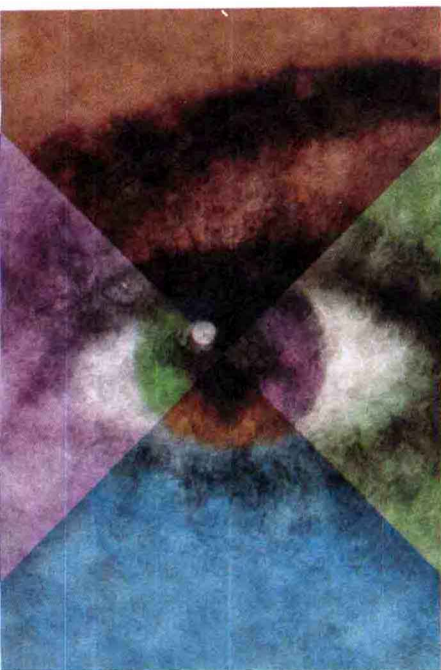


Mind on Statistics

Utts and Heckard



Mind on Statistics

JESSICA M. UTTS

University of California, Davis

ROBERT F. HECKARD

Pennsylvania State University

DUXBURY



™

THOMSON LEARNING

Australia • Canada • Mexico • Singapore • Spain • United Kingdom • United States

Sponsoring Editor: *Carolyn Crockett*
Editorial Assistant: *Jennifer Jenkins*
Marketing: *Tom Ziolkowski/*
 Samantha Cabaluna
Assistant Editor: *Ann Day*
Production Editor: *Tessa Avila*
Production Service: *Martha Emry*
Manuscript Editor: *Pamela Rockwell*

Permissions Editor: *Mary Kay Hancharick*
Interior and Cover Design: *Terri Wright*
Cover Image: *Magic Eye, Inc.*
Indexer: *Steele/Katigbak*
Print Buyer: *Vena Dyer*
Typesetting: *UG/GGS Information*
 Services, Inc.
Printing and Binding: *R. R. Donnelley/Willard*

COPYRIGHT © 2002 the Wadsworth Group. Duxbury is an imprint of the Wadsworth Group, a division of Thomson Learning Inc. Thomson Learning™ is a trademark used herein under license.

For more information about this or any other Duxbury product, contact:

DUXBURY
511 Forest Lodge Road
Pacific Grove, CA 93950 USA
www.duxbury.com
1-800-423-0563 (Thomson Learning Academic Resource Center)

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, Web distribution, or information storage and retrieval systems—without the prior written permission of the publisher.

For permission to use material from this work, contact us by

www.thomsonrights.com

fax: 1-800-730-2215

phone: 1-800-730-2214

Printed in the United States of America

1 0 9 8 7 6 5 4 3

Library of Congress Cataloging-in-Publication Data

Utts, Jessica M.

Mind on statistics / Jessica M. Utts, Robert F. Heckard.

p. cm

Includes bibliographic references and index.

ISBN 0-534-35935-3

1. Statistics. I. Heckard, R. F. II. Title.

QA276.12.U78 2002

519.5—dc21

2001025660

MINITAB is a trademark of Minitab, Inc., and is used herein with the owner's permission. Portions of MINITAB Statistical Software input and output contained in this book are printed with permission of Minitab, Inc.

Minitab, Inc.

3081 Enterprise Drive

State College, PA, 16801-3008

www.minitab.com

Sales: 800-488-3555, ext. 270

Fax: 814-238-4383

Email: MBower@minitab.com

Tech Support: 814-231-2682

All products mentioned herein are used for identification purposes only and may be trademarks or registered trademarks of their respective owners.



*To Bill Harkness—energetic, generous, and innovative
educator, guide, and friend—who launched our careers in statistics
and continues to share his vision.*



Preface

A CHALLENGE

Before you continue, think about how you would answer the question in the first bullet, and read the statement in the second bullet. We will return to them a little later in this Preface.

- ◆ What do you *really know* is true, and how do you know it?
- ◆ The diameter of the moon is about 2160 miles.

WHAT IS STATISTICS AND WHO SHOULD CARE?

Because people are curious about many things, chances are that your interests include topics to which the science of statistics has made a useful contribution. As written in Chapter 17, “information developed through the use of statistics has enhanced our understanding of how life works, helped us learn about each other, allowed control over some societal issues, and helped individuals make informed decisions. There is almost no area of knowledge that has not been advanced by statistical studies.”

Statistical methods have contributed to our understanding of health, psychology, ecology, politics, music, lifestyle choices, and dozens of other topics. A quick look through this book, especially Chapters 1 and 17, should convince you of this. Watch for the influences of statistics in your daily life as you learn this material.

Although statistics courses are often offered through mathematics departments, statistics is not a branch of mathematics. Mathematics is to statistics as wood, hammer, and nails are to a house: a partial set of materials and tools. In addition to mathematics, statistics also draws materials and tools from philosophy, graphics, computing, psychology, and language.

HOW IS THIS BOOK DIFFERENT? TWO BASIC PREMISES OF LEARNING

We wrote this book because we were tired of being told that what statisticians do is boring and difficult. We think statistics is useful and not difficult to learn, and yet the majority of college graduates we've met seem to have had a negative experience with a statistics class in college. We hope this book will help to overcome these misguided stereotypes.

Let's return to the two bullets at the beginning of this Preface. Without looking, do you remember the diameter of the moon? Unless you already had a pretty good idea, or have an excellent memory for numbers, you probably don't remember. One premise of this book is that **new material is much easier to learn and remember if it is related to something interesting or previously known**. The diameter of the moon is about the same as the air distance between Atlanta and Los Angeles, San Francisco and Chicago, London and Cairo, or Moscow and Madrid. Picture the moon sitting between any of those pairs of cities, and you are not likely to forget the size of the moon again. Throughout this book, new material is presented in the context of interesting and useful examples. The first and last chapters (1 and 17) are exclusively devoted to examples and case studies, which illustrate the wisdom that can be generated through statistical studies.

Now answer the question asked in the first bullet: What do you *really know* is true, and how do you know it? If you're like most people, you know because it's something you have experienced or verified for yourself. It's not likely to be something you were told or heard in a lecture. The second premise of this book is that **new material is easier to learn if you actively ask questions and answer them for yourself**. *Mind On Statistics* is designed to help you learn statistical ideas by actively thinking about them.

TOOLS FOR EXPANDED LEARNING

There are a number of tools provided in this book and beyond to enhance your learning of statistics.

Throughout most of the chapters there are boxes entitled *Turn on Your Mind*. Thinking about the questions in these boxes will help you discover and verify important ideas for yourself. We encourage you to think and question, rather than simply read and listen.

Special *Tech Note* boxes provide additional technical discussion as well as details about using MINITAB™ statistical software and Microsoft® Excel.



Explain in your own words what it means to say that we have 95% confidence in the interval estimate. Then give an example of something you do in your life that illustrates the same concept—you follow the same procedure each time, and it either works (most of the time) or does not work to produce the desired result. What confidence level would you assign to the procedure in your example; i.e., what percent of the time do you think it produces your desired result?

Additional Notes on
Creating Histograms



- ◆ Consider using intervals that make the range and width of each interval convenient. For instance, to create a histogram of ages at death for First Ladies, it would be convenient to use ten-year periods—died in her 30s, died in her 40s, and so on, up to 90s. This would create seven intervals.
- ◆ To show relative frequency, you can use either the proportion or the percent that are in an interval. ◆

Case Studies present real-world stories and articles about current topics followed by further discussion, so that you can see how to apply statistical thinking to issues you might encounter while reading the newspaper or talking with friends.

CASE STUDY 10.2

Nicotine Patches versus Zyban®

Some of you may know from personal experience that quitting smoking is difficult. Many people trying to quit use nicotine replacement methods like nicotine patches or nicotine gum to ease nicotine withdrawal symptoms. Recently, medical researchers have begun investigating whether the use of an antidepressant medication might be a more effective aid to those attempting to give up cigarettes. In a study reported in the March 4, 1999 *New England Journal of Medicine*, Dr. Douglas Jorenby and colleagues compared the effectiveness of nicotine patches to the effectiveness of the antidepressant bupropion, which is marketed with the brand name Zyban.

The 893 participants were randomly allocated to four treatment groups: placebo, nicotine patch only, Zyban only, and Zyban plus nicotine patch. To keep participants blind as to their treatments, they all used a patch (nicotine or placebo) and took a pill (Zyban or placebo). For instance, in the placebo-only group, participants used both a placebo patch and took placebo pills. The Zyban group also used placebo patches, and so on. The treatments were used for nine weeks.

Table 10.4 displays, for each treatment group, an approximate 95% confidence interval for the proportion not

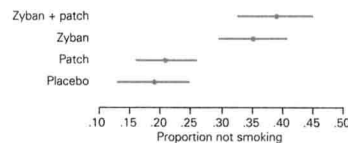


FIGURE 10.4 95% Confidence intervals for proportion not smoking after six months

smoking six months after the start of the experiment. Each interval is a range of estimates of the proportion that would not be smoking after six months in a population of individuals using that particular method. The results shown in the table indicate that Zyban improves the success rate.

The display in Figure 10.4 compares the 95% confidence intervals graphically, clearly showing an appropriate conclusion for this experiment.

Based on the graph, we can make the following generalizations:

- ♦ Zyban is effective. The Zyban groups had higher success rates and the confidence intervals for the two groups that used Zyban do not overlap the intervals for the two groups not using Zyban.
- ♦ The nicotine patch is not particularly effective. There is substantial overlap in the range of estimates in the intervals for the nicotine patch and placebo groups, so we can't conclude that the patch is better than the placebo. Similarly, there is substantial overlap between the intervals for the Zyban-only group and the Zyban-plus-patch group, so these two treatments do not significantly differ.

TABLE 10.4 95% Confidence Intervals for Proportion Not Smoking after Six Months

Treatment	Subjects	Proportion Not Smoking	Approx. 95% CI
Placebo only	160	.188	.13 to .25
Nicotine patch	244	.213	.16 to .26
Zyban	244	.348	.29 to .41
Zyban and nicotine patch	245	.388	.33 to .44

Key Terms at the end of each chapter, organized by section, function as a “quick-finder” and as a review tool.

KEY TERMS

Section 9.1

statistic, 260
parameter, 260
sampling distribution, 260

Section 9.2

sample proportion, 261
normal curve approximation
rule for sample proportions,
262, 264
sampling distribution of \hat{p} ,
264
standard deviation of \hat{p} , 265
standard error of \hat{p} , 265

Section 9.3

normal curve approximation rule
for sample means, 268
sampling distribution of \bar{x} , 268
sampling distribution of the mean,
268
standard deviation of the mean,
268
standard error of the mean,
268
law of large numbers, 270

Section 9.4

central limit theorem, 270, 271

Section 9.6

standardized statistic, 275
standardized z-statistic, 275

Section 9.7

student's t-distribution, 276
t-distribution, 276
degrees of freedom, 276

Section 9.8

statistical inference, 278
confidence intervals, 278
hypothesis testing, 278
significance testing, 278
statistical significance, 278

A CD provided in the back of the book includes many of the data sets used in this book, allowing you to explore ideas and play with the data.

Dataset Exercises



- 2.64 The data for this exercise are in the GSS-93 dataset. The variable *gunlaw* is whether a respondent favors or opposes stronger gun control laws.
- Determine the percent of respondents who favor stronger gun control laws and the percent of respondents who oppose stronger gun control laws. (Note: Not all survey participants were asked the question about gun laws so the sample size for *gunlaw* is smaller than the overall sample size.)
 - Draw a graphical summary of the *gunlaw* variable.
 - Create a two-way table of counts that shows the relationship between gender and opinion about stronger gun control laws. From looking at this table of counts, can you judge whether the two variables are related? Briefly explain.
 - What percent of females favors stronger gun control laws? What percent of males favors stronger gun control laws?
 - Do you think that gender and opinion about gun control are related? Briefly explain.

- Create a table that displays the relationship between political party and opinion about the death penalty. Calculate an appropriate set of conditional percentages for describing the relationship.
- Are the variables *polparty* and *capppun* related? Explain.

- 2.68 Use the *cholestat* dataset for this exercise. The data set contains cholesterol levels for 30 “control” patients and 28 heart attack patients at a medical facility. For the heart attack patients, cholesterol levels were measured 2 days, 4 days, and 14 days after the heart attack.
- Calculate the mean, the standard deviation, and the five-number summary for the control patients.
 - Calculate the mean, the standard deviation, and the five-number summary for the heart attack patients’ cholesterol levels 2 days after their attacks.
 - Generally, which group has the higher cholesterol levels? How much difference is there in the location of the cholesterol levels of the two groups?
 - Which group of measurements has a larger spread? Compare the groups with regard to all three measures of spread introduced in Sections 2.6 and 2.7.

Answers to Selected Exercises, provided in the back of the book, allow you to check your answers on those exercises, and guide your thinking on similar exercises.

Answers to Selected Exercises

Chapter 1


- 1.1 a. Self-selected or volunteer.
b. No. Readers with strong opinions will respond.
- 1.3 189/11,034, or about 17/1000, based on placebo group.
- 1.7 a. 150 mph
b. 55 mph
c. 80 mph
d. $\frac{1}{2}$
e. 51
- 1.11 No.
- 1.15 The base rate for that type of cancer.
- 1.19 a. $212/1525 = 0.139$
b. $1/\sqrt{1525} = 0.026$
c. 0.113 to 0.165

Chapter 2

- 2.1 a. Support ban or not; categorical.
b. Gain on verbal and math SATs after program; quantitative.
c. Smoker or not and Alzheimer sufferer or not; both categorical.
- 2.2 a. Question 1a.
- 2.4 Example: Letter grades (A, B, etc.) converted to GPA.
- 2.8 a. Explanatory is smoked or not; response is developed Alzheimer's disease or not.
- 2.10 Pie chart is more informative.
- 2.13 Example: The age of a person who is 80 years old would be an outlier at a traditional college, but not at a retirement home.
- 2.16 Whether it's the male author (then not an outlier) or the female author (then an outlier).
- 2.19 Yes. Values inconsistent with the bulk of the data will be obvious.

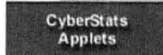
- 2.21 a. This is personal preference; some may prefer a very large family.
b. An outlier (in the high direction).
- 2.24 The median is 16.72 inches. The data values vary from 6.14 to 37.42 inches. The middle $\frac{1}{2}$ of the data is between 12.05 and 25.37 inches, so "typical" annual rainfall covers quite a wide range.
- 2.28 0, 20, 55, 175, 450; data values are skewed to the right and there is an extreme outlier of 450 CDs. (Quartile values may differ slightly if using the computer.)
- 2.31 32, 45.5, 50, 57, 74.
- 2.34 The Empirical Rule predicts about 68%, 95%, 99.7% within 1, 2, and 3 standard deviations of the mean; data show 72%, 97%, 98%, so the set of measurements fits well.
- 2.36 a. Population.
b. Population, 14.77.
- 2.40 a. Would hold without the two outliers; should still be close.
b. Yes, range is 10.75 cm, close to 6 standard deviations. Expect between 4 and 6 standard deviations.
- 2.42 Example: Male height of 73, $z = 1.00$.
- 2.44 a. $z = -0.5$; 0.3085.
b. $z = 2.5$; 0.9938.
- 2.46 50, 50, 50, 50, 50, 50, 50; no.
- 2.49 a. Mean = 51.47 years; standard deviation = 8.92 years (population standard deviation = 8.85).
b. Range is 42 years, 4.7 standard deviations, so it holds.
c. z for youngest CEO is -2.18 , z for oldest CEO is 2.53; about as expected from the Empirical Rule.
- 2.50 a. Categorical.
b. Quantitative.
- 2.52 a. Yes; night light use, for example.
b. No.
- 2.56 a. Set 2—it covers a much wider range of heights.

A website links you to useful resources and coverage of additional statistical topics, including multiple regression, computing Fisher's exact test, logistic regression, two-sample randomization tests, Kruskal-Wallis and other nonparametric methods, and the hypergeometric distribution. These resources can be found on the *Online Book Companion* site at www.duxbury.com. Instructors may contact us with suggestions for topics to be added to the website.




Utts/Heckard

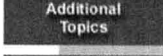
On Line Book Companions



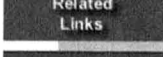
CyberStats
Applets



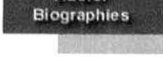
Data
Sets



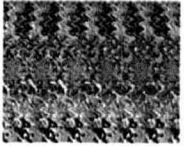
Additional
Topics



Related
Links




Author
Biographies




Mind on Statistics


Utts and Heckard




Order
Online




Review
Copies




Contact
Us




Permission
Forms




Author
Form




Reviewer
Form



Duxbury
Home



Brooks/Cole
Home



Thomson
Learning
Home

The *Instructor's Resource Manual* contains the complete solutions to all exercises, lecture suggestions and guidelines, additional examples, and other helpful ideas for teaching the course.

The *Student Solutions Manual* includes the complete solutions to all the end-of-book answers and offers many helpful hints and suggestions.

ACKNOWLEDGMENTS

We thank William Harkness, Professor of Statistics at Pennsylvania State University, for continued support and feedback throughout our careers and during the writing of this book and for his remarkable dedication to undergraduate statistics education. At Penn State, David Hunter and Steve Arnold provided many helpful insights during spirited hallway discussions; psychology professor Melvin Mark provided useful information for the butterfly ballot example in Chapter 15; and Kellie Karaky helped format several preliminary versions, usually in the final minute before a print shop deadline. Preliminary editions of *Mind On Statistics* were used at Penn State, the University of California at Davis, and Texas A & M University, and we thank the many students who provided comments and suggestions. Thanks to Dr. Melvin Morse (Valley Children's Clinic and University of Washington) for suggesting the title for Chapter 17. The following reviewers offered valuable suggestions: Patti B. Collings, Brigham Young University; James Curl, Modesto Junior College; Donald Harden, Georgia State University; Rosemary Hirschfelder, University Sound; Sue Holt, Cabrillo Community College; Tom Johnson, North Carolina University; Andre Mack, Austin Community College; D'Arcy Mays, Virginia Commonwealth; Mary Murphy, Texas A & M University; N. Thomas Rogness, Grand Valley State University; Heather Sasinouska, Clemson University; and Robert Alan Wolf, University of San Francisco. Our sincere appreciation and gratitude goes to Carolyn Crockett and the Duxbury staff, without whom this book could not have been written. Finally, for their support, patience, and numerous prepared dinners, we thank our families and friends, especially Candace Heckard, Molly Heckard, Wes Johnson, Claudia Utts-Smith, and Dennis Smith.

Jessica Utts
Robert Heckard



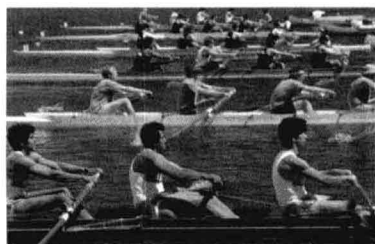
Contents

CHAPTER 1 Statistics Success Stories and Cautionary Tales 1



- 1.1 What Is Statistics? 1
- 1.2 Seven Statistical Stories with Morals 2
 - ◆ **Case Study 1.1:** Who Are Those Speedy Drivers? 2
 - ◆ **Case Study 1.2:** Disaster in the Skies? 3
 - ◆ **Case Study 1.3:** Did Anyone Ask Whom You've Been Dating? 3
 - ◆ **Case Study 1.4:** Who Are Those Angry Women? 4
 - ◆ **Case Study 1.5:** Does Prayer Lower Blood Pressure? 5
 - ◆ **Case Study 1.6:** Does Aspirin Reduce Heart Attack Rates? 5
 - ◆ **Case Study 1.7:** Does the Internet Increase Loneliness and Depression? 6
- 1.3 The Common Elements in the Seven Stories 7
 - KEY TERMS** 8
 - EXERCISES** 9
 - REFERENCES** 10

CHAPTER 2 Turning Data into Information 12



- 2.1 Raw Data 13
- 2.2 Types of Data 14
- 2.3 Summarizing One or Two Categorical Variables 18
- 2.4 Finding Information in Quantitative Data 23
- 2.5 Pictures for Quantitative Data 28
- 2.6 Numerical Summaries of Quantitative Variables 32
- 2.7 Bell-Shaped Distributions of Numbers 40

KEY TERMS	46
EXERCISES	46
REFERENCES	51

CHAPTER 3 Gathering Useful Data 52



3.1	Description or Decision? Using Data Wisely	53
3.2	Speaking the Language of Research Studies	55
◆	Case Study 3.1: Lead Exposure and Bad Teeth	60
3.3	Designing a Good Experiment	61
◆	Case Study 3.2: Kids and Weight Lifting	61
◆	Case Study 3.3: Quitting Smoking with Nicotine Patches	66
3.4	Designing a Good Observational Study	67
◆	Case Study 3.4: Baldness and Heart Attacks	67
3.5	Difficulties and Disasters in Experiments and Observational Studies	69
	KEY TERMS	74
	EXERCISES	74
	REFERENCES	78

CHAPTER 4 Sampling: Surveys and How To Ask Questions 80



4.1	The Beauty of Sampling	81
4.2	Sampling Methods	84
4.3	Difficulties and Disasters in Sampling	93
◆	Case Study 4.1: The Infamous <i>Literary Digest</i> Poll of 1936	97
4.4	How to Ask Survey Questions	98
◆	Case Study 4.2: No Opinion of Your Own? Let Politics Decide	102
	KEY TERMS	102
	EXERCISES	103
	REFERENCES	107

CHAPTER 5 Relationships Between Quantitative Variables 108



5.1	Looking for Patterns with Scatterplots	110
5.2	Describing Linear Patterns with a Regression Line	115
5.3	Measuring Strength and Direction with Correlation	122
5.4	Why Answers May Not Make Sense	128
5.5	Correlation Does Not Prove Causation	133
◆	Case Study 5.1: A Weighty Issue	135
	KEY TERMS	137
	EXERCISES	137
	REFERENCES	143

CHAPTER 6 Relationships Between Categorical Variables 144



- 6.1 Displaying Relationships Between Categorical Variables 145
- 6.2 Risk, Relative Risk, Odds Ratio, and Increased Risk 148
- 6.3 Misleading Statistics about Risk 151
- 6.4 The Effect of a Third Variable and Simpson's Paradox 154
- 6.5 Assessing the Statistical Significance of a 2×2 Table 157
 - ◆ **Case Study 6.1:** Drinking, Driving, and the Supreme Court 163
- KEY TERMS** 165
- EXERCISES** 165
- REFERENCES** 171

CHAPTER 7 Probability 172



- 7.1 Random Circumstances 173
 - ◆ **Case Study 7.1:** A Hypothetical Story—Alicia Has a Bad Day 174
- 7.2 Interpretations of Probability 176
- 7.3 Probability Definitions and Relationships 181
- 7.4 Basic Rules for Finding Probabilities 186
- 7.5 Strategies for Finding Complicated Probabilities 192
- 7.6 Using Simulation to Estimate Probabilities 199
- 7.7 Coincidences and Intuitive Judgments about Probability 202
- KEY TERMS** 209
- EXERCISES** 209
- REFERENCES** 214

CHAPTER 8 Random Variables 216



- 8.1 What is a Random Variable? 217
- 8.2 Discrete Random Variables 219
- 8.3 Expectations for Random Variables 223
- 8.4 Binomial Random Variables 228
- 8.5 Continuous Random Variables 233
- 8.6 Normal Random Variables 236
- 8.7 Approximating Binomial Distribution Probabilities 243
- 8.8 Sums, Differences, and Combinations of Random Variables 245
 - ◆ **Case Study 8.1:** Does Caffeine Enhance the Taste of Cola? 251
- KEY TERMS** 252
- EXERCISES** 252
- REFERENCES** 257

CHAPTER 9 Means and Proportions as Random Variables 258



- 9.1 Understanding Dissimilarity among Samples 259
- 9.2 Sampling Distributions for Sample Proportions 261
- 9.3 What to Expect of Sample Means 265
- 9.4 What to Expect in Other Situations: Central Limit Theorem 270
- 9.5 Sampling Distribution for Any Statistic 272
- 9.6 Standardized Statistics 274
- 9.7 Student's t -Distribution: Replacing σ with s 276
- 9.8 Statistical Inference 278
 - ◆ **Case Study 9.1:** Do Americans Really Vote When They Say They Do? 279
- KEY TERMS** 279
- EXERCISES** 280
- REFERENCES** 285

CHAPTER 10 Estimating Proportions with Confidence 286



- 10.1 The Language and Notation of Estimation 288
- 10.2 Margin of Error 289
- 10.3 Confidence Intervals 290
- 10.4 Calculating a Margin of Error for 95% Confidence 292
- 10.5 General Theory of Confidence Intervals for a Proportion 296
- 10.6 Choosing a Sample Size for a Survey 301
- 10.7 Using Confidence Intervals to Guide Decisions 302
 - ◆ **Case Study 10.1:** Extrasensory Perception Works with Movies 304
 - ◆ **Case Study 10.2:** Nicotine Patches versus Zyban® 305
 - ◆ **Case Study 10.3:** What a Great Personality 306
- KEY TERMS** 306
- EXERCISES** 307
- REFERENCES** 311

CHAPTER 11 Testing Hypotheses about Proportions 312



- 11.1 Formulating Hypothesis Statements 314
- 11.2 The Logic of Hypothesis Testing: What if the Null Is True? 316
- 11.3 Reaching a Conclusion about the Two Hypotheses 318
- 11.4 Testing Hypotheses about a Proportion 320
- 11.5 The Role of Sample Size in Statistical Significance 331

- 11.6** Real Importance versus Statistical Significance 334
 ♦ **Case Study 11.1:** The Internet and Loneliness:
 Case Study 1.7 Revisited 335
- 11.7** What Can Go Wrong: The Two Types of Errors 335
 ♦ **Case Study 11.2:** An Interpretation of a p -Value Not Fit
 to Print 339
- KEY TERMS** 340
EXERCISES 340
REFERENCES 345

CHAPTER 12 More about Confidence Intervals 346



- 12.1** Examples of Different Estimation Situations 347
- 12.2** Standard Errors 350
- 12.3** Approximate 95% Confidence Intervals 354
- 12.4** General Confidence Intervals for One Mean
 or Paired Data 357
- 12.5** General Confidence Intervals for the Difference Between
 Two Means (Independent Samples) 363
- 12.6** The Difference Between Two Proportions
 (Independent Samples) 370
- 12.7** Understanding Any Confidence Interval 373
 ♦ **Case Study 12.1:** Confidence Interval for Relative Risk:
 Case Study 3.4 Revisited 373
 ♦ **Case Study 12.2:** Premenstrual Syndrome?
 Try Calcium 374
- Summary of Formulas for Confidence Intervals 375
- KEY TERMS** 376
EXERCISES 376
REFERENCES 383

CHAPTER 13 More about Significance Tests 384



- 13.1** The General Ideas of Significance Testing 386
- 13.2** Testing Hypotheses about One Mean or
 Paired Data 387
- 13.3** Testing the Difference Between Two Means
 (Independent Samples) 396
- 13.4** Testing the Difference Between Two
 Population Proportions 404
- 13.5** The Relationship Between Significance Tests and
 Confidence Intervals 409
- 13.6** The Two Types of Errors and
 Their Probabilities 411

- 13.7** Evaluating Significance in Research Reports 415
 Summary of Chapter 13 Procedures 416
KEY TERMS 416
EXERCISES 416
REFERENCES 423

CHAPTER 14 More about Regression 424



- 14.1** Sample and Population Regression Models 426
14.2 Estimating the Standard Deviation for Regression 431
14.3 Inference about the Linear Regression Relationship 435
14.4 Predicting the Value y for an Individual 439
14.5 Estimating the Mean y at a Specified x 441
14.6 Checking Conditions for Using Regression Models for Inference 443
 ♦ **Case Study 14.1:** A Contested Election 447
KEY TERMS 449
EXERCISES 450
REFERENCES 455

CHAPTER 15 More about Categorical Variables 456



- 15.1** The Chi-square Test for Two-Way Tables 457
15.2 Analyzing 2×2 Tables 468
15.3 Testing Hypotheses about One Categorical Variable: Goodness of Fit 473
 ♦ **Case Study 15.1:** Do You Mind if I Eat the Blue Ones? 477
KEY TERMS 479
EXERCISES 479
REFERENCES 436

CHAPTER 16 Analysis of Variance 488



- 16.1** Comparing Means with an ANOVA F -Test 489
16.2 Details of One-Way Analysis of Variance 497
16.3 Other Methods for Comparing Populations 503
16.4 Two-Way Analysis of Variance 507
KEY TERMS 511
EXERCISES 511
REFERENCES 516

CHAPTER 17 Turning Information into Wisdom 518



- 17.1** Beyond the Data 519
- 17.2** Transforming Uncertainty into Wisdom 522
- 17.3** Making Personal Decisions 522
- 17.4** Control of Societal Risks 525
- 17.5** Understanding Our World 528
- 17.6** Getting to Know You 530
- 17.7** Words to the Wise 532

EXERCISES 533

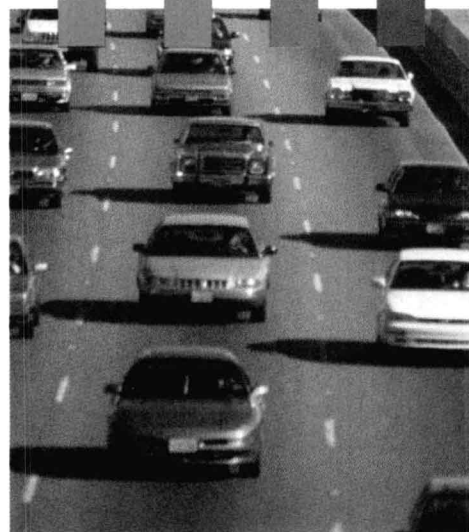
REFERENCES 535

Appendix of Tables 537

Answers to Selected Exercises 543

Index 553

The seven stories in this chapter are meant to bring life to the term *statistics*. When you are finished reading these stories, if you still think the subject of statistics is lifeless or gruesome, check your pulse!



Let's face it. You're a busy person. Why should you spend your time learning about a subject that sounds as dull as statistics? In this chapter we give seven examples of situations in which statistics either provided enlightenment or misinformation. With these examples, we hope to convince you that learning about this subject will be interesting and useful.

Each of the stories in this chapter illustrates one or more concepts that will be developed throughout the book. These concepts are given as "the moral of the story" after a case is presented. Definitions of some terms used in the story also are provided following each case. By the time you read all of these stories, you already will have an overview of what statistics is all about. ♦

1.1 WHAT IS STATISTICS?

When you hear the word *statistics* you probably think of lifeless or gruesome numbers, like the population of your state or the number of violent crimes committed in your city last year. The word *statistics*, however, actually is used to mean two different things. The better-known definition is that statistics are numbers measured for some purpose. A more complete definition, and the one that forms the substance of this book, is the following:

Statistics is a collection of procedures and principles for gathering data and analyzing information in order to help people make decisions when faced with uncertainty.

The stories in this chapter are meant to bring life to this definition. When you are finished reading them, if you still think the subject of statistics is lifeless or gruesome, check your pulse!