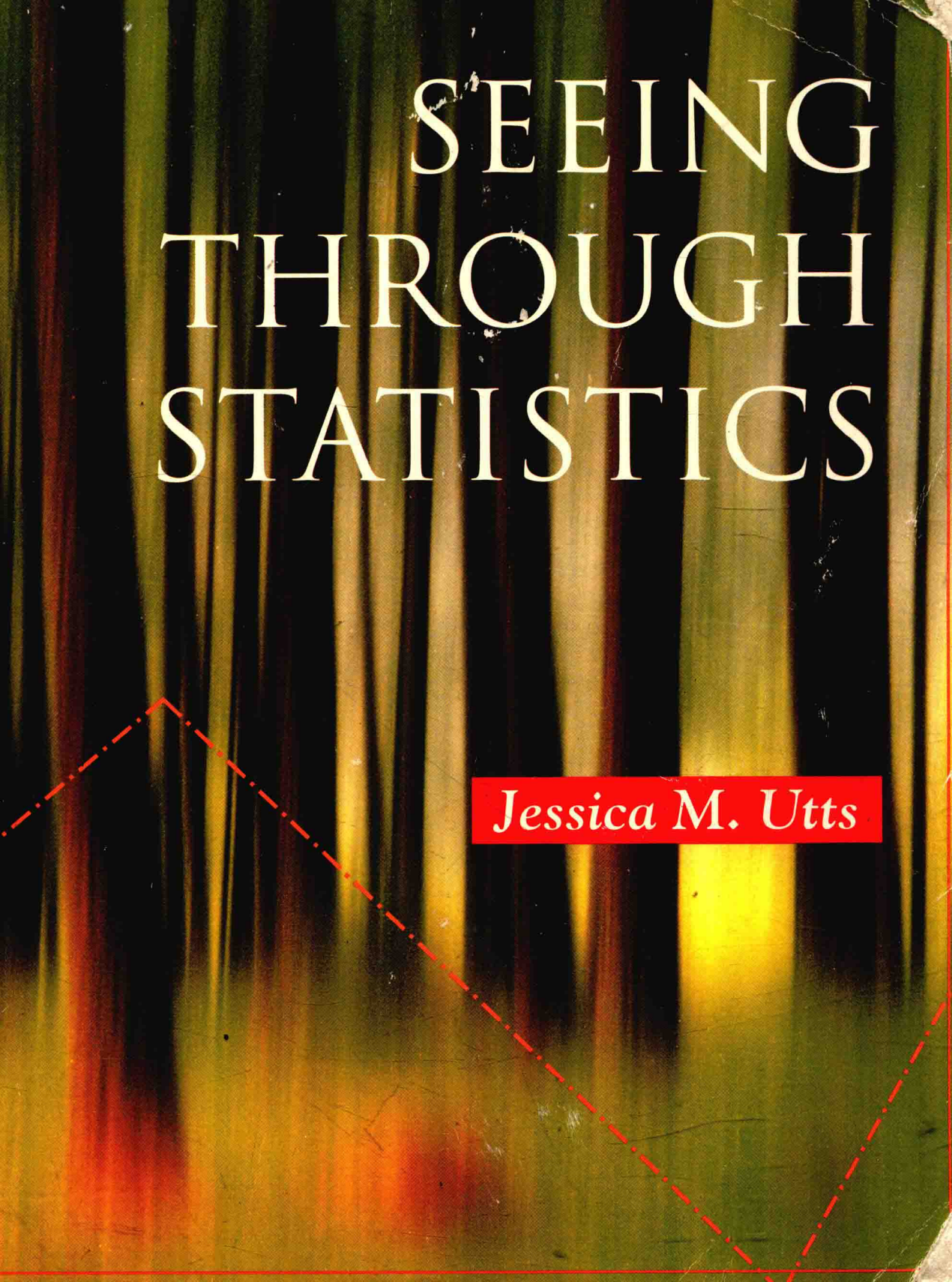
SEEING
THROUGH
STATISTICS
Jessica M. Utts

Seeing Through Statistics

JESSICA M. UTTS
University of California, Davis

An Alexander Kugushev Book

DUXBURY PRESS
AN IMPRINT OF WADSWORTH PUBLISHING COMPANY

I(T)P™ AN INTERNATIONAL THOMSON PUBLISHING COMPANY
Belmont Albany Bonn Boston Cincinnati Detroit London Madrid Melbourne
Mexico City New York Paris San Francisco Singapore Tokyo Toronto Washington

Project Development Editor: Jennifer Burger
Editorial Assistant: Janis C. Brown
Marketing Manager: Joanne M. Terhaar
Production Services Coordinator: Gary Mcdonald
Production: Robin Lockwood & Associates
Designer: Janet Bollow
Print Buyer: Barbara Britton
Permissions Editor: Peggy Meehan
Copy Editor: Rosaleen Bertolino
Technical Illustrator: Carl Brown
Cartoonist: James Stimson
Cover: Lisa Berman
Compositor: Alphatype
Printer: Malloy Lithographing, Inc.

LIBRARY OF CONGRESS CATALOGING-IN-PUBLICATION DATA

Utts, Jessica.
Seeing through statistics / Jessica Utts.
p. cm.
Includes bibliographical references (p. -) and index.
ISBN 0-534-25776-3
1. Statistics. I. Title.
QA276.12.U88 1996
001.4'22—dc20 95-22554

COPYRIGHT © 1996 by Wadsworth Publishing Company
A Division of International Thomson Publishing Inc.
I(T)P The ITP logo is a trademark under license.
Duxbury Press and the leaf logo are trademarks used under license.

This book is printed on acid-free recycled paper.

Printed in the United States of America

4 5 6 7 8 9 10—02 01 99 98

For more information, contact Duxbury Press at Wadsworth Publishing Company:

Wadsworth Publishing Company
10 Davis Drive
Belmont, California 94002, USA

International Thomson Publishing Europe
Berkshire House 168-173
High Holborn
London, WC1V7AA, England

Thomas Nelson Australia
102 Dodds Street
South Melbourne 3205
Victoria, Australia

Nelson Canada
1120 Birchmount Road
Scarborough, Ontario
Canada M1K 5G4

International Thomson Editores
Campos Eliseos 385, Piso 7
Col. Polanco
11560 México D.F. México

International Thomson Publishing GmbH
Königswinterer Strasse 418
53227 Bonn, Germany

International Thomson Publishing Asia
221 Henderson Road
#05-10 Henderson Building
Singapore 0315

International Thomson Publishing Japan
Hirakawacho Kyowa Building, 3F
2-2-1 Hirakawacho
Chiyoda-ku, Tokyo 102, Japan

All rights reserved. No part of this work covered by the copyright hereon may be reproduced or used in any form or by any means—graphic, electronic, or mechanical, including photocopying, recording, taping, or information storage and retrieval systems—without the written permission of the publisher.

To
Patrick and
his grandmother

Preface

Purpose

Statistics deals with complexity and uncertainty. We are exposed daily to information from surveys and scientific studies concerning our health, behavior, attitudes, beliefs, or scientific and technological breakthroughs. This book's first objective is to help you understand this information and to sort out the useful and the accurate from the useless and the misleading. My aim is to allow you to rely on your own interpretation of results emerging from surveys and studies; you will be able to read them with a critical eye and make your own judgments.

A second purpose of this book is to demystify statistical methods. Traditional statistics courses often place emphasis on how to compute, rather than on how to understand. This book emphasizes statistical ideas and their use in real life.

Finally, the book contains information intended to help you make better decisions when faced with uncertainty. You will learn how psychological influences can keep you from making the best decisions, and new ways to think about coincidences, gambling and other circumstances that involve chance events.

Approach

The focus of this book is on the use of statistical methods in the real world. There are dozens of real, in-depth case studies taken from various media sources as well as scores of additional real examples. The emphasis is on understanding, rather than computing, but there are also examples of how to compute important numbers when necessary.

Although the book is written as a textbook, it is also intended to be readable without the guidance of an instructor. Each concept or method is explained in plain language and supported with numerous examples.

Organization

There are 26 chapters divided into four parts. Each chapter covers material more or less equivalent to a one-hour college lecture. The final chapters of Part 1 and Part 4 consist solely of case studies, and are designed to illustrate the thought process you should follow when you read studies on your own.

此为试读，需要完整PDF请访问：www.ertongbook.com

Parts 1 and 2 are entitled "Finding Data in Life" and "Finding Life in Data." By the end of Part 1 you will already have the tools to determine whether or not the results of a study should be taken seriously. You will be able to detect false conclusions and biased results. In Part 2 you will learn how to turn numbers into useful information, and how to quantify relationships between things like aspirin consumption and heart attack rates or meditation and aging. You will also learn how to detect misleading graphs and figures, and how to interpret common economic statistics.

Part 3 is called "Understanding Uncertainty in Life" and is designed to help you do exactly that. We are all required to make decisions in the face of uncertainty on a daily basis. This part of the book will help you understand what probability and chance are all about, and describe some techniques that can help you make better decisions. The material on probability will also be useful for Part 4, "Making Judgments from Surveys and Experiments."

Part 4 is slightly more technical than the rest of the book, but once you have mastered it you will truly understand the beauty of statistical methods. Henceforth when you read the results of a statistical study, you will be able to tell whether they represent valuable advice, or flawed reasoning. Unless things have changed drastically by the time you read this, you will be amazed at the number of news reports that exhibit flawed reasoning.

Thought Questions: Using Your Common Sense

Each of the chapters, except those consisting solely of case studies, begins with a series of "thought questions." They are designed to be read and answered *before* you read the chapter. Most of the answers are based on common sense, perhaps combined with knowledge from previous chapters. Answering them in advance of reading the material will reinforce the idea that most of the information in this book is based on common sense. You will find answers to the thought questions, or similar ones, embedded in the chapters.

In the classroom, the thought questions can be used for discussion at the beginning of each class. For relatively small classes, groups of students can be assigned to discuss one question each then report back to the class. If you are taking a class in which one of these formats is used, try to answer the questions on your own before class. By doing so, you will build confidence as you learn that the material is not difficult to understand if you give it some thought.

Case Studies and Examples: Collect Your Own

The book is filled with real case studies and examples covering a very wide range of disciplines. They are intended to appeal to a broad audience. In the rare cases where technical subject matter knowledge is required it is given with the example. Sometimes the conclusion presented in the book will be different from the

conclusion presented in the original news report. This is because many news reports misinterpret statistical results.

While I hope you find the case studies and examples interesting and informative, you will learn the most by finding current examples on topics of interest to you. Follow any newspaper or news magazine for awhile and you are sure to find plenty of illustrations of the use of surveys and studies. Start collecting them now, and you can watch your understanding of them increase as you work your way through this book.

Formulas: It's Your Choice

If you are one of those who dreads mathematical formulas, you should find this book comfortably readable. In most cases where computations are required, they are presented step-by-step rather than in a formula. And the steps are accompanied by worked examples so you can see exactly how to carry them out.

On the other hand, if you prefer to work with formulas, each relevant chapter ends with a section called "For Those Who Like Formulas." The section includes all of the mathematical notation and formulas pertaining to the material in that chapter.

Exercises and Mini-Projects

There are numerous exercises at the end of each chapter. Many of them are similar to the "thought questions" and require an explanation for which there is no one correct answer. Answers to some of those with concise solutions are provided at the back of the book. *Teaching Seeing Through Statistics: An Instructor's Resource Manual* is available to instructors, and explains what is expected for each exercise.

In most chapters the exercises contain many additional real examples. However, with the idea that you learn best by doing, most chapters also contain "Mini-Projects." Some of these ask you to find examples of studies of interest to you. Others ask you to conduct your own small-scale study. If you are reading this book without the benefit of a class or instructor, I encourage you to try some of the projects on your own.

Covering the Book in a Quarter, Semester or on Your Own

This book was written for a one-quarter course taught three times a week at the University of California at Davis, as part of the general education curriculum. My aim was to allow one lecture for each chapter, and in most cases that practice works quite well. A few chapters may spill over into two lectures, still allowing

for completion of the book (and a midterm or two) in the usual 29 or 30 lecture quarter. I do not cover every detail from each chapter; I expect students to read some material on their own.

If the book is used for a semester course it can be covered at a more leisurely pace and in more depth. For instance, two classes a week can be used for covering new material and a third class for discussion, additional examples or laboratory work. Alternatively, with three regular lectures a week some chapters can be covered in two sessions instead of one.

Instructors can obtain a copy of *Teaching Seeing Through Statistics: An Instructor's Resource Manual*, containing additional information on how to cover the material in one quarter or semester. The *Resource Manual* also gives tips on teaching this material, ideas on how to cover each chapter, sample lectures, additional examples and exercise solutions.

Instructors who want to focus on more in-depth coverage of specific topics may wish to exclude others. Certain chapters can be omitted without serious consequences in later chapters. These include Chapter 9, Chapters 13 and 14 (but Chapter 14 relies on Chapter 13), Chapter 17 and Chapter 25. The chapters in Part 3 (15 to 18) rely on each other, but only parts of Chapter 15 are needed in order to understand Part 4 of the book. Thus, Chapters 16 to 18 (as a unit) could be omitted.

If you are reading this book on your own you may find that you want to concentrate on selected topics only. Parts 1 and 3 can be read alone, as can Chapters 9 and 13. Part 4 relies most heavily on Chapters 8, 12 and 15. Although it is the most technically challenging, I strongly recommend reading Part 4, since it is there that you will truly learn the beauty as well as the pitfalls of statistical reasoning. If you get stuck try to step back and reclaim the big picture. Remember, statistical methods are very powerful and are subject to abuse, but they were developed using the collective common sense of researchers who needed to figure out how to find and interpret information to understand the world. They have done the hard work; this book is intended to help you make sense of it all.

Acknowledgments

I would like to thank the following individuals for reviewing the manuscript and providing valuable insights: Mary L. Baggett, Florida State University; Paul Cantor, Lehman College, City University of New York; Deborah Delanoy, University of Edinburgh; Hariharan K. Iyer, Colorado State University; Richard G. Krutchkoff, Virginia Polytechnic Institute and State University; Scott Plous, Wesleyan University; Larry Ringer, Texas A&M University; Ralph P. Russo, University of Iowa; and Farroll T. Wright, University of Missouri.

I would also like to thank my sisters, Claudia Utts-Smith and Melissa Utts, for helping me realize the need for this book; Alex Kugushev, Duxbury Press, for persisting until I agreed to write it (and beyond); and Robert Heckard, Penn State University, for instilling in me many years ago, by example, the enthusiasm for teaching and educating that led to the development of this material.

Jessica Utts
Davis, CA

Contents

Preface xii

PART 1 *Finding Data in Life* 1

CHAPTER 1 *The Benefits and Risks of Using Statistics* 3

1.1 Statistics 4
CASE STUDY 1.1: Heart or Hypothalamus? 4
1.2 Detecting Patterns and Relationships 6
CASE STUDY 1.2: Does Aspirin Prevent Heart Attacks? 8
1.3 Don't Be Deceived by Improper Use of Statistics 8
CASE STUDY 1.3: A Mistaken Accusation of Cheating 10
1.4 Summary and Conclusions 11
Exercises 11
Mini-Projects 13

CHAPTER 2 *Reading the News* 15

2.1 The Educated Consumer of Data 16
2.2 How to Be a Statistics Sleuth: Seven Critical Components 16
2.3 Four Hypothetical Examples of Bad Reports 20
2.4 Planning Your Own Study: Defining the Components in Advance 27
CASE STUDY 2.1: Brooks Shoes Brings Flawed Study to Court 29
Exercises 31
Mini-Projects 33

CHAPTER 3 *Measurements, Mistakes and Misunderstandings* 34

3.1 Simple Measures Don't Exist 35
3.2 It's All in the Wording 35
3.3 Open or Closed Questions: Should Choices Be Given? 38
3.4 Defining What Is Being Measured 40
CASE STUDY 3.1: Questions in Advertising 42
3.5 Defining a Common Language 42
Exercises 46
Mini-Projects 49
References 49

CHAPTER 4 *How to Get a Good Sample* 51

4.1 Common Research Strategies 52
4.2 Defining a Common Language 54
4.3 The Beauty of Sampling 56
4.4 Simple Random Sampling 57
4.5 Other Sampling Methods 59
4.6 Difficulties and Disasters in Sampling 62
CASE STUDY 4.1: The Infamous *Literary Digest* Poll of 1936 65
Exercises 66
Mini-Projects 70
References 71

CHAPTER 5 *Experiments and Observational Studies* 72

5.1 Defining a Common Language 73
5.2 Designing a Good Experiment 76
CASE STUDY 5.1: Quitting Smoking with Nicotine Patches 78
5.3 Difficulties and Disasters in Experiments 79
5.4 Designing a Good Observational Study 82
CASE STUDY 5.2: Baldness and Heart Attacks 83
5.5 Difficulties and Disasters in Observational Studies 85
Exercises 87
Mini-Projects 91
References 91

CHAPTER 6 *Getting the Big Picture* 93

6.1 Final Questions 93
CASE STUDY 6.1: Mozart, Relaxation and Performance on Spatial Tasks 94
CASE STUDY 6.2: Meditation and Aging 97
CASE STUDY 6.3: Drinking, Driving and the Supreme Court 100
CASE STUDY 6.4: Smoking during Pregnancy and Child's IQ 103
CASE STUDY 6.5: For Class Discussion: Guns and Homicides at Home 106

PART 2 *Finding Life in Data* 107

CHAPTER 7 *Summarizing and Displaying Measurement Data* 109

7.1 Turning Data into Information 110
7.2 The Total Picture: Stemplots and Histograms 111
7.3 Five Useful Numbers: A Summary 118
7.4 Traditional Measures: Mean, Variance and Standard Deviation 119
CASE STUDY 7.1: Detecting Exam Cheating with a Histogram 122

此为试读,需要完整PDF请访问: www.ertongbook.com

For Those Who Like Formulas 123
Exercises 123
Mini-Projects 125
References 125

CHAPTER 8 *Bell-Shaped Curves and Other Shapes* 126

8.1 Populations, Frequency Curves and Proportions 127
8.2 The Pervasiveness of Normal Curves 129
8.3 Percentiles and Standardized Scores 129
8.4 *Z*-Scores and Familiar Intervals 133
For Those Who Like Formulas 134
Exercises 135
References 138

CHAPTER 9 *Plots, Graphs and Pictures* 139

9.1 Well-Designed Statistical Pictures 140
9.2 Pictures of Categorical Data 140
9.3 Pictures of Measurement Variables 143
9.4 Difficulties and Disasters in Plots, Graphs and Pictures 145
9.5 A Checklist for Statistical Pictures 150
Exercises 151
Mini-Projects 154
References 154

CHAPTER 10 *Relationships between Measurement Variables* 155

10.1 Statistical Relationships 156
10.2 Strength versus Statistical Significance 157
10.3 Measuring Strength through Correlation 158
10.4 Specifying Linear Relationships with Regression 161
CASE STUDY 10.1: Are Attitudes about Love and Romance Hereditary? 165
For Those Who Like Formulas 167
Exercises 168
Mini-Projects 170
References 171

CHAPTER 11 *Relationships Can Be Deceiving* 172

11.1 Illegitimate Correlations 173
11.2 Legitimate Correlation Does *Not* Imply Causation 179
11.3 Some Reasons for Relationships between Variables 180
11.4 Confirming Causation 184

Exercises 186
Mini-Projects 189
References 189

CHAPTER 12 *Relationships between Categorical Variables* 190

12.1 Displaying Relationships between Categorical Variables 191
12.2 Assessing the Statistical Significance of a 2×2 Table 194
CASE STUDY 12.1: Extrasensory Perception Works Best with Movies 201
12.3 Relative Risk, Increased Risk and Odds 202
12.4 Misleading Statistics about Risk 206
12.5 Simpson's Paradox: The Missing Third Variable 208
CASE STUDY 12.2: Assessing Discrimination in Hiring and Firing 210
For Those Who Like Formulas 212
Exercises 213
Mini-Projects 219
References 220

CHAPTER 13 *Reading the Economic News* 222

13.1 Cost of Living: The Consumer Price Index 223
13.2 Uses of the Consumer Price Index 227
13.3 Criticisms of the Consumer Price Index 228
13.4 Economic Indicators 230
CASE STUDY 13.1: Did Wages Really Go up in the Reagan-Bush Years? 232
Exercises 233
Mini-Projects 235
References 236

CHAPTER 14 *Understanding and Reporting Trends over Time* 237

14.1 Time Series 238
14.2 Components of Time Series 239
14.3 Seasonal Adjustments: Reporting the Consumer Price Index 243
14.4 Cautions and Checklist 244
CASE STUDY 14.1: If You're Looking for a Job, Try May and October 245
Exercises 247
Mini-Projects 249
References 250

PART 3 *Understanding Uncertainty in Life* 251

CHAPTER 15 *Understanding Probability* 253

15.1 Probability 254
15.2 The Relative-Frequency Interpretation 254
15.3 The Personal-Probability Interpretation 257
15.4 Calibrating Personal Probabilities of Experts 258
15.5 Probability Rules 260
CASE STUDY 15.1: Birthdays and Death Days—Is There a Connection? 262
For Those Who Like Formulas 263
Exercises 264
Mini-Projects 268
References 268

CHAPTER 16 *Possible Futures: What to Expect in the Long Run* 269

16.1 Making Use of Probability as Relative Frequency 270
16.2 When Will It Happen? 270
16.3 Long-Term Gains, Losses and Expectations 273
16.4 Using Expected Values to Make Decisions 276
CASE STUDY 16.1: How Bad Is a Bet on the British Open? 278
Exercises 280
Mini-Projects 283
References 283

CHAPTER 17 *Psychological Influences on Personal Probability* 284

17.1 Revisiting Personal Probability 285
17.2 Equivalent Probabilities; Different Decisions 285
17.3 How Personal Probabilities Can Be Distorted 287
17.4 Optimism, Reluctance to Change and Overconfidence 292
17.5 Tips for Improving Your Personal Probabilities and Judgments 294
Exercises 295
Mini-Projects 297
References 297

CHAPTER 18 *When Intuition Differs from Relative Frequency* 299

18.1 Revisiting Relative Frequency 300
18.2 Coincidences 300
18.3 The Gambler's Fallacy 304
18.4 Confusion of the Inverse 305

CASE STUDY 18.1: Streak Shooting in Basketball: Reality or Illusion? 307
For Those Who Like Formulas 309
Exercises 309
Mini-Projects 312
References 312

PART 4 *Making Judgments from Surveys and Experiments* 313

CHAPTER 19 *The Diversity of Samples from the Same Population* 315

19.1 Setting the Stage 316
19.2 What to Expect of Sample Proportions 316
19.3 What to Expect of Sample Means 321
19.4 What to Expect in Other Situations 325
CASE STUDY 19.1: Do Americans Really Vote When They Say They Do? 326
For Those Who Like Formulas 327
Exercises 328
Mini-Projects 330

CHAPTER 20 *Estimating Proportions with Confidence* 333

20.1 Confidence Intervals 334
20.2 Three Examples of Confidence Intervals from the Media 334
20.3 Constructing a Confidence Interval for a Proportion 337
CASE STUDY 20.1: A Winning Confidence Interval Loses in Court 341
For Those Who Like Formulas 343
Exercises 343
Mini-Projects 347
References 348

CHAPTER 21 *The Role of Confidence Intervals in Research* 349

21.1 Confidence Intervals for Population Means 350
21.2 Confidence Intervals for the Difference between Two Means 353
21.3 Revisiting Case Studies: How Journals Present Confidence Intervals 355
21.4 Understanding Any Confidence Interval 360
For Those Who Like Formulas 361
Exercises 361
Mini-Projects 365
References 365

CHAPTER 22 *Rejecting Chance—Testing Hypotheses in Research* 367

22.1 Using Data to Make Decisions 368
22.2 The Basic Steps for Testing Hypotheses 370
22.3 What Can Go Wrong—The Two Types of Errors 373
CASE STUDY 22.1: Testing for the Existence of Extrasensory Perception 377
Exercises 380
Mini-Projects 383
References 383

CHAPTER 23 *Hypothesis Testing—Examples and Case Studies* 384

23.1 How Hypothesis Tests Are Reported in the News 385
23.2 Testing Hypotheses about Proportions and Means 386
23.3 Revisiting Chi-Squared Tests for Categorical Variables 389
23.4 Revisiting Case Studies: How Journals Present Hypothesis Tests 392
CASE STUDY 23.1: An Interpretation of a P-Value *Not* Fit to Print 395
For Those Who Like Formulas 397
Exercises 398
Mini-Projects 402
References 403

CHAPTER 24 *Significance, Importance and Undetected Differences* 404

24.1 Real Importance versus Statistical Significance 405
24.2 The Role of Sample Size in Statistical Significance 407
24.3 No Difference versus No Statistically Significant Difference 408
CASE STUDY 24.1: Seen a UFO? You May Be Healthier Than Your Friends 410
24.4 A Summary of Warnings 411
Exercises 412
Mini-Projects 416
References 416

CHAPTER 25 *Meta-Analysis: Resolving Inconsistencies across Studies* 417

25.1 The Need for Meta-Analysis 418
25.2 Two Important Decisions for the Analyst 419
CASE STUDY 25.1: Smoking and Reduced Fertility 422
25.3 Some Benefits of Meta-Analysis 423
25.4 Criticisms of Meta-Analysis 426
CASE STUDY 25.2: Controversy over Mammograms 429
Exercises 430

Mini-Projects 432
References 433

CHAPTER 26 *Putting What You Have Learned to the Test* 435

CASE STUDY 26.1: Cranberry Juice and Bladder Infections 435
CASE STUDY 26.2: Children on the Go 436
CASE STUDY 26.3:: It Really Is True about Aspirin 438
CASE STUDY 26.4: You Can Work and Get Your Exercise at the Same Time 440
CASE STUDY 26.5: Sex, Alcohol and the First Date 440
CASE STUDY 26.6: Unpalatable Paté 441
CASE STUDY 26.7: Nursing Moms Can Exercise Too 442
CASE STUDY 26.8: So You Thought Spinach Was Good for You? 444
CASE STUDY 26.9: Chill Out—Move to Honolulu 445
CASE STUDY 26.10: So You Thought Hot Dogs Were Bad for You? 447
References 449

Solutions to Selected Exercises 450
Credits 456
Index 457

此为试读,需要完整PDF请访问: www.ertongbook.com