Statistics Unplugged

y Caldwell

C147

Statistics Unplugged

Sally Caldwell

Southwest Texas State University









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



Publisher: Vicki Knight

Assistant Editor: Jennifer Wilkinson Editorial Assistant: Monica Sarmiento

Technology Project Manager: Darin Derstine

Marketing Manager: Lori Grebe
Marketing Assistant: Laurel Anderson
Advertising Project Manager: Brian Chaffee
Signing Representative: Richard Colangelo

Project Manager, Editorial Production: Paula Berman

Print/Media Buyer: Kris Waller

COPYRIGHT © 2004 Wadsworth, a division of Thomson Learning, Inc. Thomson LearningTM is a trademark used herein under license.

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 but not limited to photocopying, recording, taping, Web distribution, information networks, or information storage and retrieval systems—without the written permission of the publisher.

Printed in Canada
2 3 4 5 6 7 07 06 05 04 03

For more information about our products, contact us at:

Thomson Learning
Academic Resource Center
1-800-423-0563

For permission to use material from this text, contact us by:

Phone: 1-800-730-2214 **Fax:** 1-800-730-2215 **Web:** http://www.thomsonrights.com

Library of Congress Control Number: 2002115694

ISBN 0-534-52113-4

Permissions Editor: Sarah Harkrader

Production Service and Compositor: Scratchgravel

Publishing Services
Text Designer: John Edeen
Copy Editor: Margaret C. Tropp
Illustrator: Gregory Draus
Cover design: Cheryl Carrington
Cover art: Kevin Tolman

Text and Cover Printer: Transcontinental/Louiseville

Wadsworth/Thomson Learning 10 Davis Drive Belmont, CA 94002-3098 USA

Thomson Learning

5 Shenton Way #01-01 UIC Building Singapore 068808

Australia

Nelson Thomson Learning 102 Dodds Street South Melbourne, Victoria 3205 Australia

Canada

Nelson Thomson Learning 1120 Birchmount Road Toronto, Ontario M1K 5G4 Canada

Europe/Middle East/Africa

Thomson Learning High Holborn House 50/51 Bedford Row London WC1R 4LR United Kingdom

Latin America

Thomson Learning Seneca, 53 Colonia Polanco 11560 Mexico D.F. Mexico

Spain/Portugal

Paraninfo Calle/Magallanes, 25 28015 Madrid, Spain In memory of Geoff Wood, whose mom wrote the book on friendship



About the Author

Sally Caldwell earned her Ph.D. in Sociology from the University of North Texas. The author of *Romantic Deception* (Adams Media, 2000), Caldwell focuses her primary research interest on the topic of deception in social relationships. Caldwell resides in a small village in the hill country of south central Texas and serves on the faculty of the Department of Sociology at Southwest Texas State University.

Preface

The idea behind this book came from my students, after I watched countless semesters unfold in a predictable fashion. The scene repeats itself each year in a classroom populated largely with panic-stricken students facing their first formal encounter with the field of statistical analysis. My passion for the subject matter is usually enough to quickly make a connection with some of the students, but certainly not all of them. A real connection, I've discovered, requires constant attention to the students' perspective, a willingness to respect the roadblocks (real or imaginary) that exist in their minds.

Over time, I've learned that there are two major obstacles for a lot of students. First, there's the *fear of the formula* factor, something that seems to plague many students. The second obstacle is the *so what* scenario: the tendency for many students to question the relevance of the subject matter—and why they have to take the course in the first place. I believe there's a way to overcome these roadblocks, and that way is what I've attempted to lay out in this book. I emphasize the *logic* behind statistical analysis and focus on an *intuitive* understanding that I believe lies within virtually every student. When the formulas take a back seat, magical things can happen. The world of statistical analysis becomes interesting, inviting, and (on occasion) enjoyable. When the magic takes hold, the exploration goes well beyond statistical techniques and into the larger realm of scientific inquiry. In the process, the subject matter is made relevant.

Acknowledgments

The idea for any book remains just that—merely an idea—until the words hit the page. In the case of this book, the words would never have made it so far were it not for the support and encouragement of a lot of people. The individuals who came to be involved in this project are numerous. Many, I suspect, were unaware along the way of just how important their involvement was.

At the university, certain individuals in strategic positions had the talent for making things happen. To Ann Marie Ellis and Lawrence Estaville, Jr., I owe a particular debt of gratitude. They are, in a word or two, door openers of the first order. To the chair of my department, Susan Day, I owe daily thanks for her tireless efforts to carve out a supportive and sustaining environment—one with a true sense of community and commitment to intellectual endeavors. My thanks go also to certain colleagues who suffered through my countless commentaries about this project. Barbara Trepagnier, for example, was always in

her office down the hall and available when I chose to drop in (uninvited), just to prattle on about the project. My car pooling pals, Sue Johnson and Pam Showalter, willingly gave me feedback as we drove through the Wimberley Valley on our way to the university.

When it came to getting words onto paper, the support was endless. First, Richard Coleangelo, Wadsworth campus representative, took my idea up the ladder and into the talented hands of Vicki Knight, Editor and Publisher at Wadsworth. Vicki knows about words, so she knows what I mean when I say words cannot express my gratitude for her constant support, humor, and patience throughout. The list goes on. Paula Berman, Production Editor at Wadsworth, and Monica Sarmiento, Editorial Assistant, teamed up to give valuable guidance and keep me on task. Vernon Boes brought his incredible talents to the cover, and Anne and Greg Draus of Scratchgravel Publishing Services painstakingly provided the finishing touches (graciously taking an endless stream of calls from me). Finally, Sara Smith, student assistant without equal, kept me organized, focused, and optimistic. Without hesitation, she would read chapter after chapter, over and over again. My debt to her can never be fully repaid.

I thank the reviewers who read through the manuscript and offered valuable comments: James Knapp, Southeastern Oklahoma State University; Paul Ansfield, University of Wisconsin Oshkosh; Lora Schlewitt-Haynes, University of Northern Colorado; Ida Mirzaie, Ohio State University; Charles Harrington, University of Southern Indiana; Steve Weinert, Cuyamaca Community College; J. Oliver Williams, North Carolina State University; Holly Straub, University of South Dakota; Faye Plascak-Craig, Marian College; Michael Hurley, George Mason University; and Susan Nolan, Seton Hall University. In addition, the tables included in Appendixes A–E and Appendixes H and I were adapted from tables constructed by Victor Bissonnette of Berry College. His generosity in supplying the materials is greatly appreciated.

On the personal side of the ledger, an army of people always understood my endless preoccupation with the project. So to Gary Hartling, John Friedli, Nancy Middlebrook, Ellen Lewis, Ron Lenamon, Marilee Wood, Tevis Grinstead, Maria Canizales, Walter Brown, Steve Klepfer, Jimmie Robinson, Sam and Barbara Dawkins, Kay Cole, and Liz Sumter, I say a heartfelt thank-you for the many priceless moments that never failed to remind me of the true joys of life.

Finally, my thanks to Suzan Abughazaleh and Eric Groves, two very special people who are a constant presence in my life, even though there are miles in between. Like my parents, who never ran short with their encouragement, Suzan and Eric have always understood my pursuits. The mere thought of them always brings forth a smile.

Sally Caldwell

If you would like to give us your opinion and suggestions regarding this text, please go to http://psychology.wadsworth.com/caldwell1e/ and click on "Evaluation Form" in the left menu to complete the student's evaluation form.

Brief Contents

1	The What and How of Statistics 4	
2	Describing Data and Distributions 17	
3	The Shape of Distributions 46	
4	The Normal Curve 64	
5	Four Fundamental Concepts 84	
6	Confidence Intervals 98	
7	Hypothesis Testing With a Single Sample 137	
8	Hypothesis Testing With Two Samples (Mean Difference and Difference of Means) 165	
9	Beyond the Null Hypothesis 188	
10	Analysis of Variance 206	
11	The Chi-Square Test 238	
12	Correlation and Regression 262	
Appendix A	Table of Areas Under the Normal Curve (Distribution of Z) 29	5
Appendix B	Family of t Distributions (Two-Tailed Test) 297	
Appendix C	Family of t Distributions (One-Tailed Test) 298	
Appendix D	Distribution of F (.05 Level of Significance) 299	
Appendix E	Distribution of F (.01 Level of Significance) 300	
Appendix F	Distribution of Q (.05 Level of Significance) 301	
Appendix G	Distribution of Q (.01 Level of Significance) 302	
Appendix H	Critical Values for Chi-Square (χ^2) 303	
Appendix I	Critical Values of r (Correlation Coefficient) 304	
	Answers to Chapter Problems 305	
	References 310	
	Glossary 311	
	Index 316	

 $oldsymbol{0}$ Introduction: Methods, Material, and Moments to Remember

Contents

0	Introduction: Methods, Material, and Moments to Remember	1
1	The What and How of Statistics	4
	A World of Information 4 Levels of Measurement 6 Samples and Populations 8 The Purposes of Statistical Analysis 11 Descriptive Statistics 11 Inferential Statistics 12 Chapter Summary 14 Some Other Things You Should Know 14 Key Terms 15 Chapter Problems 15	
2	Describing Data and Distributions Measures of Central Tendency 18 The Mean 18 The Median 21 The Mode 24 Measures of Variability or Dispersion 25 The Range 26 Deviations From the Mean 26 The Mean Deviation 29 The Variance 31 The Standard Deviation 34 n Versus n - 1 40 Chapter Summary 42 Some Other Things You Should Know 43 Key Terms 44 Chapter Problems 44	17

7		
/3	The Shape of Distributions	46
	The Basic Elements 47 Beyond the Basics: Comparisons and Conclusions 50 A Special Curve 54 Chapter Summary 61 Some Other Things You Should Know 61 Key Terms 62 Chapter Problems 62	
/ 4	The Normal Curve	64
	Real-World Normal Curves 65 Into the Theoretical World 68 The Table of Areas Under the Normal Curve 71 Finally, an Application 76 Chapter Summary 81 Some Other Things You Should Know 82 Key Terms 82 Chapter Problems 82	
5	Four Fundamental Concepts	84
	Fundamental Concept #1: Random Sampling 85 Fundamental Concept #2: Sampling Error 87 Fundamental Concept #3: The Sampling Distribution of Sample Means Fundamental Concept #4: The Central Limit Theorem 91 Chapter Summary 95 Some Other Things You Should Know 95 Key Terms 96	90
	Chapter Problems 96	
6	Confidence Intervals Confidence Interval for the Mean 99 Confidence Interval for the Mean With σ Known 100 An Application 100 Reviewing Z Values 102 Z Values and the Width of the Interval 103	98

ix

Bringing in the Standard Error of the Mean 104 The Relevance of the Central Limit Theorem and the Standard Error 107 Confidence and Interval Width 109 A Brief Recap 112 Confidence Interval for the Mean With σ Unknown 113 Estimating the Standard Error of the Mean The Family of t Distributions 116 The Table for the Family of t Distributions 118 An Application 122 A Final Comment About the Interpretation of a Confidence Interval for the Mean 124 A Final Comment About Z Versus t 124 Confidence Intervals for Proportions 125 An Application 126 Margin of Error 128 Chapter Summary 130 Some Other Things You Should Know 132 Key Terms 133 Chapter Problems 133

7 Hypothesis Testing With a Single Sample 137

Setting the Stage 138

A Hypothesis as a Statement of Your Expectations:

The Case of the Null Hypothesis 139

Single Sample Test With σ Known 141

Refining the Null and Phrasing It the Right Way 142

The Logic of the Test 143

Applying the Test 145

Levels of Significance, Critical Values, and the Critical Region

But What If We're Wrong? 151

Another Example 153

Single Sample Test With σ Unknown 156

Applying the Test 157

Some Variations on a Theme 159

Chapter Summary 160

Some Other Things You Should Know 161

Key Terms 161

Chapter Problems 162

8	Hypothesis Testing With Two Samples (Mean Difference and Difference of Means)	165
	Related Samples 166 The Logic of the Test 168 The Null Hypothesis 170 Combining the Logic and the Null 170 The Standard Error of the Mean Difference 171 Applying the Test 172 Interpreting the Results 172 Some Additional Examples 173	
	Independent Samples 175 The Logic of the Test 175 The Null Hypothesis 178 Combining the Logic and the Null 178 The Standard Error of the Difference of Means 178 Applying the Test 181 Interpreting the Results 182 Some Additional Examples 182	
	Chapter Summary 184 Some Other Things You Should Know 184 Key Terms 185 Chapter Problems 185	
9	Beyond the Null Hypothesis Research or Alternative Hypotheses 189 One-Tailed and Two-Tailed Test Scenarios 191 Testing a Nondirectional Research Hypothesis 192 Testing a Directional Research Hypothesis 194 Power and Effect 198 Chapter Summary 202 Some Other Things You Should Know 202 Key Terms 203 Chapter Problems 203	188

10 Analysis of Variance

206

The Logic of ANOVA 208
From Curves to Data Distributions 210
The Different Means 211
From Different Means to Different Types of Variation 212

The Application 215 Calculating the Within-Groups Sum of Squares (SS_W) 217 Calculating the Between-Groups Sum of Squares (SS_B) 219 From Sums of Squares to Estimates of Variance 221 Calculating the F Ratio 226 The Interpretation 227 Interpretation of the F Ratio 227 Post Hoc Testing 229 Chapter Summary 233 Some Other Things You Should Know 234 Key Terms 235 Chapter Problems 235	
The Chi-Square Test	238
The Chi-Square Test of Independence 239 The Logic of the Test 240 A Focus on the Departure From Chance 244 The Null Hypothesis 245 The Application 246 The Formula 250 The Calculation 250 Conclusion and Interpretation 251 The Chi-Square Goodness-of-Fit Test 253 The Null Hypothesis 253 The Application 254 The Formula and the Calculation 255 Conclusion and Interpretation 256 Chapter Summary 257 Some Other Things You Should Know 257 Key Terms 258 Chapter Problems 258	
Correlation and Regression	262
Scatter Plots 263 Linear Associations: Direction and Strength 264 Other Types of Association 267 Correlation Analysis 268 Two Variables: X and Y 269 The Logic of Correlation 271	

The Null Hypothesis 215

11

12

The Formula for Pearson's r 272
Application 275
Interpretation 277
An Additional Step: Testing the Null 279
Conclusion and Interpretation 280
Regression Analysis 281
An Application 281
The Logic of Prediction and the Line of Best Fit 283
The Regression Equation 285
The Standard Error of the Estimate 288
Chapter Summary 290
Some Other Things You Should Know 290
Key Terms 291
Chapter Problems 292

Appendix A Table of Areas Under the Normal Curve (Distribution of *Z*) 295 **Appendix B** Family of *t* Distributions (Two-Tailed Test) 297 **Appendix C** Family of *t* Distributions (One-Tailed Test) 298 **Appendix D** Distribution of *F* (.05 Level of Significance) 299 **Appendix E** Distribution of *F* (.01 Level of Significance) 300 **Appendix F** Distribution of Q (.05 Level of Significance) 301 **Appendix G** Distribution of Q (.01 Level of Significance) 302 **Appendix H** Critical Values for Chi-Square (χ^2) **Appendix I** Critical Values of *r* (Correlation Coefficient) 304 Answers to Chapter Problems 305 References 310 Glossary 311 Index 316

Introduction: Methods, Material, and Moments to Remember

Statistics, Quantitative Methods, Statistical Analysis—words, phrases, and course titles that can shake the confidence of nearly any student. Let me put your mind at ease right away. Your experience with statistics doesn't have to be a horror story. In fact, your experience with statistics can be an enjoyable one—a venture into a new way of thinking and looking at the world. It's all a matter of how you approach the material.

Having taught statistics to legions of undergraduate students, I've spent a lot of time trying to understand how students react to the material and why they react the way they do. In the process, I've developed my own approach to the subject matter, and that's what I've tried to lay out in this book. As we get started, let me tell you a little more about what to expect as you work your way through this book.

First, let me explain my method. I'm committed to the idea that the subject matter of statistics can be made understandable, but I'm also convinced that it takes a method based on *repetition*. Important ideas and concepts can be introduced, but they have to be reintroduced and reemphasized if a student is to get the connection between one concept and the next. Repetition—that's the method I've used in this book, so you should be prepared for that.

At times you may wonder why you're rereading material that was emphasized at an earlier point. Indeed, you'll likely start muttering "not that again!" If that happens, enjoy the moment. It signals that you're beginning to develop a sense of familiarity with the central concepts.

I've also tried to incorporate *simplicity* into the method—particularly in the examples I've used. Some examples will probably strike you as extremely simplistic—particularly the examples that are based on just a few cases and the ones that involve numbers with small values. I trust that simplistic examples won't offend you. The goal here is to cement a learning process, not to master complicated mathematical operations. That brings us to the matter of the material you're going to encounter.

As it turns out, it's often the logic behind statistics that proves to be the key to success or failure. You can be presented with formulas—simple or complex—and you can, with enough time and commitment, memorize a string of them. All of that is well and good, but your ability to grasp the logic behind the formulas is a different matter altogether. I'm convinced that it's impossible to truly understand what statistics is all about unless you understand the logic behind the procedures. Consequently, it's the *logic* that I've tried to emphasize in this book.

Indeed, it's safe to say that numbers and formulas have taken a back seat in this book. Of course you'll encounter some formulas and numbers, but that's not where the emphasis is. Make no mistake about it—the emphasis in this book is on the conceptual basis behind the calculations.

There's one other thing about the material that deserves comment. Like it or not, the traditional approach to learning new material may come up short when you want to learn about statistical analysis. The reason is a simple one: The field of statistics is very different from other subjects you've studied in the past.

If, for example, you were to taking a course to learn a foreign language, you'd probably figure out the goal of the course fairly early. You'd quickly sense that you'd be learning the basics of grammar and vocabulary, trying to increase your command of both over time. I suspect you'd have a similar experience if you signed up for a history course. You'd quickly sense that you were being introduced to names, dates, places, and overall context with the goal of increasing your understanding of the how and why behind events.

Unfortunately, the field of statisticalal analysis doesn't fit that learning model very well. You may be able to immediately sense where you're going in a lot of courses, but that's not necessarily the case in the field of statistics. In fact, my guess is that a command of statistical analysis is probably best achieved when you're willing to go along for the ride without really knowing at first where you're going. A statement like that is close to heresy in the academic world, so let me explain.

There is an end game to statistical analysis. People use statistical analysis to describe information and to carry out research in an objective, quantifiable way. Indeed, the realm of statistical analysis is fundamental to scientific inquiry. But the eventual application of statistical analysis requires that you first have a firm grasp of some highly abstract concepts. You can't even begin to appreciate the very special way in which scientists pose research questions if you don't have the conceptual background.

For a lot of students (indeed, most students, I suspect), it's a bit much to tackle concepts and applications at the same time. The process has to be broken down into two parts—first the conceptual understanding, and then the applications. And that's the essence of my notion that you're better off if you don't focus at the outset on where you're going. Concentrate on the conceptual basis first. Allow yourself to become totally immersed in an abstract, conceptual world, without any thought about direct applications. In my judgment, that's the best way to conquer the field of statistical analysis.

If you're the sort of student who demands an immediate application of concepts—if you don't have much tolerance for abstract ideas—let me strongly suggest that you lighten up a bit. If you're going to master statistics—even at the introductory level—you'll have to open your mind to the world of abstract thinking.

Toward that end, let me tell you in advance that I'll occasionally ask you to take a moment to seriously think about one notion or another. Knowing students the way I do, I suspect there's a chance (if only a small chance) that you'll ignore my suggestion and just move ahead. Let me warn you. The approach of trying to get from Point A to Point B as quickly as possible usually doesn't work in the field of statistics. When the time comes to really think about a concept, take whatever time is necessary.

Indeed, many of my students eventually come to appreciate what I mean when I tell them that a particular concept or idea requires a "dark room moment." In short, some statistical concepts or ideas are best understood if contemplated in a room that is totally dark and void of any distractions. Those should become your moments to remember. I'm totally serious about that, so let me explain why.

Many statistical concepts are so abstract that a lot of very serious thought is required if you really want to understand them. Moreover, many of those abstract concepts turn out to be central to the statistical way of reasoning. Simply reading about the concepts and telling yourself that you'll remember what they're all about won't do it. And that's the purpose behind a dark room moment.

If I could give you a single key to the understanding of statistics, it would be this: Take the dark room moments seriously. Don't be impatient, and don't think a few dark room experiences are beneath your intellectual dignity. If I tell you that this concept or that idea may require a dark room moment, heed the warning. Head for a solitary environment—a private room, or even a closet. Turn out the lights, if need be, and undertake your contemplation in a world void of distractions. You may be amazed how it will help your understanding of the topic at hand.

Now, having said all of that as background, it's time to get started. Welcome to the world of statistics—in this case, *Statistics Unplugged*!