ALL-IN-ONE

Statistics

for dummies®

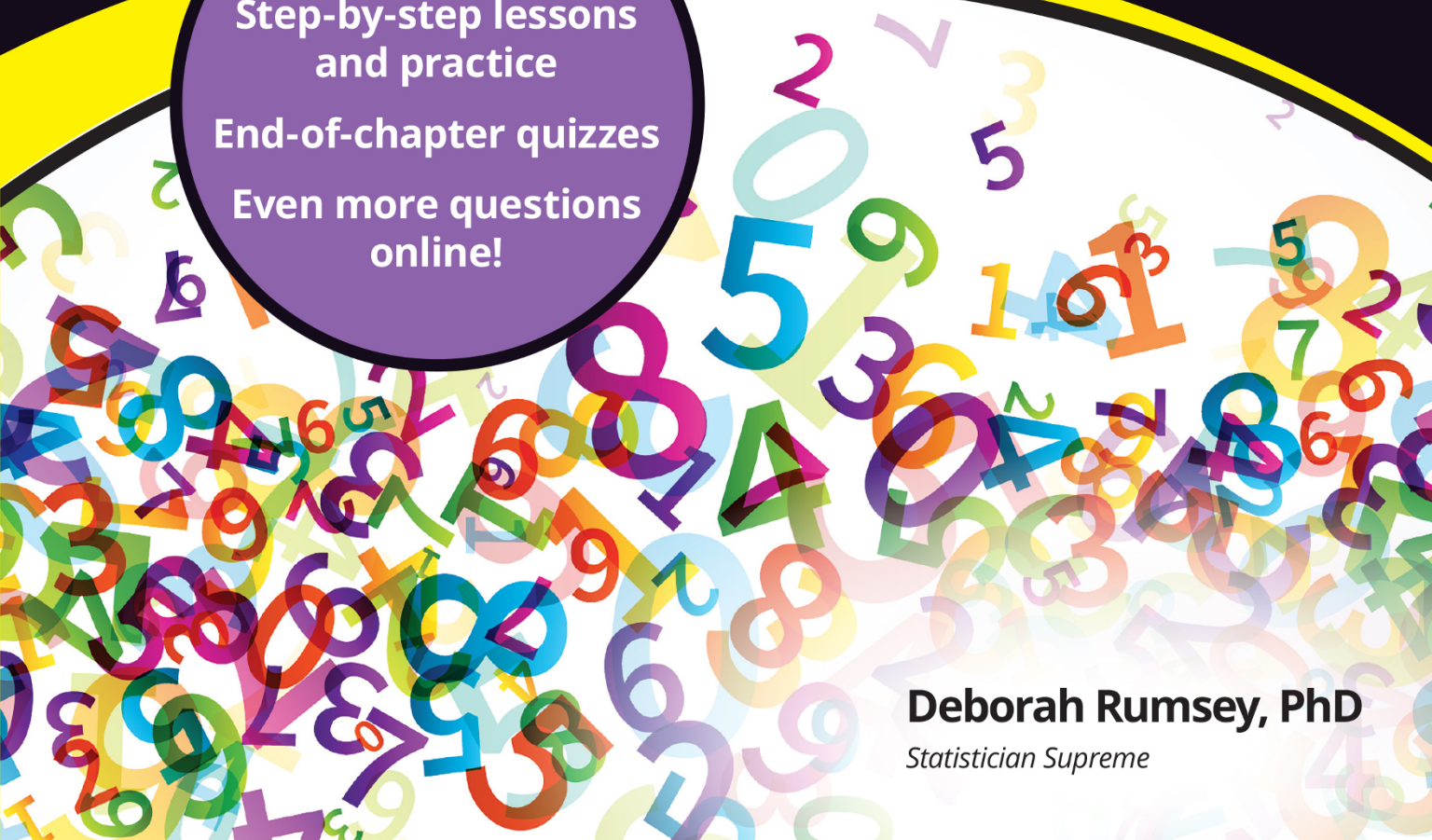
A Wiley Brand

FEATURES

Step-by-step lessons and practice

End-of-chapter quizzes

Even more questions online!



Deborah Rumsey, PhD

Statistician Supreme

This book comes with access to more content online.

Quiz yourself, track your progress,
and improve your grade!

Register your book or ebook at
www.dummies.com/go/getaccess.

Select your product, and then follow the prompts
to validate your purchase.

You'll receive an email with your PIN and instructions.

Statistics

ALL-IN-ONE

Statistics

ALL-IN-ONE

by Deborah J. Rumsey, PhD

for dummies®
A Wiley Brand

Statistics All-in-One For Dummies®

Published by: **John Wiley & Sons, Inc.,** 111 River Street, Hoboken, NJ 07030-5774, www.wiley.com

Copyright © 2023 by John Wiley & Sons, Inc., Hoboken, New Jersey

Published simultaneously in Canada

No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning or otherwise, except as permitted under Sections 107 or 108 of the 1976 United States Copyright Act, without the prior written permission of the Publisher. Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, (201) 748-6011, fax (201) 748-6008, or online at http://www.wiley.com/go/permissions.

Trademarks: Wiley, For Dummies, the Dummies Man logo, Dummies.com, Making Everything Easier, and related trade dress are trademarks or registered trademarks of John Wiley & Sons, Inc., and may not be used without written permission. All other trademarks are the property of their respective owners. John Wiley & Sons, Inc., is not associated with any product or vendor mentioned in this book.

LIMIT OF LIABILITY/DISCLAIMER OF WARRANTY: WHILE THE PUBLISHER AND AUTHORS HAVE USED THEIR BEST EFFORTS IN PREPARING THIS WORK, THEY MAKE NO REPRESENTATIONS OR WARRANTIES WITH RESPECT TO THE ACCURACY OR COMPLETENESS OF THE CONTENTS OF THIS WORK AND SPECIFICALLY DISCLAIM ALL WARRANTIES, INCLUDING WITHOUT LIMITATION ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. NO WARRANTY MAY BE CREATED OR EXTENDED BY SALES REPRESENTATIVES, WRITTEN SALES MATERIALS OR PROMOTIONAL STATEMENTS FOR THIS WORK. THE FACT THAT AN ORGANIZATION, WEBSITE, OR PRODUCT IS REFERRED TO IN THIS WORK AS A CITATION AND/OR POTENTIAL SOURCE OF FURTHER INFORMATION DOES NOT MEAN THAT THE PUBLISHER AND AUTHORS ENDORSE THE INFORMATION OR SERVICES THE ORGANIZATION, WEBSITE, OR PRODUCT MAY PROVIDE OR RECOMMENDATIONS IT MAY MAKE. THIS WORK IS SOLD WITH THE UNDERSTANDING THAT THE PUBLISHER IS NOT ENGAGED IN RENDERING PROFESSIONAL SERVICES. THE ADVICE AND STRATEGIES CONTAINED HEREIN MAY NOT BE SUITABLE FOR YOUR SITUATION. YOU SHOULD CONSULT WITH A SPECIALIST WHERE APPROPRIATE. FURTHER, READERS SHOULD BE AWARE THAT WEBSITES LISTED IN THIS WORK MAY HAVE CHANGED OR DISAPPEARED BETWEEN WHEN THIS WORK WAS WRITTEN AND WHEN IT IS READ. NEITHER THE PUBLISHER NOR AUTHORS SHALL BE LIABLE FOR ANY LOSS OF PROFIT OR ANY OTHER COMMERCIAL DAMAGES, INCLUDING BUT NOT LIMITED TO SPECIAL, INCIDENTAL, CONSEQUENTIAL, OR OTHER DAMAGES.

For general information on our other products and services, please contact our Customer Care Department within the U.S. at 877-762-2974, outside the U.S. at 317-572-3993, or fax 317-572-4002. For technical support, please visit https://hub.wiley.com/community/support/dummies.

Wiley publishes in a variety of print and electronic formats and by print-on-demand. Some material included with standard print versions of this book may not be included in e-books or in print-on-demand. If this book refers to media such as a CD or DVD that is not included in the version you purchased, you may download this material at http://booksupport.wiley.com. For more information about Wiley products, visit www.wiley.com.

Library of Congress Control Number: 2022945239

ISBN 978-1-119-90256-0 (pbk); ISBN 978-1-119-90257-7 (ebk); ISBN 978-1-119-90258-4 (ebk)

Contents at a Glance

Introduction 1

Unit 1: Getting Started with Statistics 5

CHAPTER 1: The Statistics of Everyday Life 7
CHAPTER 2: Taking Control: So Many Numbers, So Little Time 15
CHAPTER 3: Tools of the Trade 25

Unit 2: Number-Crunching Basics 47

CHAPTER 4: Crunching Categorical Data 49
CHAPTER 5: Means, Medians, and More 65
CHAPTER 6: Getting the Picture: Graphing Categorical Data 99
CHAPTER 7: Going by the Numbers: Graphing Numerical Data 125

Unit 3: Distributions and the Central Limit Theorem 167

CHAPTER 8: Coming to Terms with Probability 169
CHAPTER 9: Random Variables and the Binomial Distribution 199
CHAPTER 10: The Normal Distribution 219
CHAPTER 11: The *t*-Distribution 251
CHAPTER 12: Sampling Distributions and the Central Limit Theorem 263

Unit 4: Guesstimating and Hypothesizing with Confidence 285

CHAPTER 13: Leaving Room for a Margin of Error 287
CHAPTER 14: Confidence Intervals: Making Your Best Guesstimate 305
CHAPTER 15: Claims, Tests, and Conclusions 341
CHAPTER 16: Commonly Used Hypothesis Tests: Formulas and Examples 361

Unit 5: Statistical Studies and the Hunt for a Meaningful Relationship 389

CHAPTER 17: Polls, Polls, and More Polls 391
CHAPTER 18: Experiments and Observational Studies: Medical Breakthroughs or Misleading Results? 413
CHAPTER 19: Looking for Links: Correlation and Regression 435
CHAPTER 20: Two-Way Tables and Independence 467
APPENDIX: Tables for Reference 503

Index 513

Table of Contents

INTRODUCTION 1
About This Book 1
Foolish Assumptions 2
Icons Used in This Book 3
Beyond the Book 3
Where to Go from Here 4

UNIT 1: GETTING STARTED WITH STATISTICS 5

CHAPTER 1: The Statistics of Everyday Life 7
Statistics and the Media: More Questions than Answers? 8
Probing popcorn problems 8
Venturing into viruses 8
Comprehending crashes 9
Mulling malpractice 9
Belaboring the loss of land 10
Scrutinizing schools 10
Scanning sports 10
Banking on business news 11
Touring the travel news 12
Surveying sexual stats 12
Breaking down weather reports 13
Using Statistics at Work 13
Delivering babies — and information 13
Posing for pictures 14
Poking through pizza data 14
Statistics in the office 14

CHAPTER 2: Taking Control: So Many Numbers, So Little Time 15
Detecting Errors, Exaggerations, and Just Plain Lies 16
Checking the math 16
Uncovering misleading statistics 17
Looking for lies in all the right places 22
Feeling the Impact of Misleading Statistics 23

CHAPTER 3: Tools of the Trade 25
Thriving in a Statistical World 25
Statistics: More than Just Numbers 26
Designing Appropriate Studies 28
Surveys (Polls) 28
Experiments 29
Collecting Quality Data 30
Sample, random, or otherwise 30
Bias 32

Grabbing Some Basic Statistical Jargon . . . 33
Data . . . 33
Data set . . . 34
Variable . . . 34
Population . . . 34
Statistic . . . 35
Parameter . . . 35
Mean (Average) . . . 36
Median . . . 36
Standard deviation . . . 36
Percentile . . . 37
Standard score . . . 37
Distribution and normal distribution . . . 38
Central Limit Theorem . . . 38
z-values . . . 39
Margin of error . . . 39
Confidence interval . . . 40
Hypothesis testing . . . 41
p-values . . . 42
Statistical significance . . . 42
Correlation, regression, and two-way tables . . . 43
Drawing Credible Conclusions . . . 44
Reeling in overstated results . . . 44
Questioning claims of cause and effect . . . 44
Becoming a Sleuth, Not a Skeptic . . . 45

UNIT 2: NUMBER-CRUNCHING BASICS . . . 47

CHAPTER 4: **Crunching Categorical Data** . . . 49
Summing Up Data with Descriptive Statistics . . . 49
Crunching Categorical Data: Tables and Percents . . . 50
Counting on the frequency . . . 50
Relating with percentages . . . 52
Two-way tables: Summarizing multiple measures . . . 54
Interpreting counts and percents with caution . . . 55
Practice Questions Answers and Explanations . . . 57
Whaddya Know? Chapter 4 Quiz . . . 60
Answers to Chapter 4 Quiz . . . 62

CHAPTER 5: **Means, Medians, and More** . . . 65
Measuring the Center with Mean and Median . . . 65
Averaging out to the mean . . . 66
Splitting your data down the median . . . 68
Comparing means and medians: Histograms . . . 70
Accounting for Variation . . . 74
Reporting the standard deviation . . . 75
Being out of range . . . 78
Examining the Empirical Rule (68-95-99.7) . . . 79
Measuring Relative Standing with Percentiles . . . 83
Calculating percentiles . . . 83
Interpreting percentiles . . . 85

Gathering a five-number summary . . . 88
Exploring interquartile range . . . 89
Practice Questions Answers and Explanations . . . 90
Whaddya Know? Chapter 5 Quiz . . . 97
Answers to Chapter 5 Quiz . . . 98

CHAPTER 6: **Getting the Picture: Graphing Categorical Data** . . . 99
Take Another Little Piece of My Pie Chart . . . 100
Tallying personal expenses . . . 100
Bringing in a lotto revenue . . . 101
Ordering takeout . . . 102
Projecting age trends . . . 103
Raising the Bar on Bar Graphs . . . 108
Tracking transportation expenses . . . 108
Making a lotto profit . . . 110
Tipping the scales on a bar graph . . . 110
Pondering pet peeves . . . 111
Practice Questions Answers and Explanations . . . 117
Whaddya Know? Chapter 6 Quiz . . . 123
Answers to Chapter 6 Quiz . . . 124

CHAPTER 7: **Going by the Numbers: Graphing Numerical Data** . . . 125
Handling Histograms . . . 126
Making a histogram . . . 126
Interpreting a histogram . . . 130
Putting numbers with pictures . . . 137
Detecting misleading histograms . . . 139
Examining Boxplots . . . 143
Making a boxplot . . . 143
Interpreting a boxplot . . . 146
Tackling Time Charts . . . 152
Interpreting time charts . . . 152
Understanding variability: Time charts versus histograms . . . 153
Spotting misleading time charts . . . 153
Practice Questions Answers and Explanations . . . 159
Whaddya Know? Chapter 7 Quiz . . . 165
Answers to Chapter 7 Quiz . . . 166

UNIT 3: DISTRIBUTIONS AND THE CENTRAL LIMIT THEOREM . . . 167

CHAPTER 8: **Coming to Terms with Probability** . . . 169
A Set Notation Overview . . . 169
Noting outcomes: Sample spaces . . . 170
Noting subsets of sample spaces: Events . . . 171
Noting a void in the set: Empty sets . . . 172
Putting sets together: Unions, intersections, and complements . . . 172
Probabilities of Events Involving A and/or B . . . 173
Probability notation . . . 174
Marginal probabilities . . . 175
Union probabilities . . . 175

Intersection (joint) probabilities . . . 175
Complement probabilities . . . 176
Conditional probabilities . . . 176
Understanding and Applying the Rules of Probability . . . 178
The complement rule (for opposites, not for flattering a date) . . . 178
The multiplication rule (for intersections, not for rabbits) . . . 179
The addition rule (for unions of the nonmarital nature) . . . 180
Recognizing Independence in Multiple Events . . . 182
Checking independence for two events with the definition . . . 182
Using the multiplication rule for independent events . . . 183
Including Mutually Exclusive Events . . . 184
Recognizing mutually exclusive events . . . 184
Simplifying the addition rule with mutually exclusive events . . . 185
Distinguishing Independent from Mutually Exclusive Events . . . 186
Comparing and contrasting independence and exclusivity . . . 186
Checking for independence or exclusivity in a 52-card deck . . . 187
Avoiding Probability Misconceptions . . . 189
Predictions Using Probability . . . 190
Practice Questions Answers and Explanations . . . 192
Whaddya Know? Chapter 8 Quiz . . . 196
Answers to Chapter 8 Quiz . . . 197

CHAPTER 9: **Random Variables and the Binomial Distribution** . . . 199
Defining a Random Variable . . . 200
Discrete versus continuous . . . 200
Probability distributions . . . 202
The mean and variance of a discrete random variable . . . 202
Identifying a Binomial . . . 203
Checking binomial conditions step by step . . . 204
No fixed number of trials . . . 204
More than success or failure . . . 205
Trials are not independent . . . 205
Probability of success (p) changes . . . 205
Finding Binomial Probabilities Using a Formula . . . 207
Finding Probabilities Using the Binomial Table . . . 210
Finding probabilities for specific values of X . . . 210
Finding probabilities for X greater-than, less-than, or between two values . . . 211
Checking Out the Mean and Standard Deviation of the Binomial . . . 212
Practice Questions Answers and Explanations . . . 214
Whaddya Know? Chapter 9 Quiz . . . 216
Answers to Chapter 9 Quiz . . . 217

CHAPTER 10: **The Normal Distribution** . . . 219
Exploring the Basics of the Normal Distribution . . . 219
Meeting the Standard Normal (Z-) Distribution . . . 223
Checking out Z . . . 223
Standardizing from X to Z . . . 223
Finding probabilities for Z with the Z-table . . . 225
Finding Probabilities for a Normal Distribution . . . 227
Knowing Where You Stand with Percentiles . . . 230

Finding *X* When You Know the Percent . . . 232
Figuring out a percentile for a normal distribution . . . 232
Translating tricky wording in percentile problems . . . 234
Normal Approximation to the Binomial . . . 236
Practice Questions Answers and Explanations . . . 239
Whaddya Know? Chapter 10 Quiz . . . 248
Answers to Chapter 10 Quiz . . . 249

CHAPTER 11: **The *t*-Distribution** . . . 251
Basics of the *t*-Distribution . . . 251
Comparing the *t*- and *Z*-distributions . . . 251
Discovering the effect of variability on *t*-distributions . . . 252
Using the *t*-Table . . . 253
Finding probabilities with the *t*-table . . . 253
Figuring percentiles for the *t*-distribution . . . 256
Picking out *t**-values for confidence intervals . . . 257
Studying Behavior Using the *t*-Table . . . 258
Practice Questions Answers and Explanations . . . 259
Whaddya Know? Chapter 11 Quiz . . . 260
Answers to Chapter 11 Quiz . . . 261

CHAPTER 12: **Sampling Distributions and the Central Limit Theorem** . . . 263
Defining a Sampling Distribution . . . 264
The Mean of a Sampling Distribution . . . 265
Measuring Standard Error . . . 266
Sample size and standard error . . . 266
Population standard deviation and standard error . . . 267
Looking at the Shape of a Sampling Distribution . . . 269
Case 1: The distribution of *X* is normal . . . 269
Case 2: The distribution of *X* is not normal — Enter the Central Limit Theorem . . . 269
Finding Probabilities for the Sample Mean . . . 273
The Sampling Distribution of the Sample Proportion . . . 275
Finding Probabilities for the Sample Proportion . . . 278
Practice Questions Answers and Explanations . . . 280
Whaddya Know? Chapter 12 Quiz . . . 283
Answers to Chapter 12 Quiz . . . 284

UNIT 4: GUESSTIMATING AND HYPOTHESIZING WITH CONFIDENCE . . . 285

CHAPTER 13: **Leaving Room for a Margin of Error** . . . 287
Seeing the Importance of that Plus or Minus . . . 287
Finding the Margin of Error: A General Formula . . . 289
Measuring sample variability . . . 289
Calculating margin of error for a sample proportion . . . 291
Reporting results . . . 293
Calculating margin of error for a sample mean . . . 293
Being confident you're right . . . 294

Determining the Impact of Sample Size . . . 296
Sample size and margin of error . . . 296
Bigger isn't always (that much) better! . . . 297
Keeping margin of error in perspective . . . 297
Practice Questions Answers and Explanations . . . 300
Whaddya Know? Chapter 13 Quiz . . . 303
Answers to Chapter 13 Quiz . . . 304

CHAPTER 14: **Confidence Intervals: Making Your Best Guesstimate** . . . 305
Not All Estimates Are Created Equal . . . 306
Linking a Statistic to a Parameter . . . 306
Getting with the Jargon . . . 307
Interpreting Results with Confidence . . . 308
Zooming In on Width . . . 310
Choosing a Confidence Level . . . 310
Factoring In the Sample Size . . . 312
Counting On Population Variability . . . 313
Calculating a Confidence Interval for a Population Mean . . . 315
Case 1: Population standard deviation is known . . . 315
Case 2: Population standard deviation is unknown and/or n is small . . . 316
Figuring Out What Sample Size You Need . . . 318
Determining the Confidence Interval for One Population Proportion . . . 319
Creating a Confidence Interval for the Difference of Two Means . . . 322
Case 1: Population standard deviations are known . . . 322
Case 2: Population standard deviations are unknown and/or sample sizes are small . . . 323
Estimating the Difference of Two Proportions . . . 326
Spotting Misleading Confidence Intervals . . . 329
Practice Questions Answers and Explanations . . . 330
Whaddya Know? Chapter 14 Quiz . . . 337
Answers to Chapter 14 Quiz . . . 339

CHAPTER 15: **Claims, Tests, and Conclusions** . . . 341
Setting Up the Hypotheses . . . 342
Defining the null . . . 342
What's the alternative? . . . 343
Gathering Good Evidence (Data) . . . 343
Compiling the Evidence: The Test Statistic . . . 344
Gathering sample statistics . . . 344
Measuring variability using standard errors . . . 344
Understanding standard scores . . . 345
Calculating and interpreting the test statistic . . . 345
Weighing the Evidence and Making Decisions: *p*-Values . . . 346
Connecting test statistics and *p*-values . . . 346
Defining a *p*-value . . . 347
Calculating a *p*-value . . . 347
Making Conclusions . . . 349
Setting boundaries for rejecting H_0 . . . 350
Testing varicose veins . . . 350

Assessing the Chance of a Wrong Decision . . . 352
Making a false alarm: Type I errors . . . 353
Missing out on a detection: Type II errors . . . 353
Practice Questions Answers and Explanations . . . 356
Whaddya Know? Chapter 15 Quiz . . . 359
Answers to Chapter 15 Quiz . . . 360

CHAPTER 16: **Commonly Used Hypothesis Tests: Formulas and Examples** . . . 361
Testing One Population Mean . . . 362
Handling Small Samples and Unknown Standard Deviations: The *t*-Test . . . 363
Putting the *t*-test to work . . . 364
Relating *t* to *Z* . . . 365
Handling negative *t*-values . . . 365
Examining the not-equal-to alternative . . . 365
Drawing conclusions using the critical value . . . 366
Testing One Population Proportion . . . 368
Comparing Two (Independent) Population Averages . . . 371
Case 1: Difference of two population means when population standard deviations are known . . . 371
Case 2: Difference of two population means when population standard deviations are unknown . . . 374
Testing for an Average Difference (The Paired *t*-Test) . . . 375
Comparing Two Population Proportions . . . 378
Practice Questions Answers and Explanations . . . 382
Whaddya Know? Chapter 16 Quiz . . . 387
Answers to Chapter 16 Quiz . . . 388

UNIT 5: STATISTICAL STUDIES AND THE HUNT FOR A MEANINGFUL RELATIONSHIP . . . 389

CHAPTER 17: **Polls, Polls, and More Polls** . . . 391
Recognizing the Impact of Polls . . . 392
Getting to the source . . . 392
Surveying what's hot . . . 393
Impacting lives . . . 394
Behind the Scenes: The Ins and Outs of Surveys . . . 395
Planning and designing a survey . . . 396
Selecting the sample . . . 399
Carrying out a survey . . . 402
Interpreting results and finding problems . . . 405
Practice Questions Answers and Explanations . . . 408
Whaddya Know? Chapter 17 Quiz . . . 411
Answers to Chapter 17 Quiz . . . 412

CHAPTER 18: **Experiments and Observational Studies: Medical Breakthroughs or Misleading Results?** . . . 413
Boiling Down the Basics of Studies . . . 414
Looking at the lingo of studies . . . 414
Observing observational studies . . . 415
Examining experiments . . . 415

Designing a Good Experiment . . . 417
Designing the experiment to make comparisons . . . 417
Selecting the sample size . . . 419
Choosing the subjects . . . 421
Making random assignments . . . 421
Controlling for confounding variables . . . 422
Respecting ethical issues . . . 423
Collecting good data . . . 424
Analyzing the data properly . . . 425
Interpreting Experiment Results . . . 428
Making appropriate conclusions . . . 428
Making informed decisions . . . 429
Practice Questions Answers and Explanations . . . 431
Whaddya Know? Chapter 18 Quiz . . . 433
Answers to Chapter 18 Quiz . . . 434

CHAPTER 19: **Looking for Links: Correlation and Regression** . . . 435
Picturing a Relationship with a Scatterplot . . . 436
Making a scatterplot . . . 436
Interpreting a scatterplot . . . 437
Quantifying Linear Relationships Using the Correlation . . . 440
Calculating the correlation . . . 440
Interpreting the correlation . . . 441
Examining properties of the correlation . . . 443
Working with Linear Regression . . . 445
Figuring out which variable is *X* and which is *Y* . . . 445
Checking the conditions . . . 446
Calculating the regression line . . . 446
Interpreting the regression line . . . 449
Putting it all together: The regression line for the crickets . . . 451
Making Proper Predictions . . . 453
Checking the conditions . . . 453
Staying in-bounds . . . 454
Regression Analysis: Understanding the Output . . . 456
Residing with Residuals . . . 457
Explaining the Relationship: Correlation versus Cause and Effect . . . 458
Practice Questions Answers and Explanations . . . 460
Whaddya Know? Chapter 19 Quiz . . . 465
Answers to Chapter 19 Quiz . . . 466

CHAPTER 20: **Two-Way Tables and Independence** . . . 467
Organizing a Two-Way Table . . . 468
Setting up the cells . . . 469
Figuring the totals . . . 469
Interpreting Two-Way Tables . . . 472
Singling out variables with marginal distributions . . . 472
Examining all groups — a joint distribution . . . 476
Comparing groups with conditional distributions . . . 478

Checking Independence and Describing Dependence 483
Checking for independence 484
Describing a dependent relationship 486
Cautiously Interpreting Results 489
Checking for legitimate cause and effect 489
Projecting from sample to population 490
Making prudent predictions 491
Resisting the urge to jump to conclusions 491
Practice Questions Answers and Explanations 492
Whaddya Know? Chapter 20 Quiz 498
Answers to Chapter 20 Quiz 500

APPENDIX: **Tables for Reference** 503
The *Z*-Table 503
The *t*-Table 506
The Binomial Table 508

INDEX 513

Introduction

You get hit with an incredible amount of statistical information on a daily basis. You know what I'm talking about: charts, graphs, tables, and headlines that talk about the results of the latest poll, survey, experiment, or other scientific study. The purpose of this book is to develop and sharpen your skills in sorting through, analyzing, and evaluating all that info, and to do so in a clear, fun, and pain-free way with tons of opportunities to practice. You also gain the ability to decipher and make important decisions about statistical results (for example, the results of the latest medical studies), while being ever aware of the ways that people can mislead you with statistics. And you see how to do it right when it's your turn to design the study, collect the data, crunch the numbers, and/or draw the conclusions.

This book is also designed to help those of you who are looking to get a solid foundation in introductory statistics or those taking a statistics class and wanting some backup. You'll gain a working knowledge of the big ideas of statistics and gather a boatload of tools and tricks of the trade that'll help you get ahead of the curve, especially for taking exams.

This book is chock-full of real examples from real sources that are relevant to your everyday life — from the latest medical breakthroughs, crime studies, and population trends to the latest U.S. government reports. I even address a survey on the worst cars of the millennium! By reading this book, you'll understand how to collect, display, and analyze data correctly and effectively, and you'll be ready to critically examine and make informed decisions about the latest polls, surveys, experiments, and reports that bombard you every day. You will even find out how to use crickets to gauge temperature!

You will also get to climb inside the minds of statisticians to see what's worth taking seriously and what isn't to be taken so seriously. After all, with the right skills and knowledge, you don't have to be a professional statistician to understand introductory statistics. You can be a data guru in your own right.

About This Book

This book departs from traditional statistics texts, references, supplemental books, and study guides in the following ways:

- It includes practical and intuitive explanations of statistical concepts, ideas, techniques, formulas, and calculations found in an introductory statistics course.
- It shows you clear and concise step-by-step procedures that explain how you can intuitively work through statistics problems.
- It features interesting real-world examples relating to your everyday life and workplace.

- It contains plenty of excellent practice problems crafted in a straightforward manner to lead you down the path of success.
- It offers not only answers, but also clear, complete explanations of the answers. Explanations help you know exactly how to approach a problem, what information you need to solve it, and common problems you need to avoid.
- It includes tips, strategies, and warnings based on my vast experience with students of all backgrounds and learning styles.
- It gives you upfront and honest answers to your questions like, "What does this really mean?" and "When and how will I ever use this?"

As you work your way through the lessons and problems in this book, you should be aware of four conventions that I've used.

- **Dual use of the word *statistics:*** In some situations, I refer to statistics as a subject of study or as a field of research, so the word is a singular noun. For example, "Statistics is really quite an interesting subject." In other situations, I refer to statistics as the plural of *statistic,* in a numerical sense. For example, "The most commonly used statistics are the mean and the standard deviation."
- **Use of the word *data:*** You're probably unaware of the debate raging among statisticians about whether the word *data* should be singular ("data is ") or plural ("data are "). It got so bad that one group of statisticians had to develop two versions of a statistics T-shirt: "Messy Data Happens" and "Messy Data Happen." I go with the plural version of the word *data* in this book.
- **Use of the term *standard deviation:*** When I use the term *standard deviation,* I mean *s,* the sample standard deviation. (When I refer to the population standard deviation, I let you know.)
- **Use of *italics:*** I use *italics* to let you know a new statistical term is appearing on the scene. Look for a definition accompanying its first appearance.

Foolish Assumptions

I don't assume that you've had any previous experience with statistics, other than the fact that you're a member of the general public who gets bombarded every day with statistics in the form of numbers, percents, charts, graphs, "statistically significant" results, "scientific" studies, polls, surveys, experiments, and so on.

What I do assume is that you can do some of the basic mathematical operations and understand some of the basic notation used in algebra, such as the variables x and y, summation signs (Σ), taking the square root, squaring a number, and so on. If you need to brush up on your algebra skills, check out *U Can Algebra I For Dummies* by Mary Jane Sterling (Wiley).

I don't want to mislead you: You do encounter formulas in this book, because statistics does involve a bit of number crunching. But don't let that worry you. I take you slowly and carefully

through each step of any calculations you need to do, explaining things both with notation and without. I also provide practice questions for you to work so you can become familiar and comfortable with the calculations and make them your own.

Icons Used in This Book

You'll see the following five icons throughout the book:

EXAMPLE

Each example is a stat question based on the discussion and explanation, followed by a solution. Work through these examples, and then refer to them to help you solve the practice problems that follow them as well as the quiz questions at the end of the chapter.

REMEMBER

This icon points out important information that you need to focus on. Make sure you understand this information fully before moving on. You can skim through these icons when reading a chapter to make sure you remember the highlights.

TIP

Tips are hints that can help speed you along when answering a question. See whether you find them useful when working on practice problems.

WARNING

This icon flags common mistakes that students make if they're not careful. Take note and proceed with caution!

YOUR TURN

When you see this icon, it's time to put on your thinking cap and work out a few practice problems on your own. The answers and detailed solutions are available so you can feel confident about your progress.

Beyond the Book

In addition to the material in the print or e-book you're reading right now, this book also comes with a handy online Cheat Sheet. Use it when you need a quick refresher on a formula or the next step in conducting a hypothesis test. To get this Cheat Sheet, simply go to `www.dummies.com` and type **Statistics All in One For Dummies Cheat Sheet** in the Search box.

You'll also have access to online quizzes related to each chapter, beginning with Unit 2, Chapter 4. These quizzes provide a whole new set of problems for practice and confidence-building. To access the quizzes, follow these simple steps:

1. **Register your book or ebook at Dummies.com to get your PIN. Go to `www.dummies.com/go/getaccess`.**
2. **Select your product from the drop-down list on that page.**
3. **Follow the prompts to validate your product, and then check your email for a confirmation message that includes your PIN and instructions for logging in.**

If you do not receive this email within two hours, please check your spam folder before contacting us through our Technical Support website at `http://support.wiley.com` or by phone at 877-762-2974.

Now you're ready to go! You can come back to the practice material as often as you want — simply log on with the username and password you created during your initial login. No need to enter the access code a second time.

Your registration is good for one year from the day you activate your PIN.

Where to Go from Here

This book is written in such a way that you can start anywhere and still be able to understand what's going on. So you can take a peek at the table of contents or the index, look up the information that interests you, and flip to the page listed. However, if you have a specific topic in mind and are eager to dive into it, here are some directions:

- To work on interpreting graphs, charts, means or medians, and the like, head to Unit 2.
- To find info on the normal, *Z*-, *t*-, or binomial distributions or the Central Limit Theorem, see Unit 3.
- To focus on confidence intervals and hypothesis tests of all shapes and sizes, flip to Unit 4.
- To delve into surveys, experiments, regression, and two-way tables, see Unit 5.

Or if you aren't sure where you want to start, start with Chapter 1 for the big picture and then plow your way through the rest of the book. Ready, set, go!

1 Getting Started with Statistics

In This Unit . . .

CHAPTER 1: **The Statistics of Everyday Life**

Statistics and the Media: More Questions than Answers?
Using Statistics at Work

CHAPTER 2: **Taking Control: So Many Numbers, So Little Time**

Detecting Errors, Exaggerations, and Just Plain Lies
Feeling the Impact of Misleading Statistics

CHAPTER 3: **Tools of the Trade**

Thriving in a Statistical World
Statistics: More than Just Numbers
Designing Appropriate Studies
Collecting Quality Data
Grabbing Some Basic Statistical Jargon
Drawing Credible Conclusions
Becoming a Sleuth, Not a Skeptic

IN THIS CHAPTER

- Raising questions about statistics you see in everyday life
- Encountering statistics in the workplace

Chapter 1
The Statistics of Everyday Life

Today's society is completely taken over by numbers. Numbers are everywhere you look, from billboards showing the on-time statistics for a particular airline, to sports shows discussing the Las Vegas odds for upcoming football games. The evening news is filled with stories focusing on crime rates, the expected life span of junk-food junkies, and the president's approval rating. On a normal day, you can run into 5, 10, or even 20 different statistics (with many more on election night). Just by reading a Sunday newspaper all the way through, you come across literally hundreds of statistics in reports, advertisements, and articles covering everything from soup (how much does an average person consume per year?) to nuts (almonds are known to have positive health effects — what about other types of nuts?).

In this chapter we discuss the statistics that often appear in your life and work, and talk about how statistics are presented to the general public. After reading this chapter, you'll realize just how often the media hits you with numbers and how important it is to be able to unravel the meaning of those numbers. Like it or not, statistics are a big part of your life. So, if you can't beat 'em, join 'em. And if you don't want to join 'em, at least try to understand 'em.

Statistics and the Media: More Questions than Answers?

Open a newspaper and start looking for examples of articles and stories involving numbers. It doesn't take long before those numbers begin to pile up. Readers are inundated with results of studies, announcements of breakthroughs, statistical reports, forecasts, projections, charts, graphs, and summaries. The extent to which statistics occur in the media is mind-boggling. You may not even be aware of how many times you're hit with numbers nowadays.

This section looks at just a few examples from one Sunday paper's worth of news that I read the other day. When you see how frequently statistics are reported in the news without providing all the information you need, you may find yourself getting nervous, wondering what you can and can't believe anymore. Relax! That's what this book is for — to help you sort out the good information from the bad (the chapters in Unit 2 give you a great start on that).

Probing popcorn problems

The first article I came across that dealt with numbers was "Popcorn plant faces health probe," with the subheading: "Sick workers say flavoring chemicals caused lung problems." The article describes how the Centers for Disease Control (CDC) expressed concern about a possible link between exposure to chemicals in microwave popcorn flavorings and some cases of fixed obstructive lung disease. Eight people from one popcorn factory alone contracted this lung disease, and four of them were awaiting lung transplants.

According to the article, similar cases were reported at other popcorn factories. Now, you may be wondering, what about the folks who eat microwave popcorn? According to the article, the CDC found "no reason to believe that people who eat microwave popcorn have anything to fear." (Stay tuned.) The next step is to evaluate employees in more depth, including conducting surveys to determine health and possible exposures to the flavoring chemicals, checks of lung capacity, and detailed air samples. The question here is: How many cases of this lung disease constitute a real pattern, compared to mere chance or a statistical anomaly? (You find out more about this in Chapter 15.)

Venturing into viruses

A second article discussed a recent cyber attack: A wormlike virus made its way through the Internet, slowing down web browsing and email delivery around the world. How many computers were affected? The experts quoted in the article said that 39,000 computers were infected, and they in turn affected hundreds of thousands of other systems.

Questions: How did the experts get that number? Did they check each computer out there to see whether it was affected? The fact that the article was written less than 24 hours after the attack suggests the number is a guess. Then why say 39,000 and not 40,000 — to make it seem less like a guess? To find out more on how to guesstimate with confidence (and how to evaluate someone else's numbers), see Chapter 14.

Comprehending crashes

Next in the paper was an alert about the soaring number of motorcycle fatalities. Experts said that the *fatality rate* — the number of fatalities per 100,000 registered vehicles — for motorcyclists has been steadily increasing, as reported by the National Highway Traffic Safety Administration (NHTSA). In the article, many possible causes for the increased motorcycle death rate were discussed, including age, gender, size of engine, whether the driver had a license, alcohol use, and state helmet laws (or lack thereof). The report was very comprehensive, showing various tables and graphs with the following titles:

- Motorcyclists killed and injured, and fatality and injury rates by year, per number of registered vehicles, and per millions of vehicle miles traveled
- Motorcycle rider fatalities by state, helmet use, and blood alcohol content
- Occupant fatality rates by vehicle type (motorcycles, passenger cars, light trucks), per 10,000 registered vehicles and per 100 million vehicle miles traveled
- Motorcyclist fatalities by age group
- Motorcyclist fatalities by engine size (displacement)
- Previous driving records of drivers involved in fatal traffic crashes by type of vehicle (including previous crashes, DUI convictions, speeding convictions, and license suspensions and revocations)

This article was very informative and provided a wealth of detailed information regarding motorcycle fatalities and injuries in the U.S. However, the onslaught of so many tables, graphs, rates, numbers, and conclusions can be overwhelming and confusing and lead you to miss the big picture. With a little practice, and help from Unit 2, you'll be better able to sort out graphs, tables, and charts and all the statistics that go along with them. For example, some important statistical issues come up when you see rates versus counts (such as death rates versus number of deaths). As I address in Chapter 2, counts can give you misleading information if they're used when rates would be more appropriate.

Mulling malpractice

Further along in the newspaper was a report about a recent medical malpractice insurance study: Malpractice cases affect people in terms of the fees doctors charge and the ability to get the healthcare they need. The article indicates that one in five Georgia doctors have stopped doing risky procedures (such as delivering babies) because of the ever-increasing malpractice insurance rates in the state. This is described as a "national epidemic" and a "health crisis" around the country. Some brief details of the study are included, and the article states that of the 2,200 Georgia doctors surveyed, 2,800 of them — which they say represent about 18 percent of those sampled — were expected to stop providing high-risk procedures.

Wait a minute! That can't be right. Out of 2,200 doctors, 2,800 don't perform the procedures, and that is supposed to represent 18 percent? That's impossible! You can't have a bigger number on the top of a fraction, and still have the fraction be under 100 percent, right? This is one

of many examples of errors in media reporting of statistics. So what's the real percentage? There's no way to tell from the article. Chapter 4 nails down the particulars of calculating these kinds of statistics so you can know what to look for and immediately tell when something's not right.

Belaboring the loss of land

In the same Sunday paper was an article about the extent of land development and speculation across the United States. Knowing how many homes are likely to be built in your neck of the woods is an important issue to get a handle on. Statistics are given regarding the number of acres of farmland being lost to development each year. To further illustrate how much land is being lost, the area is also listed in terms of football fields. In this particular example, experts said that the mid-Ohio area is losing 150,000 acres per year, which is 234 square miles, or 115,385 football fields (including end zones). How do people come up with these numbers, and how accurate are they? And does it help to visualize land loss in terms of the corresponding number of football fields? I discuss the accuracy of data collected in more detail in Chapter 17.

Scrutinizing schools

The next topic in the paper was school proficiency — specifically, whether extra school sessions help students perform better. The article stated that 81.3 percent of students in this particular district who attended extra sessions passed the writing proficiency test, whereas only 71.7 percent of those who didn't participate in the extra school sessions passed it. But is this enough of a difference to account for the $386,000 price tag per year? And what's happening in these sessions to cause an improvement? Are students in these sessions spending more time just preparing for those exams rather than learning more about writing in general? And here's the big question: Were the participants in the extra sessions student volunteers who may be more motivated than the average student to try to improve their test scores? The article didn't say.

Studies like this appear all the time, and the only way to know what to believe is to understand what questions to ask and to be able to critique the quality of the study. That's all part of statistics! The good news is, with a few clarifying questions, you can quickly critique statistical studies and their results. Chapter 18 helps you do just that.

Scanning sports

The sports section is probably the most numerically jam-packed section of the newspaper. Beginning with game scores, the win/loss percentages for each team, and the relative standing for each team, the specialized statistics reported in the sports world are so deep that they require wading boots to get through. For example, basketball statistics are broken down by team, by quarter, and by player. For each player, you get minutes played, field goals, free throws, rebounds, assists, personal fouls, turnovers, blocks, steals, and total points.

Who needs to know this stuff, besides the players' mothers? Apparently, many fans do. Statistics are something that sports fans can never get enough of and players often can't stand to hear about. Stats are the substance of water-cooler debates and the fuel for armchair quarterbacks around the world.