvonne Barry (John Tyler Community College), Russ Espinoza (California State University Fullerton), Donna Goyer (California State University San Marcos), Dr. Jocelyn E. Holden (Ball State University), Alisha Janowsky (University of Central Florida), Rachel W. Kallen (University of Cincinnati), Laura Little (University of Washington), Caitlin Powell (Georgia College & State University), George Tremblay (Antioch University New England), and Diana Young (Georgia College).

I want to thank all of those users (instructors and students alike) who have written me with suggestions and who have pointed out errors. I don't have the space to thank them individually, but many are listed along with the errors they found, on the Web pages labeled "Errata."

I owe thanks to my past colleagues at the University of Vermont. I retired from there in May of 2002, but still consider the University to be my intellectual home. I most certainly want to thank colleagues at the University of Bristol, England, where part of a sabbatical leave was devoted to completing the first edition of the book. Most of all, however, I owe a debt to all of my students who, over the years, have helped me to see where problems lie and how they can best be approached. Their encouragement has been invaluable. Finally, I want to thank the Biometrika trustees for permission to reproduce the table of Wilcoxon's W statistic.

David C. Howell
St. George, Utah
November, 2012
E-mail: David.Howell@uvm.edu



Introduction

Students usually come to any course with some doubt about just what will be involved and how well they will do. This chapter will begin by laying out the kinds of material that we will, and will not, cover. I will then go on to make a distinction between statistics and mathematics, which, for the most part, really are not the same thing at all. As I will point out, all of the math that you need for this course, you learned in high school—though you may have forgotten a bit of it. I will then go on to lay out why we need statistical procedures and what purpose they serve, and to provide a structure for all of the procedures we will cover. Finally, the chapter will provide an introduction to computer analyses of data.

For many years, when I was asked at parties and other social situations what I did for a living, I would answer that I was a psychologist (now retired). Even though I quickly added that I was an experimental psychologist, people would make comments about being careful what they said and acted as if I was thinking all sorts of thoughts that would never occur to me. So finally I changed tactics and started telling people that I taught statistics—an answer that is also perfectly true. That answer solved one problem—people no longer look at me with blatant suspicion—but it created another. Now they tell me how terrible they are in math, and how successful they were in avoiding ever taking a statistics course—not a very tactful remark to make to someone who spent his professional life teaching that subject. Now I just tell them that I taught research methods in psychology for 35 years, and that seems to satisfy them.

Let's begin by asking what the field of statistics is all about. After all, you are about to invest a semester in studying statistical methods, so it might be handy to know what you are studying. The word *statistics* is used in at least three different ways. As used in the title of this book, *statistics* refers to a set of procedures and rules (not always computational or mathematical) for reducing large masses



aplia™
Engage. Prepare. Educate.

Aplia for *Fundamental Statistics for the Behavioral Sciences*

After reading, go to the end of this chapter for instructions on how to use Aplia's homework and learning resources.

of data to manageable proportions and for allowing us to draw conclusions from those data. That is essentially what this book is all about.

A second, and very common, meaning of the term is expressed by such statements as “statistics show that the number of people applying for unemployment benefits has fallen for the third month in a row.” In this case *statistics* is used in place of the much better word *data*. For our purposes, *statistics* will never be used in this sense.

A third meaning of the term is in reference to the result of some arithmetic or algebraic manipulation applied to data. Thus the mean (average) of a set of numbers is a statistic. This perfectly legitimate usage of the term will occur repeatedly throughout this book.

We thus have two proper uses of the term: (1) a set of procedures and rules and (2) the outcome of the application of those rules and procedures to samples of data. You will always be able to tell from the context which of the two meanings is intended.

The term *statistics* usually elicits some level of math phobia among many students, but mathematics and mathematical manipulation do not need to, and often don't, play a leading role in the lives of people who work with statistics. (Indeed, Jacob Cohen, one of the clearest and most influential writers on statistical issues in the behavioral sciences, suggested that he had been so successful in explaining concepts to others precisely because his knowledge of mathematical statistics was so inadequate. People could actually understand what he was saying.) Certainly you can't understand any statistical text without learning a few formulae and understanding many more. But the required level of mathematics is not great. You learned more than enough in high school. Those who are still concerned should spend a few minutes going over Appendix A. It lays out some very simple rules of mathematics that you may have forgotten, and a small investment of your time will be more than repaid in making the rest of this book easier to follow. I know—when I was a student I probably wouldn't have looked at it either, *but you really should!* A more complete review of arithmetic, which is perhaps more fun to read, can be found by going to the website for this book at



https://www.uvm.edu/~dhowell/fundamentals7/ArithmeticReview/review_of_arithmetic_revised.html.

Something far more important than worrying about algebra and learning to apply equations is thinking of statistical methods and procedures as ways to tie the results of some experiment to the hypothesis that led to that experiment. Several editions ago, I made a major effort to remove as much mathematical material as possible when that material did not contribute significantly to your understanding of data analysis. I also simplified equations by going back to definitional formulae rather than present formulae that were designed when we did everything with calculators. This means that I am asking you to think a bit more about the logic of what you are doing. I don't mean just the logic of a hypothesis test. I mean the logic behind the way you approach a problem. It doesn't do any good to be able to ask if two groups have different means (averages) if a difference in means has nothing to say about the real question you hoped to ask. And it does no good to say that a difference is not due to chance without also giving me some idea of

how large the difference is and whether it makes an important difference. When we put too much emphasis on formulae, there is a tendency to jump in and apply those formulae to the data without considering what the underlying question really is. One reviewer whose work I respect has complained that I am trying to teach critical thinking skills along with statistics. The reviewer is right, and I enthusiastically plead guilty. You will never be asked to derive a formula, but you will be asked to think. I leave it to you to decide which skill is harder to learn.

Another concern that some students have, and I may have contributed to that concern in the preceding paragraph, is the belief that the only reason to take a course in statistics is to be able to analyze the results of experimental research. Certainly your instructor hopes many of you will use statistical procedures for that purpose, but those procedures and, more importantly, the ways of thinking that go with them have a life beyond standard experimental research. This is my plea to get the attention of those, like myself, who believe in a liberal arts education. Much of the material we will cover here will be applicable to whatever you do when you finish college. People who work for large corporations or small family-owned businesses have to work with data. People who serve on a town planning commission have to be able to ask how various changes in the town plan will lead to changes in residential and business development. They will have to ask how those changes will in turn lead to changes in school populations and the resulting level of school budgets, and on and on. Those people may not need to run an analysis of variance (Chapters 16 through 18), though some acquaintance with regression models (Chapters 9 through 11) may be helpful, but the logical approach to data required in the analysis of variance is equally required when dealing with town planning. (And if you mess up town planning, you have everybody mad at you.)

A course in statistics is not something you take because it is required and then promptly forget. (Well, that probably is why many of you are taking it, but I hope you expect to come away with more than just three credits on your transcript.) If taught well, knowledge of statistics is a job skill you can use (and market). That is largely why I have tried to downplay the mathematical foundations of the field. Those foundations are important, but they are not what will be important later. Being able to think through the logic and the interpretation of an experiment or a set of data is an important skill that will stay with you; being able to derive the elements of a regression equation is not. That is why most of the examples used in this book relate to work that people actually do. Work of that type requires thought. It may be easier to understand an example that starts out, "Suppose we had three groups labeled A, B, and C" than it is to understand an actual experiment. But the former is boring and doesn't teach you much. A real-life example is more interesting and has far more to offer.

I.1 A Changing Field

People are often puzzled when I say that I am working on a revision of a text. They assume that statistical procedures stay pretty much constant over time. Fortunately that is not the case. Not only do methods for carrying out more complex and interesting

analyses continue to develop, but over the years we have changed the way we look at the results of experimental research. When I was in graduate school and for quite a few years beyond, researchers in the behavioral sciences were primarily concerned with whether a difference that they found between experimental groups (or the relationship they found between two or more variables) was *reliable*. If they ran the study over again, would they still find that the experimental group outperformed a control group? After a while the field slowly began to change by going further and asking if a difference was *meaningful*. Perhaps the groups really were different, but the difference was too small to matter to anyone. That led to the development of a number of different indices of importance, or *effect size*. That was a step forward for the field. Some disciplines were ahead of us in that transition, while other fields were slower to ask that question about meaningfulness.

In the late 1980s, which really is a lot closer to us than you might imagine, a few people in psychology began asking a slightly different question. If the results we found are reliable, and if they are meaningful, what have other guys found? Perhaps there are 20 studies on a particular theoretical question, but people are finding different results. Or perhaps most studies agree, at least in a general sense. This idea of looking at multiple studies on a topic has been extremely important in medicine, where we now speak of “evidence-based practice.” Let’s combine all of the studies on dealing with a particular type of cancer and see if there is agreement on the best form of treatment. This approach is called “meta-analysis,” and I have added a non-technical discussion of it later in the book. It is time that we stopped acting as if the study we ran is the only study of interest. As you can see, the field has moved from “Is this difference reliable?” to “Is this difference meaningful?” to “Is this what other people are finding as well?”

1.2 The Importance of Context

Let’s start with an example that has a great deal to say in today’s world. It may be an old study, but it is certainly an important one, and one that is still cited in the literature on drug use. Drug use and abuse is a major problem in our society. Heroin addicts die every day from overdoses. Psychologists should have something to contribute to understanding the problem of drug overdoses, and, in fact, we do. I will take the time to describe an important line of research in this area because a study that derives from that line of research can be used to illustrate a number of important concepts in this chapter and the next. Many of you will know someone who is involved with heroin, and because heroin is a morphine derivative, this example may have particular meaning to you.

We will take as an example a study similar to an important experiment on morphine tolerance by Shepard Siegel (1975). Morphine is a drug that is frequently used to alleviate pain. Repeated administrations of morphine, however, lead to morphine tolerance, in which morphine has less and less of an effect (pain reduction) over time. (You may have experienced the same thing if you eat spicy food very often. You will find that the more you eat it, the hotter you have to make it to taste the way it did when you started.) A common experimental task that demonstrates morphine tolerance involves placing a rat on an uncomfortably warm surface. When the heat

becomes too uncomfortable, the rat will lick its paws, and the latency of the paw-lick is used as a measure of the rat's sensitivity to pain. A rat that has received a single morphine injection typically shows a longer paw-lick latency, indicating a reduced pain sensitivity. The development of morphine tolerance is indicated by a progressive shortening of paw-lick latencies (indicating increased sensitivity, or decreased insensitivity) with repeated morphine injections.

Siegel noted that there are a number of situations involving drugs other than morphine in which *conditioned* (learned) drug responses are opposite in direction to the unconditioned (natural) effects of the drug. For example, an animal injected with atropine will usually show a marked decrease in salivation. However, if physiological saline (which should have no effect whatsoever) is suddenly injected (*in the same physical setting*) after repeated injections of atropine, the animal will show an *increase* in salivation. It is as if the animal was compensating for the anticipated effect of atropine. In such studies, it appears that a learned compensatory mechanism develops over trials and counterbalances the effect of the drug. (You experience the same thing if you leave the seasoning out of food that you normally add seasoning to. It will taste unusually bland, though the Grape Nuts you eat for breakfast does not taste bland—and I hope that you don't put seasoning on Grape Nuts.)

Siegel theorized that such a process might help to explain morphine tolerance. He reasoned that if you administered a series of pre trials in which the animal was injected with morphine and placed on a warm surface, morphine tolerance would develop. Thus, if you again injected the subject with morphine on a subsequent test trial, the animal would be as sensitive to pain as would be a naive animal (one who had never received morphine) because of the tolerance that has fully developed. Siegel further reasoned that if on the test trial you instead injected the animal with physiological saline, which should have no effect, *in the same test setting* as the normal morphine injections, the conditioned hypersensitivity that results from the repeated administration of morphine would not be counterbalanced by the presence of morphine, and the animal would show very short paw-lick latencies and heightened sensitivity.

You may think that an experiment conducted 40 years ago, which is before most of the readers of this book were born, is too old to be interesting. But a quick Internet search will reveal a great many recent studies that have derived directly from Siegel's early work. A particularly interesting one by Mann-Jones, Ettinger, Baisden, and Baisden (2003) has shown that a drug named dextromethorphan can counteract morphine tolerance. That becomes interesting when you learn that dextromethorphan is an important ingredient in cough syrup. This suggests that heroin addicts don't want to be taking cough syrup any more than they want to be administering heroin in novel environments. The study can be found at <http://www.eou.edu/psych/re/morphinetolerance.doc>.

But what do mice on a warm surface have to do with drug overdose? First, heroin is a derivative of morphine. Second, heroin addicts show clear tolerance effects with repeated use and, as a result, often increase the amount of each injection. By Siegel's theory, they are protected from the dangerous effects of the large (and to you and me, lethal) dose of heroin by the learned compensatory mechanism associated with the context in which they take the drug. But if they take what has come to be their standard dose in an entirely new setting, they would not benefit from that

protective compensatory mechanism, and what had previously been a safe dose could now be fatal. In fact, Siegel noted that many drug overdose cases occur when an individual injects heroin in a novel environment. We're talking about a serious issue here, and drug overdoses often occur in novel settings.

If Siegel is right, his theory has important implications for the problem of drug overdose. One test of Siegel's theory, which is a simplification of studies he actually ran, is to take two groups of mice who have developed tolerance to morphine and whose standard dosage has been increased above normal levels. One group is tested in the same environment in which they previously have received the drug. The second group is treated exactly the same, except that they are tested in an entirely new environment. If Siegel is correct, the animals tested in the new environment will show a much higher pain threshold (the morphine will have more of an effect) than the animals injected in their usual environment. This is the basic study we will build on.

Our example of drug tolerance illustrates a number of important statistical concepts. It also will form a useful example in later chapters of this book. Be sure you understand what the experiment demonstrates. It will help if you think about what events in your own life or the lives of people around you illustrate the phenomenon of tolerance. What effect has tolerance had on behavior as you (or they) developed tolerance? Why is it likely that you probably feel more comfortable with comments related to sexual behavior than do your parents? Would language that you have come to ignore have that same effect if you heard it in a commencement speech?

1.3 Basic Terminology

Statistical procedures can be separated into roughly two overlapping areas: descriptive statistics and inferential statistics. The first several chapters of this book will cover descriptive statistics, and the remaining chapters will examine inferential statistics. We will use the simplified version of Siegel's morphine study to illustrate the differences between these two terms.

Descriptive Statistics

Whenever your purpose is merely to *describe* a set of data, you are employing descriptive statistics. A statement about the average length of time it takes a normal mouse to lick its paw when placed on a warm surface would be a descriptive statistic, as would be the time it takes a morphine-injected mouse to do the same thing. Similarly, the amount of change in the latency of paw-licks once morphine has been administered and the variability of change among mice would be other descriptive statistics. Here we are simply reporting measures that describe average latency scores or their variability. Examples from other situations might include an examination of dieting scores on the Eating Restraint Scale, crime rates as reported by the Department of Justice, and certain summary information concerning examination grades in a particular course. Notice that in each of these examples we are just describing what the data have to say about some phenomenon.

Inferential Statistics

All of us at some time or another have been guilty of making unreasonable generalizations on the basis of limited data. If, for example, one mouse showed shorter latencies the second time it received morphine than it did the first, we might try to claim clear evidence of morphine tolerance. But even if there were no morphine tolerance, or environmental cues played no role in governing behavior, there would still be a 50-50 chance that the second trial's latency would be shorter than that of the first, assuming that we rule out tied scores. Or you might hear or read that tall people tend to be more graceful than short people, and conclude that that is true because you once had a very tall roommate who was particularly graceful. You conveniently forget about the 6' 4" klutz down the hall who couldn't even put on his pants standing up without tripping over them. Similarly, the man who says that girls develop motor skills earlier than boys because his daughter walked at 10 months and his son didn't walk until 14 months is guilty of the same kind of error: generalizing from single (or too limited) observations.

Small samples or single observations may be fine when we want to study something that has very little variability. If we want to know how many legs a cow has, we can find a cow and count its legs. We don't need a whole herd—one will do. However, when what we want to measure varies from one individual to another, such as the amount of milk a cow will produce or the change in response latencies with morphine injections in different contexts, we can't get by with only one cow or one mouse. We need a bunch. This relates to an important principle in statistics—variability. The difference between how we determine the number of legs on a cow versus the milk production of cows depends critically on the degree of variability in the thing we want to measure. Variability will follow you throughout this course.

When the property in question varies from animal to animal or trial to trial, we need to take multiple measurements. However, we can't make an unlimited number of observations. If we want to know whether morphine injected in a new context has a greater effect, how much milk cows generally produce, or when girls usually start to walk, we must look at more than one mouse, one cow, or one girl. But we cannot possibly look at all mice, cows, or girls. We must do something in between—we must draw a *sample* from a *population*.

DEFINITION **Population:** Complete set of events in which you are interested

POPULATIONS, SAMPLES, PARAMETERS, AND STATISTICS: A **population** can be defined as the entire collection of events in which you are interested (e.g., the scores of all morphine-injected mice, the milk production of all cows in the country, the ages at which every girl first began to walk). Thus if we were interested in the stress levels of all adolescent Americans, then the collection of all adolescent Americans' stress scores would form a population, in this case a population of more than 50 million numbers. If, on the other hand, we were interested only in the stress scores of the sophomore class in Fairfax, Vermont (a town of approximately 2,300 inhabitants), the population would contain about 60 numbers and could be obtained quite easily in its entirety. If we were interested in paw-lick latencies of

mice, we could always run another mouse. In this sense the population of scores theoretically would be infinite.

The point is that a population can range from a relatively small set of numbers, which is easily collected, to an infinitely large set of numbers, which can never be collected completely. The populations in which we are interested are usually quite large. The practical consequence is that we can seldom, if ever, collect data on entire populations. Instead, we are forced to draw a **sample** of observations from a population and to use that sample to infer something about the characteristics of the population.

When we draw a sample of observations, we normally compute numerical values (such as averages) that summarize the data in that sample. When such values are based on the sample, they are called **statistics**. The corresponding values in the population (e.g., population averages) are called **parameters**. The major purpose of inferential statistics is to draw inferences about parameters (characteristics of populations) from statistics (characteristics of samples).¹

- Descriptive statistics: Simply describe the set of data at hand.
- Inferential statistics: Use statistics, which are measures on a sample, to infer values of parameters, which are measures on a population.

DEFINITION **Sample:** Set of actual observations; subset of a population.

Statistics: Numerical values summarizing sample data.

Parameters: Numerical values summarizing population data.

Random sample: A sample in which each member of the population has an equal chance of inclusion.

We usually act as if a sample is a truly **random sample**, meaning that each and every element of the population has an equal chance of being included in the sample. If we have a true random sample, not only can we estimate parameters of the population, but we can also have a very good idea of the accuracy of our estimates. To the extent that a sample is not a random sample, our estimates may be meaningless, because the sample may not accurately reflect the entire population. In fact, we rarely take truly random samples because that is impractical in most settings. We usually take samples of convenience (volunteers from Introductory Psychology, for example) and hope that their results reflect what we would have obtained in a truly random sample.

¹The word *inference* as used by statisticians means very much what it means in normal English usage—a conclusion based on logical reasoning. If three-fourths of the people at a picnic suddenly fall ill, I am likely to draw the (possibly incorrect) inference that something is wrong with the food. Similarly, if the average social sensitivity score of a random sample of fifth-grade children is very low, I am likely to draw the inference that fifth graders in general have much to learn about social sensitivity. Statistical inference is generally more precise than everyday inference, but the basic idea is the same.

A problem arises because one person's sample might be another person's population. For example, if I were to conduct a study into the effectiveness of this book as a teaching instrument, the scores of one class on an exam might be considered by me to be a sample, though a nonrandom one, of the population of scores for all students who are or might be using this book. The class instructor, on the other hand, cares only about her own students and would regard the same set of scores as a population. In turn, someone interested in the teaching of statistics might regard my population (the scores of everyone using this book) as a nonrandom sample from a larger population (the scores of everyone using *any* textbook in statistics). Thus the definition of a population depends on what you are interested in studying. Notice also that when we speak about populations, we speak about populations of *scores*, not populations of *people* or *things*.

The fact that I have used nonrandom samples here to make a point should not lead the reader to think that randomness is not important. On the contrary, it is the cornerstone of much statistical inference. As a matter of fact, one could define the relevant population as the collection of numbers from which the sample has been *randomly* drawn.

INFERENCE: We previously defined inferential statistics as the branch of statistics that deals with inferring characteristics of populations from characteristics of samples. This statement is inadequate by itself because it leaves the reader with the impression that all we care about is determining population parameters such as the average paw-lick latency of mice under the influence of morphine. There are, of course, times when we care about the exact value of population parameters. For example, we often read about the incredible number of hours per day the average high school student spends sending text messages, and that is a number that is meaningful in its own right. But if that were all there were to inferential statistics, it would be a pretty dreary subject, and the strange looks I get at parties when I admit to teaching statistics would be justified.

In our example of morphine tolerance in mice, we don't really care what the average paw-lick latency of mice is. But we do care whether the average paw-lick latency of morphine-injected mice tested in a novel context is greater or less than the average paw-lick latency of morphine-injected mice tested in the same context in which they had received previous injections. And for this we need to estimate the corresponding population means. In many cases inferential statistics is a tool used to estimate parameters of two or more populations, more for the purpose of finding if those parameters are different than for the purpose of determining the actual numerical values of the parameters.

Notice that in the previous paragraph it was the population parameters, not the sample statistics, that I cared about. It is a pretty good bet that if I took two different *samples* of mice and tested them, one sample mean (average) would be larger than another. (It's hard to believe that they would come out absolutely equal.) But the real question is whether the sample mean of the mice tested in a novel context is sufficiently larger than the sample mean of mice tested in a familiar context to lead me to conclude that the corresponding *population* means are also different.

And don't lose sight of the fact that we really don't care very much about drug addiction in mice. What we do care about are heroin addicts. But we probably wouldn't be very popular if we gave heroin addicts overdoses in novel settings to see