Statistical Methods for Communication Science

ANDREW F. HAYES
The Ohio State University


LAWRENCE ERLBAUM ASSOCIATES, PUBLISHERS
2005 Mahwah, New Jersey London

Senior Acquisitions Editor: Linda Bathgate
Assistant Editor: Karin Wittig Bates
Cover Design: Kathryn Houghtaling Lacey
Textbook Production Manager: Paul Smolenski
Text and Cover Printer: Hamilton Printing Company

This book was typeset by the author using the LaTeX language and delivered to the publisher as camera-ready copy.

A web page supporting this book can be found at
http://www.comm.ohio-state.edu/ahayes/smcs/

Lawrence Erlbaum Associates, Inc., Publishers
10 Industrial Avenue
Mahwah, New Jersey 07430

Library of Congress Cataloging-in-Publication Data
Hayes, Andrew F.

Statistical methods for communication science/Andrew F. Hayes
p. cm.
Includes bibliographic references and index
ISBN 0-8058-5487-8 (hard cover: alk. paper)
1. Communication—Statistical Methods. 2. Communication—Research
I. Hayes, Andrew F. II. Title
P93.7.H39 2005
302.2′02 dc–22 2005040570

Books published by Lawrence Erlbaum Associates are printed on acid-free paper, and their bindings are chosen for strength and durability.

Printed in the United States of America
10 9 8 7 6 5 4 3 2 1

To **Carole**

CONTENTS

Contents **vii**

Preface **xv**

1 Statistics and Communication Science **1**
1.1 Welcome . . . 1
1.2 Why Do Science? . . . 2
1.3 Assumptions and Philosophies of Scientific Investigation . . . 5
1.4 Building Your Statistical Vocabulary . . . 8
1.5 The Role of Statistics in Scientific Investigation . . . 12
1.6 Summary . . . 14

2 Fundamentals of Measurement **16**
2.1 Measurement Concepts . . . 16
2.1.1 Methods of Measurement . . . 17
2.1.2 Operationalization . . . 18
2.1.3 Levels of Measurement . . . 20
2.1.4 Measurement Precision . . . 23
2.1.5 Qualitative Data versus Quantitative Measurement . . . 24
2.2 Measurement Quality . . . 24
2.2.1 Reliability of Measurement . . . 24
2.2.2 Validity of Measurement . . . 25
2.3 Summary . . . 30

3 Sampling **31**
3.1 Population Inference . . . 31
3.1.1 The *Literary Digest* Poll: Population Inference Gone Awry . . . 33
3.1.2 Population Inference Through Representativeness . . . 35
3.2 Sampling Methods . . . 36
3.2.1 Nonprobability Sampling . . . 36
3.2.2 Probability Sampling . . . 38
3.3 Is Nonprobability Sampling Really So Bad? . . . 41
3.4 Summary . . . 44

4 Data Description and Visualization **45**
4.1 Graphical and Tabular Descriptions of Data . . . 45
4.1.1 Frequency Tables . . . 45
4.1.2 The Histogram . . . 48
4.1.3 Describing the Shape of a Distribution . . . 49
4.2 Measures of Central Tendency . . . 51
4.2.1 The Mode . . . 51
4.2.2 The Median . . . 52
4.2.3 The Arithmetic Mean . . . 53
4.2.4 Choosing a Measure of Central Tendency . . . 54
4.3 Measures of Variation . . . 56
4.3.1 The Range and Interquartile Range . . . 57
4.3.2 The Standard Deviation . . . 57
4.3.3 The Variance . . . 59
4.4 Quantifying Skewness and Kurtosis . . . 59
4.5 Another Graphical Tool: The Box Plot . . . 60
4.6 Standardization . . . 61
4.7 Describing Association Between Quantitative Variables . . . 63
4.7.1 Pearson's Coefficient of Correlation . . . 64
4.7.2 Alternative Measures of Association . . . 68
4.7.3 Cautions When Interpreting Correlation . . . 70
4.8 Visualizing Correlation: The Scatterplot . . . 73
4.9 Descriptively Comparing Groups . . . 76
4.10 Data Screening and Missing Data . . . 78
4.11 Introducing Some Common Symbolic Notation . . . 80
4.12 Summary . . . 81

5 Fundamentals of Probability **83**
5.1 Defining Probability . . . 84
5.2 Laws of Probability . . . 86
5.2.1 The Additive Law of Probability . . . 86
5.2.2 The Multiplicative Law of Probability . . . 88
5.3 Probability Distributions . . . 90
5.3.1 The Binomial Probability Distribution . . . 91
5.3.2 The Normal Probability Distribution . . . 94
5.3.3 Chebychev's Theorem . . . 99
5.4 Random Variables and Expected Values . . . 99
5.5 Summary . . . 102

6 Assessing and Quantifying Reliability **103**
6.1 Classical Test Theory . . . 105
6.1.1 Partioning Measurements into Their Components . . . 105
6.1.2 The Definition of Reliability Under Classical Test Theory . . . 108
6.2 Estimating the Reliability of Quantitative Measurements . . . 110
6.2.1 Estimating Reliability From Repeated Measurements Over Time . . . 110
6.2.2 Estimating Reliability From Internal Consistency of Indicator Scores . . . 111
6.2.3 Reliability of Method or of Measurement? . . . 117
6.3 Reliability of Subjective Categorical Judgments . . . 118

6.3.1 Holsti's Method . . . 121
6.3.2 Correcting for Chance Agreement: Scott's π and Cohen's κ . . . 121
6.3.3 Using an Agreement Index . . . 126
6.4 How High is High Enough? . . . 128
6.5 Summary . . . 128

7 Parameter Estimation 130
7.1 The Theory of Estimation . . . 130
7.1.1 The Sampling Distribution . . . 131
7.1.2 Properties of the Sampling Distribution of the Sample Mean . . . 135
7.1.3 Deriving the Probability of Obtaining Certain Sample Means . . . 141
7.2 Interval Estimates . . . 143
7.2.1 The Confidence Interval . . . 144
7.2.2 A More Realistic Approach to Interval Estimation . . . 146
7.2.3 The Relationship Between Confidence and Precision . . . 149
7.2.4 Computing the Probability of a Sample Mean, Revisited . . . 150
7.2.5 Interval Estimates Derived From Samples of Convenience . . . 151
7.3 Estimating a Population Proportion . . . 151
7.4 Bootstrapping a Confidence Interval . . . 154
7.5 Summary . . . 157

8 Hypothesis Testing Concepts 158
8.1 Hypothesis Testing Steps . . . 160
8.1.1 Step 1: Translate the Research Hypothesis or Question into a Statistical Hypothesis . . . 160
8.1.2 Step 2: Quantify the Obtained Result . . . 162
8.1.3 Step 3: Derive the p-value . . . 163
8.1.4 Step 4: Decide Between the Null and Alternative Hypothesis . . . 165
8.1.5 Step 5: Interpret the Result of the Test in Substantive Terms . . . 167
8.2 Testing a Hypothesis About a Population Proportion . . . 167
8.2.1 Testing a Nondirectional ("Two-tailed") Hypothesis . . . 168
8.2.2 Testing a Directional ("One-tailed") Hypothesis . . . 173
8.3 Decision Errors, Power, and Validity . . . 175
8.3.1 Type I, Type II, and Type III errors . . . 175
8.3.2 The Validity and Power of a Statistical Test . . . 178
8.4 Hypothesis Test or Confidence Interval? . . . 179
8.5 Summary . . . 182

9 Testing a Hypothesis About a Single Mean 183
9.1 The One-Sample t test . . . 183
9.1.1 Testing a Directional Hypothesis About a Single Mean . . . 185
9.1.2 Testing a Nondirectional Hypothesis . . . 191
9.1.3 Conducting the One-Sample t test With a Computer . . . 193
9.1.4 Statistical Assumptions . . . 194
9.1.5 Large Versus Small Samples . . . 195
9.1.6 Confidence Intervals . . . 195
9.1.7 Bootstrapping the p-value . . . 196
9.2 Comparing the Means of Paired Responses . . . 196
9.2.1 The Paired-Samples t test . . . 197
9.2.2 Paired-Sample Inference From Nonrandom Samples . . . 201

9.3 Summary 209

10 Comparing Two Independent Groups 210
10.1 The Independent Groups t test 211
10.1.1 The Pooled Variance Approach 213
10.1.2 The Welch-Satterthwaite Approach 217
10.1.3 The Conditional Decision Rule 218
10.1.4 The Behrens-Fisher Problem 219
10.1.5 Violations of the Normality Assumption 220
10.1.6 Confidence Intervals for the Mean Difference 221
10.1.7 Bootstrapping Confidence Intervals and p-values 221
10.1.8 Effect Size 223
10.2 Testing for Group Differences in Variability 224
10.2.1 Levene's Test 225
10.2.2 The Brown-Forsythe Test 226
10.2.3 The F-ratio Test: A Test to Avoid 228
10.3 Comparing Two Groups from Nonrandom Samples 228
10.3.1 The Random Assignment Model of Chance 233
10.3.2 Inference Without Random Sampling or Random Assignment . . 237
10.4 Thinking Clearly About Inference 241
10.5 Comparing Two Independent Proportions 242
10.6 Summary 242

11 Some Tests for Categorical Variables 244
11.1 Testing a Hypothesis About a Frequency Distribution 244
11.1.1 Testing the Hypothesis of Equal Relative Frequencies 245
11.1.2 Testing the Fit of Observed Frequencies to Any Expectation . . 250
11.1.3 Statistical Assumptions 251
11.1.4 Alternative Tests for the $k = 2$ Case 252
11.1.5 Testing Goodness of Fit for Process Inference 254
11.2 Association Between Two Categorical Variables 256
11.2.1 Testing for Independence in a 2×2 Crosstabulation 257
11.2.2 Quantifying Strength of Association in a 2×2 Table 262
11.2.3 An Alternative Test: The Z test for Independent Proportions . . 264
11.2.4 Another Example With a 3×2 Crosstabulation 265
11.2.5 Testing for Independence in Larger Tables 269
11.2.6 Quantifying Association in a Table of Any Size 269
11.3 Summary 270

12 Simple Linear Regression 271
12.1 The Simple Linear Regression Model 274
12.1.1 The Simple Regression Line 275
12.1.2 The Least Squares Criterion 279
12.1.3 The Standard Error of Estimation 282
12.1.4 The Standardized Regression Equation 283
12.1.5 Variance Explained by a Regression Model 285
12.1.6 More on Residuals 287
12.1.7 The Dangers of Extrapolating Away From the Data 289
12.2 Population Inference in Linear Regression 289
12.2.1 Testing a Hypothesis About the Population Regression Weight . 290

12.2.2 Confidence Intervals for β . . . 293
12.2.3 Reframing Inference In Terms of the Population Correlation . . . 293
12.2.4 Statistical Assumptions . . . 296
12.3 Inference in Nonrandom Samples: The Permutation Test . . . 301
12.4 Detecting Influential Cases . . . 305
12.4.1 Distance, Leverage, and Influence . . . 305
12.4.2 Influence as Change in the Model When a Case is Excluded . . . 306
12.5 Summary . . . 309

13 Multiple Linear Regression 310
13.1 The Multiple Regression Model . . . 312
13.2 Quantifying and Testing for Multivariate Association . . . 314
13.2.1 R^2 as an Index of Multivariate Association . . . 314
13.2.2 Average Estimation Error . . . 319
13.2.3 Population Inferences About Multivariate Association . . . 319
13.2.4 Adjusted R^2 . . . 322
13.3 Partial Association . . . 323
13.3.1 Experimental Versus Statistical Control . . . 323
13.3.2 Mechanisms That Can Produce Association Between Two Variables . . . 325
13.3.3 Measures of Partial Association . . . 328
13.3.4 Partial Versus Semipartial Correlation . . . 334
13.3.5 The Standardized Partial Regression Coefficient . . . 336
13.3.6 Inference for Measures of Partial Association . . . 337
13.3.7 Confidence Intervals for Measures of Partial Association . . . 339
13.3.8 Assumptions for Statistical Inference and Detecting Influential Cases . . . 340
13.4 Setwise Partial Association and Hierarchical Regression . . . 342
13.4.1 Setwise Partial and Semipartial Correlations . . . 343
13.4.2 Calculating SR^2 and PR^2 . . . 344
13.4.3 Inference for Measures of Setwise Partial Association . . . 345
13.4.4 Interpreting Regression Statistics for the Predictor Variables . . . 347
13.5 Modeling and Testing for Nonlinearity . . . 349
13.5.1 The Quadratic Regression Model . . . 350
13.5.2 Interpreting the Coefficients in a Quadratic Regression Model . . . 352
13.6 Miscellaneous Issues in Multiple Regression . . . 353
13.6.1 A Reminder About Causality . . . 353
13.6.2 Number of Cases Required for a Regression Analysis . . . 354
13.6.3 Multicollinearity . . . 354
13.6.4 Selecting Predictor Variables to Include in a Regression Model . . . 357
13.6.5 The "Relative Importance" of Predictor Variables . . . 359
13.6.6 The Effects of Measurement Error . . . 361
13.6.7 Categorical, Ordinal, and Bounded Outcome Variables . . . 363
13.7 Summary . . . 365

14 Single Factor Analysis of Variance 366
14.1 Analysis of Variance . . . 367
14.1.1 Partitioning the Outcome Variable into its Sources of Variation . . . 368
14.1.2 Total, Between-, and Within-Group Variability in Y . . . 370
14.1.3 The F ratio . . . 371

14.1.4 Underlying Statistical Theory . . . 373
14.1.5 Statistical Assumptions of ANOVA . . . 374
14.1.6 Revisiting the Pooled Variance t test . . . 376
14.1.7 Quantifying Effect Size . . . 376
14.1.8 Why Not Multiple t tests? . . . 378
14.1.9 Inference with Nonrandom Samples . . . 380
14.2 Pairwise Mean Comparisons . . . 384
14.2.1 The Multiple Test Problem Resurfaces . . . 385
14.2.2 The Bonferroni Correction . . . 386
14.2.3 Holm's Sequential Rejection Method . . . 387
14.2.4 The Games-Howell Method . . . 389
14.2.5 Using a Pooled Error Term . . . 389
14.3 Focused Contrasts . . . 391
14.3.1 Focused t tests . . . 391
14.3.2 Contrast Coefficients . . . 392
14.3.3 Scheffe's Test . . . 397
14.4 ANOVA as a Special Case of Multiple Regression . . . 399
14.4.1 Coding a Categorical Variable for a Regression Analysis . . . 399
14.4.2 Testing the Omnibus Null Hypothesis Using Regression . . . 400
14.4.3 Interpreting the Regression Model . . . 401
14.5 Some Controversies in the Comparison of Multiple Groups . . . 402
14.5.1 Planned Versus Unplanned Comparisons: To Correct or Not? . . . 402
14.5.2 Are We A Bit Fickle About The Multiple Test Problem? . . . 404
14.5.3 Do We Really Need ANOVA? . . . 405
14.6 Summary . . . 406

15 Analysis of Covariance: ANOVA With Statistical Controls 407
15.1 Analysis of Covariance as Multiple Regression . . . 408
15.1.1 Conducting an ANCOVA With Regression . . . 409
15.1.2 Partitioning Variability in Y in ANCOVA . . . 410
15.1.3 Measures of Effect Size . . . 414
15.1.4 Adjusted Means . . . 416
15.1.5 Focused Contrasts Between Adjusted Means . . . 417
15.1.6 Statistical Assumptions . . . 418
15.1.7 Multiple Covariates . . . 418
15.2 Analysis of Covariance in Experimental Designs . . . 421
15.2.1 Controlling for Variables Unaffected by the Manipulation . . . 421
15.2.2 Controlling for Variables Affected by the Manipulation . . . 424
15.2.3 Mediation . . . 425
15.3 Summary . . . 427

16 Interaction 428
16.1 Interaction in Communication Research and Theory . . . 429
16.2 Factorial Analysis of Variance . . . 433
16.2.1 Partitioning Variance in Y in a Balanced Factorial Design . . . 435
16.2.2 Main and Interaction Effects . . . 438
16.2.3 The Regression Equivalent of a 2×2 Factorial ANOVA . . . 440
16.2.4 Factorial ANOVA and Unbalanced Designs . . . 441
16.2.5 Probing an Interaction in a Factorial ANOVA . . . 446
16.2.6 Quantifying Effect Size . . . 449

16.3 Moderated Multiple Regression 451
16.3.1 Interaction Between a Dichotomous and a Quantitative Variable 452
16.3.2 Interaction Between Two Quantitative Variables 458
16.3.3 Interaction Between a Quantitative and a Multicategorical Variable 461
16.3.4 Mean Centering of Predictors 465
16.3.5 Differences In Regression Coefficients vs. Differences in Correlations 468
16.3.6 Statistical Control of Covariates 469
16.3.7 Required Terms in a Moderated Multiple Regression Model . . . 469
16.3.8 The Pitfalls of Subgroup Analyses to Test Moderation 470
16.4 Simplifying the Hunt for Interactions 471
16.5 Why Not to Categorize Prior to Testing for Interaction 473
16.5.1 Classification Errors 474
16.5.2 Smaller Effect Sizes and Reduced Statistical Power 475
16.5.3 Spurious Statistical Significance 477
16.5.4 Artifactual Failures to Replicate Findings 477
16.5.5 Is Categorization Ever Sensible? 479
16.6 Summary 479

Appendices **483**
A: Table of Right-Tail Normal Probabilities 483
B: Table of Right-Tail Critical t Values 484
C: Table of Right-Tail Critical χ^2 Values 486
D1: Table of Critical F Ratios, $p \leq .05$ 487
D2: Table of Critical F Ratios, $p \leq .01$ 489

References **491**

Index **507**

Preface

Why do we need yet another book on statistical methodology when there are so many books already available? When I started writing this book, I entertained this question myself and, obviously, decided that another book was needed. My primary motivation for writing this book is that, prior to its publication, no statistical methods book on the market was tailored to the field of communication. Typically we send students to other departments to get statistical training—psychology, sociology, or statistics, for instance. This is a shame because there is plenty of evidence that suggests that people learn better when information is personalized. Of course, when planning and teaching my courses I could have simply adopted a book written for another field or for the general social sciences and then use communication examples during lectures. I tried this for a couple of years and found that it didn't work well. I couldn't find a single book that students liked or that didn't focus excessively on concepts that I felt weren't particularly relevant to the way that communication scientists conduct their business. Without a statistical methods book written for the field of communication, I found it difficult to engage students in the material and maximize their success in acquiring what may be some of the more difficult material that burgeoning communication scientists need to understand—and understand thoroughly.

So that is why I wrote this book. In the spirit of these reasons, I have tried to couch most of my examples in this book in terms of communication research and questions in the hope that this will cognitively involve students (and faculty, and practitioners) in the field more and make them see the relevance, indeed the tremendous importance of statistics to the field of communication and to success as a communication scientist. I wrote this book thinking in particular of the reader who has some background in research methodology and perhaps even some prior exposure to introductory statistics. However, even without this background, anyone should be able to read and understand the material in these pages. Although this book is suitable as a text, my hope is that the writing and examples are clear enough that any motivated reader can come away having learned something useful without the formal guidance provided in a classroom context.

I have always believed that although a good understanding of the mathematical basis of statistics is helpful to mastery of statistical methodology, it is by no means necessary. Perhaps several years or so ago I wouldn't have said this because hand computation predominated statistics education. But the age of high-speed, low-cost computers is here. The result of this widely available technology is an increasing use of complex statistical procedures because they are available to all at, literally, the touch

of a button. Of course, the danger of the convenience and user-friendly nature of statistical software is the temptation to use a statistical procedure with which you may not be familiar merely because the computational obstacles no longer exist. There is no substitute for conceptual understanding because it is necessary to make sense of what a computer tells you. This book is motivated by a desire to bring that conceptual understanding to everyone interested and motivated, regardless of their mathematical sophistication.

There is a CD that comes with this book as well as a web page supporting the book that I believe you will find useful. The URL is

http://www.comm.ohio-state.edu/ahayes/smcs/

Using the data files on the CD and the web page (available in multiple formats), you should be able to replicate most of the analyses reported in this book. The web page contains some files mentioned in this book that are not on the CD as well as additional documents that you will find useful that were developed after this book went to print.

To the Student

One of the student's greatest barriers to mastering statistics is fear of mathematics. Many students just lock up with anxiety when they are asked to do any computation, and this anxiety typically interferes with the ultimate goal of conceptual understanding. I hope you will not let this happen to you. Statistics has a bad reputation that is perpetuated by the way most statistical methods books are written. In this book, I depart company with many of my book-writing colleagues by minimizing the computational aspects of statistical methods while maximizing their conceptual presentation. You will find as you read through this book that it is possible to understand conceptually how the statistical procedures discussed in this book are used without completely understanding their mathematical basis. Many mathematicians I know can barely add and subtract without the help of a calculator. Their brilliance resides in their ability to solve problems using their knowledge of mathematics concepts and theory. Because computers are used for statistical computations, in my opinion, your time is best used by developing an understanding of the concepts presented here.

This is not to say that mathematics is irrelevant to complete mastery of statistics. Mathematics is the language of statistics. If you do want to study statistics formally, you eventually will have to come to terms with some complex mathematical operations and the mathematical basis of statistical theory. My assumption is that if you are reading this book, your primary goal is to understand how to use statistics intelligently rather than understand the mathematics behind the statistics. An additional assumption I make is that if you are interested in the mathematics, you will seek that information on your own. There are many good books similar to this one but with a greater focus on mathematics that will undoubtedly serve you well as a supplement to this book.

To the Instructor

As a science educator, you may have a great interest in if not a total passion (such as I do) for the nitty-gritty of statistics. On the other hand, most likely your students don't share this passion. Many of them will be taking your course because they have to. And many of your students probably have developed a fear of statistics, perhaps because of bad experiences they've had in the past. And most certainly, mathematics

is something that the typical communication student would prefer to avoid. To be sure, one or two students in an average size class will really enjoy the elegance and subtle details of mathematics and statistical theory. But those students are in the minority, and so it makes no sense to design a course with only them in mind. But one thing all students in your class have in common is a need to use statistics. Thus, a class and text needs to teach students how to use statistics to accomplish their research objectives. A student's interests are better served, in my opinion, by teaching them how to use statistical methods intelligently rather than trying to turn them into miniature statisticians. (Those who have such an interest or ability will undoubtedly seek out more information on their own). My approach in this book is one where the concepts are presented in such a way that students, regardless of their background or motivations, will acquire the skills needed to use statistics.

I believe that professional methodologists such as myself spend a whole lot of time thinking, writing, and arguing about things that for the user of statistics rarely matter that much. I believe that when teaching statistical methodology, one should focus on the things that matter much or most of the time rather than spending time on things that matter rarely. What does matter to the student is to understand how to use their statistical knowledge. So you will find that many of the lively and interesting controversies and debates in the statistical methodology literature are given little to no treatment in this book (although I do occasionally make reference to that literature) because those controversies and debates just won't matter to how most students end up using statistics.

I had the student and not the professional methodologist in mind when I wrote this book. The end result is that rather than trying to satisfy everyone by talking a little bit about everything, I emphasize certain things that I think are more important while under emphasizing or completely ignoring others, knowing that the instructor who feels that important things are missing or who prefers to spend more time than I on certain topics can supplement the book with additional readings. Many good sources of supplementary reading are available. I have found the *Quantitative Applications in the Social Sciences* series published by Sage Publications most helpful both as an educator and as a professional methodologist. These small volumes are ideal supplements because they provide further detail while still being both readable and friendly on the wallet.

The advent of high-speed desktop computers has revolutionized the way that statisticians think about statistics and inference. Many of the procedures communication scientists use are old and, frankly, a bit outdated. To be sure, they work relatively well, but they are a bit klunky and conceptually or philosophically ill-suited to how communication research is actually done. For example, although some communication researchers do conduct research using random samples, most researchers do not. Yet almost all statistical methods communication researchers use are based on what I call the "population model" of inference. According to this model, the goal of statistics is to give the researcher a means of inferring some unknown characteristic of a population by randomly sampling from that population, as if we are all closet public opinion pollsters. But we aren't. Furthermore, many of the methods used by communication scientists make questionable assumptions. Modern and computationally intensive methods of data analysis, such as randomization tests and bootstrapping, are largely unknown to communication scientists but conceptually better suited to the way we conduct research. To be sure, classical methods are important to understand, and I still emphasize them in this book. But the time is right to introduce some of these

modern developments into the introductory classroom. You will find discussions of some of these new conceptualizations of statistical inference and what "chance" means scattered at appropriate places throughout the book. If you choose not to cover these materials in your course, no harm done. But I strongly encourage you to consider exposing your students to some of these exciting even if seemingly (but not actually) unorthodox ways of thinking about statistical inference.

Acknowledgments

Many people have contributed to this book either directly or indirectly. Rather than list them all, let me simply say thank you, knowing that you know who you are. But a few people whose contributions were most profound should be acknowledged. First, I want to acknowledge the patience of several years of graduate students in the M.A. and Ph.D. programs in communication at The Ohio State University. They suffered through various drafts of these chapters and somehow managed to make sense of their contents in spite of the typos, poor grammar, and verbal belches. Next, Richard Darlington needs acknowledgment because he profoundly influenced the way I think about data analysis. I entered Ph.D. study at Cornell University thinking I was pretty smart when it came to statistical methods. After exposure to Dick's thinking and writing, I eventually realized how little I actually did know. I'm sure he and others familiar with his work will see his influence on me in these pages. Robert Witte also influenced me during my undergraduate study at San Jose State University in ways that are hard to overstate. Rather than scaring me away from statistics, he made the topic interesting and lively. Linda Bathgate and the production staff at Lawrence Erlbaum Associates contributed in an important way by giving me the freedom and a sufficiently large page budget to produce the kind of book I wanted to produce. I've learned that a book is never complete—you simply run out of time. But had I been forced to cut corners, this book would have turned out differently than I had originally envisioned it. I also thank my wife Carole and several anonymous reviewers who read and reread earlier drafts. They contributed tremendously to the production of this book, and each suggested modification improved the final product. Carole, my son Conor, and my daughter Julia have also contributed through their support, encouragement, and patience over the years it took to complete this book. Finally, the hard work of everyone who has been involved in the production of the LaTeX system must be acknowledged for their selflessness in providing their code free of charge for the world to use.

Andrew F. Hayes
Columbus, Ohio
April 2005