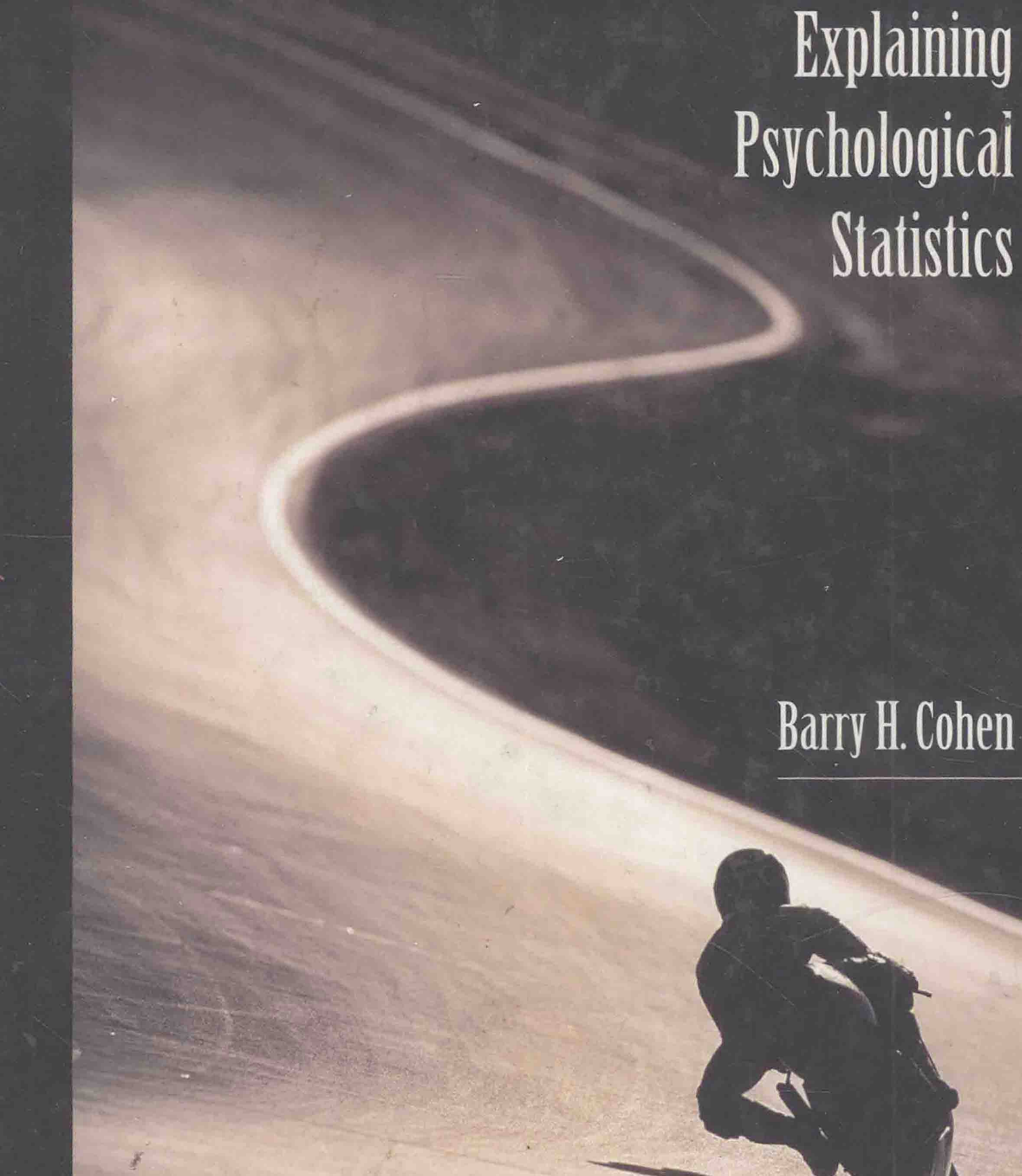
Explaining Psychological Statistics

Barry H. Cohen



EXPLAINING PSYCHOLOGICAL STATISTICS

Barry Cohen
New York University

Brooks/Cole Publishing Company
I(T)P™ An International Thomson Publishing Company
Pacific Grove • Albany • Bonn • Boston • Cincinnati • Detroit • London • Madrid • Melbourne
Mexico City • New York • Paris • San Francisco • Singapore • Tokyo • Toronto • Washington

Sponsoring Editor: *Jim Brace-Thompson*
Marketing Team: *Gay Meixel, Romi Fineroff*
Marketing Representative: *Connie Jirovsky*
Editorial Assistant: *Jodi Hermans, Terry Thomas*
Production Coordinator: *Kirk Bomont*
Project Management: *Lifland et al., Bookmakers*
Manuscript Editor: *Quica Ostrander*
Interior Design: *Quica Ostrander*
Interior Illustration: *Gail Magin*
Cover Design: *Roy R. Neuhaus*
Cover Photo: *Anton Want/Tony Stone Images*
Art Coordinator: *Quica Ostrander*
Indexer: *Denise Throckmorton*
Typesetting: *Weimer Graphics*
Cover Printing: *Color Dot*
Printing and Binding: *Quebecor/Fairfield*

For more information, contact:

BROOKS/COLE PUBLISHING COMPANY
511 Forest Lodge Road
Pacific Grove, CA 93950
USA

International Thomson Publishing Europe
Berkshire House 168-173
High Holborn
London WC1V 7AA
England

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

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

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

International Thomson Publishing GmbH
Königswinterer Strasse 418
53227 Bonn
Germany

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

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

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

Library of Congress Cataloging-in-Publication Data

Cohen, Barry H., [date]
Explaining psychological statistics / Barry H. Cohen.
p. cm.
Includes bibliographical references and index.
ISBN 0-534-20076-1
1. Psychometrics. 2. Psychology—Mathematical models.
3. Statistics—Study and teaching (Higher) I. Title.
BF39.C56 1995
150'.1'5195—dc20 95-35752
CIP

DEDICATION

To the memory of
Rose and Charles
and
Fannie and Louis,
without whose journey to America
this book would not have been possible,
and
to Leona

PREFACE TO THE INSTRUCTOR

This text evolved from my lecture notes for a one-semester statistics course (Intermediate Psychological Statistics) required of students in the M.A. psychology program at New York University. Many of my students were former English, theater, music, or arts majors headed for doctoral programs in clinical psychology. Most of these students were very bright but had very weak mathematical backgrounds. This text reflects my efforts to teach sophisticated statistical concepts to highly intelligent students who either had had no mathematical training since high school or could scarcely remember anything of their undergraduate statistics course, even if they had taken it recently. In preparing the text, I decided to include all of the introductory material that would be required if this text were used for an undergraduate course. At the same time, enough advanced material has been included to make this text suitable for a course on the master's level.

Order of Topics

The topics follow the usual order for an introductory course, with a few exceptions. First, correlation (and regression) is not included among the descriptive statistics. Correlation coefficients are so often used to test hypotheses in psychological research, especially in the "softer" areas, that it seemed counterproductive to introduce this topic as a descriptive statistic only to return to it later under the category of significance testing. Also, the topic of bivariate statistics can interrupt the smooth flow from explanations of the z score and the normal distribution to the use of those concepts for one-sample hypothesis testing. Similarly, the topic of probability as applied to discrete events can also interrupt the flow from descriptions of the normal distribution to explanations of hypothesis tests based on that distribution. Therefore, in this text, discrete probability is not discussed until Chapter 19, when it is needed for an explanation of nonparametric statistics. On the other hand, the rules of probability as applied to smooth distributions are relevant to parametric hypothesis testing, so these rules are presented at the end of the part of the text that deals with descriptive statistics.

The most unusual change in the normal order of topics is the placement of the information on the matched (or paired) t-test after the chapters on correlation and regression, rather than immediately after either the chapter on the one-sample t-test or the chapter on the t-test for two independent samples. The rationale for this placement is that the power of the matched t-test, and the degree to which it is better than the independent-samples t-test, can best be explained in terms of linear correlation. (And the placement of the material on correlation was constrained by the considerations mentioned above.) For my purposes, the conceptual link between the matched t-test and correlation is more important than the computational link between the matched t-test and the one-sample t-test. However, as students generally enter the course with a good intuitive understanding of correlation, it would not create any problems to cover the material on the

matched *t*-test immediately after the material on the one-sample or the two-sample *t*-test.

ABCD Format

You will notice that each chapter is divided into four sections, labeled A, B, C, and D. Section A focuses on definitional formulas—their structure and their relationship to other definitional formulas—so that students can gain some insight into why and how statistical formulas work. Students seem to have better retention of statistical formulas that do not appear arbitrary. Section B presents computational formulas and step-by-step procedures, adding the details students need to analyze data and interpret results. Section B also includes information on how to report statistical results in the latest APA format, and most of the B sections contain an excerpt from a journal article illustrating the use and reporting of the test discussed in that chapter. Section C presents optional material that is usually more advanced and that can be omitted without any loss in continuity. Section D provides a comprehensive summary of the previous three sections, along with definitions of key terms and a list of all the major formulas in that chapter.

The Content of the Text

The text is organized into six parts. Part I contains only one chapter, which introduces all the fundamental concepts of the text. Section A describes the different scales of measurement, which will be referred to chiefly in Chapters 3 and 4 and again in Part VI. Chapter 1 also introduces some concepts of experimental design. (This text contains a good deal of information on experimental design woven into the discussions of statistical tests.) Section B introduces the rules of summation, which will be used many times in the text to illustrate the derivation of one formula from another. Section C introduces the notion of double and triple summations. These are used only in Chapter 14, Section C, but are presented in order to prepare the student for more advanced statistical texts.

Part II covers descriptive statistics. Chapter 2, on frequency distributions, can easily be omitted in its entirety for more advanced courses with no loss in continuity. Section A describes the construction of frequency distributions and the finding of percentiles for ungrouped distributions, and Section B goes over the same material for the more complex case of grouped distributions. Because beginning students often find the concept of linear interpolation confusing and intimidating, this topic has been placed in Section C, which also includes a discussion of Tukey's stem-and-leaf display—not because it is more advanced, but because it is newer and less often encountered than the frequency distribution. In Chapter 3, Sections A and B contain the traditional material on the measurement of central tendency. Section C discusses Tukey's box-and-whisker plot as a means of identifying outliers and suggests what can be done to lessen the impact of outliers. Sections A and B of Chapter 4 contain material on the measurement of variability. Section C includes a description of the standard measures of skewness and kurtosis—not because these measures are frequently used in

psychological research, but because they are frequently included in statistical software packages and because they serve to deepen the students' understanding of how distributions can differ. Section A of Chapter 5 combines a description of standardized scores with a description of the normal distribution, and Section B demonstrates how these topics can be used together to solve practical problems. Finally, Section C presents the basic rules of probability (including conditional probability) as applied to smooth distributions, in preparation for the discussions of hypothesis tests in Part III.

Part III covers hypothesis tests involving one or two sample means. An entire chapter, Chapter 6, is devoted to the critical concepts of the sampling distribution and the Central Limit Theorem. The logic of null hypothesis testing is introduced in Section A of Chapter 7, and a step-by-step procedure to carry out hypothesis tests is introduced in Section B. Section C attempts to give the student a deeper understanding of p values as conditional probabilities and the fallacies that can arise in dealing with the topic of null hypothesis testing. In Chapter 8, Section A describes the one-sample t-test, and Section B focuses on interval estimation and confidence intervals for a single population mean. Section C contains some more advanced information concerning sampling distributions and the properties of estimators, but this material is meant only to prepare students for more advanced statistical treatments and is not needed in later chapters. Chapter 9, in Sections A and B, explains both the separate-variances and the pooled-variances t-tests, and Section C contains a discussion of when to use which test. Finally, Chapter 10 is devoted to the central concept of power. Section A explains power conceptually, and Section B delves into the calculations of power analysis. Note that the value of the power calculations goes beyond their actual use in experimental design; these calculations foster a deeper understanding of null hypothesis testing (e.g., knowing when a failure to reject the null hypothesis should be taken seriously). Section C deals with the serious concerns that have been raised about the overreliance on null hypothesis testing, and refers to classic and recent journal articles that have contributed to this important debate. Some defense of null hypothesis testing is offered.

The theme uniting the chapters of Part IV is that each chapter deals with a situation in which there are two dependent measures for each subject. Chapter 11 deals with linear correlation, including its limitations and its uses as a research tool. In addition to its discussion of power in relation to testing correlation coefficients for significance, Section C introduces Fisher's Z transformation and its applications for finding confidence intervals and comparing correlations from two independent samples. Sections A and B of Chapter 12 present the basic concepts and mechanics of linear regression, including confidence intervals for predictions. Section C covers two major topics: the point-biserial r and multiple regression. Point-biserial r is described as an important complement to the t value from a two-sample test and a useful alternative measure of effect size. Multiple regression is covered in some detail for the case of two predictors, and then the complications that arise with three or more predictors are briefly outlined. Instructors who choose to cover an introduction to multiple regression may want to postpone this topic until the end of the course because of its complexity. Finally,

Chapter 13 describes the matched, or repeated-measures, *t*-test and includes a good deal of material on experimental design. The calculation of the matched *t*-test is shown in terms of both Pearson's correlation coefficient and the direct-difference method. Section C demonstrates the connection between the power of the matched *t*-test and the degree of linear correlation between the two sets of scores.

Part V is devoted to the analysis of variance—both one-way and two-way designs. Section A of Chapter 14 develops the one-way ANOVA as a generalization of the *t*-test, and Section B presents computational formulas and other details. Section C deals with the concepts of power and effect size as applied to ANOVA and the use of the *F* ratio for testing homogeneity of variance. Chapter 15 discusses various methods used to control Type I errors when following an ANOVA with multiple comparisons but focuses on the computational details of two tests in particular: Fisher's protected *t*-tests and Tukey's "honestly significant difference." Section C tackles the more difficult concepts behind complex comparisons, planned comparisons, and orthogonal contrasts. Section A of Chapter 16 develops the two-way ANOVA by adding a grouping factor to a one-way ANOVA. Section B concentrates on computational formulas and on the varieties of experimental designs that call for a two-way ANOVA. Two important topics are presented in Section C: follow-up tests with and without a significant interaction, and the unweighted means solution for an unbalanced design. Section A of Chapter 17 shows how the one-way repeated measures ANOVA can be analyzed as though it were a two-way ANOVA with one observation per cell and draws connections between the one-way RM ANOVA and both the one-way independent-samples ANOVA and the matched *t*-test. As usual, Section B deals with computational formulas, experimental design, and publishing practices. Section C covers several important topics, including the Geisser-Greenhouse correction for a lack of sphericity, post hoc comparisons, and techniques for counterbalancing. Finally, Chapter 18 deals with the commonly encountered two-way ANOVA with one between-group factor and one within-subjects factor. Section C of this chapter, in addition to discussing post hoc comparisons and violations of assumptions, introduces two important advanced topics: analysis of covariance and multivariate analysis of variance. Both are discussed in terms of their simplest cases, and the section gives an indication of the complexity these designs can attain.

Part VI introduces the most often used of the nonparametric statistical procedures. Section A of Chapter 19 describes the binomial distribution, and Section B applies this distribution to a nonparametric alternative to the matched *t*-test: the sign test. Section C reviews the probability rules first presented in Chapter 5 and applies them to discrete events. Permutations and combinations are also introduced as ways of counting to determine the probabilities of discrete events. Chapter 20 describes the chi-square test, with Section A covering the one-way (goodness-of-fit) test and Section B detailing the two-way (independence) test. Section C deals with several topics, the most important of which is the measurement of the strength of association between the variables in a chi-square test. Finally, Chapter 21 presents two of the most common statistical tests for ordinal

data: the Wilcoxon (Mann-Whitney) rank-sum test and the Spearman correlation coefficient. Section C describes two additional tests for multigroup designs: the Kruskal-Wallis test and the Friedman test.

The Content of Introductory Versus Intermediate Courses

For an introductory, one-semester course, I would suggest skipping Section C in all chapters, with the possible exception of chapters 2, 5, 6, 7, and 10. Chapter 10 can be skipped entirely (though I recommend including at least Section A from that chapter), as can Chapters 18 and 19 and possibly Chapters 17 and 21, depending on the instructor's priorities. For an intermediate (i.e., master's level) course, I would suggest including Section C for all chapters but eliminating Chapter 2 entirely. In fact, Chapters 1 through 6 could be reviewed briefly in one lecture.

Appendixes

Appendix A contains all the statistical tables that are needed to cover the material in this text. Appendix B consists of a two-part review of the basic arithmetic and algebraic operations that are required to work the exercises in each chapter. You may want to urge your students to take the diagnostic quiz at the beginning of Appendix B during the first week of your course, so students will know as soon as possible if they need to review basic math. Appendix C presents a statistical decision tree that helps to answer the frequently asked question, Which statistical test should I use? If used near the end of the course, this appendix can aid students in summarizing and organizing what they have learned. Appendix C also contains exercises that students can use to test their own ability to choose the appropriate test for a particular experimental design. Finally, Appendix D contains answers to selected exercises (those marked with asterisks in the text) or, in some cases, selected parts of selected exercises. Note that answers are always given for exercises from earlier chapters that are referred to in later chapters. If an exercise you would like to assign refers to a previous exercise, it is preferable to assign the previous exercise as well, but it is not necessary; students can use Appendix D to obtain the answer to the previous exercise for comparison purposes.

Supplements

Exercises for which answers are not given in Appendix D are solved completely in the Instructor's Manual. Also included in the manual is a rationale for each exercise (whether it is answered in Appendix D or not), explaining the principles that are being illustrated by that exercise and detailing how that exercise relates to the other exercises and the concepts being taught in the course. Finally, the Instructor's Manual contains teaching tips for new instructors.

If you plan to use computer data analysis as an adjunct to your course, students can purchase a Computer Guide separately. The Computer Guide contains sample programs and detailed explanations of how to use SPSS and SAS for data

analysis. The Computer Guide covers both mainframe and PC/DOS versions of both SPSS and SAS, as well as SPSS for Windows. The chapters of the guide match the chapters of the text in terms of the statistical procedures described. However, the guide contains its own exercises based on real data sets. In the context of these exercises, students learn to use the data-handling features of SPSS and SAS as well as the basic statistical procedures.

Barry Cohen

PREFACE TO THE STUDENT

About the Size of This Text

One of the first things you will have noticed about this text is how thick it is. This book is larger than most introductory or intermediate statistics texts, and some students may find the sheer number of pages intimidating. The book does contain a great deal of information about statistics, but there are several reasons why the number of pages can give you a misleading impression. First, note that concrete analogies and detailed verbal explanations take up more room than mathematical language and statistical jargon. This text is not designed to be mathematically elegant; it is designed to communicate with students whose mathematical background and experience with psychological research are minimal. (If your mathematics background is weak, be sure to take the quiz in Appendix B.) Second, a good deal of redundancy is built into the structure of the text. You may have noticed in skimming through the pages that each chapter is divided into four sections. Material in one section may be covered from a different angle in another section of the same chapter. Third, the text is comprehensive enough to be used either as an undergraduate text for a first course in statistics or as an intermediate text for a course on the master's level. Your instructor may not assign all of the chapters or all of the sections of all of the chapters. But the early chapters may be useful for master's students (as a refresher course), and the later chapters may be useful to undergraduates (for future reference).

The Goals of This Text

Most psychology students take a course in statistics because it is required—and it is required because statistics plays an important role in most areas of psychological research. Any research you conduct after your statistics course will strengthen and deepen your understanding of statistical concepts. And even if you do not have the opportunity or inclination to conduct research after graduation, you will want to read the results of research published in psychological journals. Therefore, this text not only teaches you how to use statistics to summarize and draw conclusions from the psychological data you collect in the future, it also emphasizes how to read and interpret the statistical results published in psychological journals. This emphasis should prove especially useful to the many students who will eventually enter careers in which they will need to be critical consumers of research results presented in a statistical form. On the other hand, those who enter careers in which they will be collecting psychological data will undoubtedly perform most of their statistical analyses using sophisticated computer software. These students will appreciate this text's emphasis on understanding the meaning of statistical results rather than on the details of calculating those results.

Although I wrote this text for the student with little mathematical background and perhaps some anxiety about math, I have not given up on the goal of teaching

the basic mathematical structure of statistical formulas. Rather, I have undertaken to explain some fairly complex mathematical relationships verbally. In my experience, the material in a statistics course is more memorable if the rationale for each formula and the connections between apparently dissimilar formulas are made clear. To attain the goals described above, I have used a unique ABCD structure in each chapter.

The ABCD Format

Section A provides the *conceptual foundation* for each chapter. I believe that in order to teach the concepts completely, it is necessary to explain the structure of the basic statistical formulas. However, the emphasis in this section is not on computational shortcuts or the details of performing a statistical test. Rather, the focus is on definitional formulas: what the different pieces of a formula represent and how one formula is related to those covered in previous chapters. The prose in Section A can be a bit "long-winded," because I am using simple language, trying to stay concrete, and expressing the same concept in several ways in order to get my point across. The exercises at the end of Section A help to solidify the truly fundamental concepts in your mind before you encounter the additional details presented in Section B.

Section B presents the *basic statistical procedures* for each chapter. This section covers essentially the same material as Section A but includes important details and qualifications, and the emphasis is on step-by-step procedures for performing statistical tests. Alternative formulas are introduced because they are either easy to use with a calculator or frequently encountered in other texts or manuals. The exercises at the end of Section B ensure that you can perform the statistical tests covered in the chapter and also remind you of the important concepts related to those tests.

Section C contains *optional material* that should be tackled only after you have successfully completed the exercises of Section B. Often the material in Section C is more advanced conceptually and potentially confusing. In some cases, concepts or procedures are in Section C because they are not often used or because there is a good chance that your instructor will ignore them in order to devote time to more central issues. Your instructor may want to cover Section C for some chapters but not others. However, the material in Section C is always useful, and you may decide to try some of the C sections not assigned by your instructor. The exercises at the end of this section will help you test your mastery of this additional material.

Finally, Section D contains a *summary* of each of the three preceding sections. This text does not come with a separate study guide; the D section of each chapter is designed to function as a built-in study guide. The summary of Section A is designed to be quite thorough, and if you absorbed the concepts of that section on the first reading you may not have to read Section A over again for study purposes—the summary should be sufficient. It is recommended that you read the summary of Section A immediately after reading Section A for the first time and before proceeding with Section B. The summary of Section B streamlines

the procedures of that section and presents an additional example to broaden your understanding. The summary will often be sufficient as a guide to solving the exercises for Section B. The summary of Section C can serve not only as a review of that material but also as a basis for deciding whether you want to tackle that section on your own. If you decide not to read an unassigned Section C, the summary can give you some slight familiarity with material you may need to learn in the future. Section D concludes with definitions of the most important terms in the chapter and a list of the important formulas. Any term that is defined at the end of a chapter appears in **boldface** in the text and in the index, and the page on which it is defined appears in boldface in the index, as well.

Additional Features

The chapters of this text are grouped into six parts, each based on a theme. Each part begins with its own introduction that will orient you to the theme and give you some idea of the content of each section in each chapter. In addition, each chapter begins by listing the symbols, concepts, and formulas from previous chapters that you will need to use in that chapter. You will notice the cumulative nature of the material in this text; most of what you learn in the early chapters will be used again in subsequent chapters.

Appendixes

Appendix A contains all of the statistical tables you will need to solve the exercises in each chapter. Appendix B contains a review of the arithmetic and algebraic tools you will need to solve the exercises. If you think that your math background is weak, or you don't remember much of it, take the diagnostic quiz at the beginning of Appendix B as soon as possible. If you need to refresh your math, it is best to do this before you become immersed in your statistics course. The math review and exercises contained in Appendix B will be all you need to handle the mathematics of the statistical tests in this text. Appendix C is a statistical decision tree that can help you decide which statistical summary or test to use for a particular situation. This appendix will be most useful near the end of your statistics course, to summarize and organize what you have learned and help you to apply your knowledge to the analysis of actual research designs. Finally, Appendix D contains the answers to selected exercises—those marked by an asterisk in the text. (When exercises that are starred have multiple parts, Appendix D may include answers for only some of the parts.) Many of the exercises in later chapters of the text refer to exercises in earlier chapters and ask you to compare answers or use earlier results as a stepping stone to solving the later exercise. Later exercises refer only to earlier exercises whose answers are in Appendix D. Therefore, when an exercise refers to a previous one that was not assigned, or that you could not solve, you can look up its answer in Appendix D and use that answer to help you solve the later exercise.

As a preview of the content of this text, you may want to read the Preface to the Instructor.

Acknowledgments

This text began in the summer of 1988 when Dr. R. Brooke Lea, then my teaching assistant, both encouraged and assisted me in transferring my lectures to written form for class distribution. Dr. Lea deserves much credit as well as gratitude for the existence of this book.

Over the past seven years the text has evolved with the help of many interested students and teaching assistants, including Ann Demarais, Andrew Hilford, Naomi Kubo, Jeannette Levine, Kathryn Perko, Katya Rascovsky, Kay Rothman, Rosalind Sackoff, Jeremy Stone, Inna Subotin, Jennifer Thomas, Kay Wenzel, and Cheryl Williams. Thanks to all those I mentioned, and both thanks and apologies to those I may have accidentally left out.

My lecture notes would never have become a published text without the support and encouragement of the editors at Brooks/Cole. Philip Curson was the first to make me think that publishing my text was a possibility, and Marianne Taflinger kept me going through the early stages of review and text development. James Brace-Thompson continued the regimen of patience and encouragement, and to all of them I owe my heartfelt thanks. Another debt of gratitude is owed the entire production staff at Brooks/Cole and Lifland et al., Bookmakers, especially Kirk Bomont, who oversaw the production, Quica Ostrander, who contributed doubly by copyediting the manuscript and designing the text, Denise Throckmorton, who prepared the index, and Roy Neuhaus, who designed the striking cover. Several instructors of statistics were enormously helpful in reviewing the text at various stages, offering useful suggestions and corrections: Linda Allred, East Carolina University; Karen Caplovitz Barrett, Colorado State University; Michael Biderman, University of Tennessee at Chattanooga; Dennis Bonge, University of Arkansas; Dennis Cogan, Texas Tech University; Lawrence T. DeCarlo, Fordham University; George Domino, University of Arizona; Rick Froman, John Brown University; Tom Gilovich, Cornell University; Debra K. Harless, Bethel College; Kelli Klebe, University of Colorado–Colorado Springs; John B. Murray, St. John's University; and Peter Taylor, Cornell University.

I am grateful to the Longman Group UK Ltd., on behalf of the Literary Executor of the late Sir Ronald A. Fisher, F.R.S., and Dr. Frank Yates, F.R.S., for permission to reproduce Table IV from *Statistical Tables for Biological, Agricultural and Medical Research,* 6th Ed. (1974).

Finally, the Department of Psychology at New York University was very supportive while I was writing this book. I especially want to thank Bettie Brewer, the Graduate Secretary, for helping me in countless ways. All of my family and friends deserve much thanks, as well, for their patience and understanding—most of all my wife, Leona, who deserves thanks also for tirelessly typing the entire manuscript.

Barry Cohen

CONTENTS