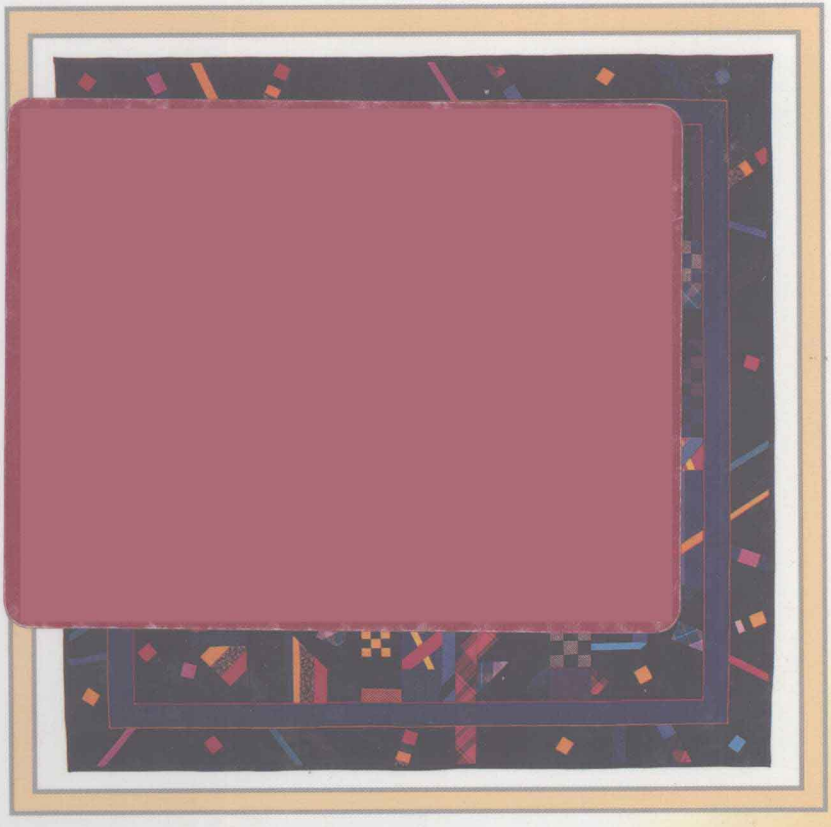
SEVENTH EDITION

BASIC STATISTICAL ANALYSIS



RICHARD C. SPRINTHALL

SEVENTH EDITION

Basic Statistical Analysis

RICHARD C. SPRINTHALL

American International College

Boston New York San Francisco
Mexico City Montreal Toronto London Madrid Munich Paris
Hong Kong Singapore Tokyo Cape Town Sydney

Series Editor: *Kelly May*
Editorial Assistant: *Marlana Voerster*
Marketing Manager: *Brad Parkins*
Editorial Production Service: *Chestnut Hill Enterprises, Inc.*
Manufacturing Buyer: *JoAnne Sweeney*
Cover Administrator: *Linda Knowles*
Electronic Composition: *Omegatype Typography, Inc.*

For related titles and support materials, visit our online catalogue at www.ablongman.com.

Between the time Website information is gathered and published, some sites may have closed. Also, the transcription of URLs can result in typographical errors. The publisher would appreciate notification where these occur so that they may be corrected in subsequent editions.

Library of Congress Cataloging-in-Publication Data

Sprinthall, Richard C.
Basic statistical analysis / Richard C. Sprinthall. — 7th ed.
p. cm.
Includes bibliographical references and index.
ISBN 0-205-36066-1 (alk. paper)
1. Social sciences—Statistical methods. 2. Statistics. I. Title.
HA29 .S658 2003
519.5'024'3—dc21 2002066543

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1 07 06 05 04 03 02

Preface

As with any new edition, significant changes and additions have been incorporated into this seventh edition of *Basic Statistical Analysis,* and yet the major thrust of the coverage remains essentially the same. To that extent it may seem like deja vu, or as some naive critics of statistical analysis might say, deja voo doo. In fact, this new version may even seem, as Yogi Berra once said, like experiencing "deja vu all over again." (That's redundant, sir, isn't it, isn't it?) And if a double redundancy isn't enough, let's remember Howard Cosell, who once offered a triple-play on words when he asked his sports viewers "to reflect back nostalgically on the past." And so, as was the case with past editions, this book is still intended for students in the behavioral and health sciences who are confronting their first "stats" course. Although the book was originally written with psychology majors in mind, this edition features an infusion of new material that should be appropriate for other majors, especially criminal justice majors, education majors, and majors in the health sciences. Despite the overall continuity with previous editions, the text has changed in a number of important ways.

New features in this edition are many, but the major addition is the inclusion of the *SPSS Statistical Program.* This is probably the most popular and creative statistical program ever produced. The inclusion of this tool should add immeasurably to the increasing statistical sophistication of today's students. This edition of the text has been rewritten to fit seamlessly with SPSS and an appendix has been included that should provide students new to computing with a complete user's guide to this powerful program. For any of you who prefer to use the *AB-Stat* program that has accompanied this book in the past, it can be downloaded from the Allyn-Bacon Web site (www.ablongman.com). Other new features include:

1. More examples of statistical misdirection in Chapter 1
2. New techniques for scoring questionnaire research in Chapter 3
3. A chance to have some fun with your calculator in Chapter 4
4. A creative device for calculating various unknowns for the *z*-score equation in Chapter 5
5. New material for finding the *SD* or *s* from the *SEm*
6. A tightening of the research chapter, Chapter 9, with additional material in this chapter on the placebo effect
7. A simplified version of the Bonferroni test in Chapter 12

8. More on coefficient alpha and its relationship to the Cyril Hoyt in Chapter 17; also, new material on construct validity and test ethics
9. More coverage of confidence intervals for the single sample *t*, the two-sample independent *t* and the paired *t*
10. Many new and interesting examples from the fields of psychology, law enforcement, health sciences, education, and sociology. Also, all problems now have a research basis and include a full write-up of the conclusions.
11. And finally, we present two new members of the statistical hall of fame: Ada Lovelace, the daughter of the poet Lord Byron, who has been admitted because of her pioneering efforts in computer programming and Gertrude Cox, the founder of the first experimental statistics department in the United States.

For many students, approaching that first statistics course is a rather scary and mystifying experience. The prospect is laced with a large dose of ambivalence, like a two-edged sword that has "can't live with it" on one side and "can't live without it" on the other. The "can't live with it" feeling often owes its genesis to the memory of past traumas, perhaps a ninth-grade algebra course in which everything went smoothly "until we got to the word problems." And the "can't live without it" conviction may be even more direct and compelling, since the statistics course is probably required, or, at the very least, "strongly recommended."

To help alleviate some of the pangs of anxiety, many students adopt one of two main defenses, both of which revolve around the same core notion that this is a course that is simply not needed. First, there are those who are convinced they do not need it because the intricacies of statistical proof are irrelevant. To them reading the conclusions of a research report is enough: "If it's got numbers in it, especially significant numbers, then I believe it." Understanding whether the logic of the design warrants those significant conclusions is deemed a waste of time. Students often feel that this is like asking them to understand the theory of the internal combustion engine when all they want to do is drive the car. The attitude of these students resembles that of the person who earnestly intones that "they say to starve a cold," without having the vaguest notion as to who "they" are ("they" are always the anonymous voices of supreme and unchallenged authority). This group of students is blithely willing to accept all statistical results. Second, there are those students who are convinced that they do not need the course because statistics is all a pack of lies. To them, statistical proof is tantamount to no proof at all and so they smugly and cynically reject all statistical evidence. Their attitudes are like those of persons who say, "I don't think it works this way," and when questioned as to how it does work, they reply, "some other way."

There is of course a germ of truth in both arguments. Significant numbers do mean something, although not always what the researcher or research consumer concludes. Some statistical proofs are like a sleight-of-hand show: Now you see it, now you don't. In order to know what the significant results mean and when the statistical charlatan is operating, at least one course in statistics is essential. Newspapers, magazines, and television newscasts constantly bombard us with the results of research studies: drink-

ing coffee causes this, drinking beer causes that, washing clothes in a certain detergent causes something else, and on and on. Determining cause-and-effect relationships should be made on the basis of fundamental and sound research practices, not by falling prey to a human need to quantify that is so strong as to embrace even dubious math as evidence of analytical rigor. Just to be an informed citizen and voter, some knowledge of statistical methdology is of enormous benefit. One must be made alert to politicized statistics, such as when a new administraiton announces in September that there are more people working than when they took over last January (ignoring the obvious seasonal variations in employment).

The avowed goal of this book is to demystify statistics—to state the case for statistical analysis and inference in clear, no-frills language. The student is told specifically what an X is and what a Y is, and whether the twain shall ever meet. As in law school, the student is presented with rules of evidence and the logic behind those rules. The focus will constantly remain on how statistical techniques can be used. It will not be a case of presenting the best method for calculating a standard deviation and then leaving it up to the student to find some use for it. Statistical concepts are embedded in the hard rock of research methodology. The student will learn at a practical level how to read and do statistical research. The student will be given a guided tour of the most important and practical exhibits in the statistician's showcase: not to create feelings of awe, but to teach for understanding.

Part of the power of being a professional has been the ability to use, take control of and protect from public scrutiny the language of the trade. The way to keep any discipline closed is to make the concepts obscure enough to only be understood by the so-called experts, thereby excluding the lay audience. Using and creating terms that have the dazzling sound of super sophistication become intimidating moats, designed to keep the great unwashed from entering the statistical fortress. This book, unlike so many others, crosses the moat.

For most students it will be easy to read and, at times, perhaps even fun to read. The book assumes little or no background in mathematics. The student will not be stunned by finding an elegant, but tricky, derivation on page 3 or by finding that the author suddenly assumes on page 5 that everyone remembers enough calculus to integrate the normal curve equation. The student does not even have to remember arithmetic, let alone calculus.

The use of this book does require, however, that the student own an electronic calculator. Although the calculator need not be expensive, it must at least have a square-root key. Pressing the square-root key is easier and more accurate than looking up and interpolating table values. Therefore, the back of this book is not cluttered with pages of square and square-root tables.

Although the text is designed primarily for a one-semester, beginning course, enough added material is presented to allow its use by students taking more advanced courses. Chapters 1 to 13 contain topics usually covered in a one-semester course and if this is what is needed, the course can end with Chapter 13. At this point, the student will have gained enough understanding of statistical reasoning and research methodology to be able to read and comprehend a large part of the research in the social-science literature. Because many students must later take courses in experimental

psychology or in research methods in education and/or sociology, topics sometimes found in the more advanced courses are also included here in Chapters 14 through 19.

The book is divided into three major units: Descriptive Statistics, Inferential Statistics, and Advanced Topics in Inferential Statistics. How can a book of this size cover so much? Because some topics will not be covered here. First, little consideration will be given to grouped data problems. Finding class intervals and standard deviations from the frequency data inherently creates some error and also loses track of the individual score. When the amount of data is so large that grouped data techniques are really needed, statisticians turn to computers anyway. Second, the coverage of probability theory will be shortened. Not that probability theory is unimportant; it is absolutely crucial. But the only probability concepts found in this book are those that bear directly on practical statistical tests of significance. What is practical? How to calculate and understand the logic behind such things as z scores, the t test, ANOVA, chi square, and the Pearson r, and regression.

Special features of this book include the following:

1. *Definitions of key concept in the glossary.* Brief, but thorough, definitions are conveniently presented in the glossary. Experience has clearly shown that much of the trauma experienced by students taking their first "stats" course can be traced to confusion over terminology. Conscientious use of the glossary can alleviate most if not all of this confusion.

2. *A programmed approach to the computation of each statistical test.* Computational procedures are set forth in a step-by-step programmed format. A student who can follow a recipe or build a simple model plane can do an ANOVA.

3. *Constant stressing of the interaction between statistical tests and research methodology.* Examples are used from the literature of the social sciences to illustrate strategic methodological problems. Statistical analysis, if not carried out in the context of methodology, can degenerate into an empty and sterile pursuit. Three chapters have been specially designed to bridge the analysis-methodology gap. Chapter 9 focuses directly on the essentials of the research enterprise, Chapter 17 relates the statistical analyses to measurement theory and Chapter 19 presents 26 research simulations that are programmed in such a way as to lead directly to the appropriate statistical analysis.

4. *A literary style that is both easy to read and attention getting.* This book attempts to generate a feeling of excitement and enthusiasm, by talking directly to the students and spotlighting the student's own life space. Students obviously learn best when their interests are aroused.

5. *A large number of problems and test questions.* John Dewey's "learn by doing" axiom was never more true than in the field of statistics. Over 400 problems and test questions are placed both within and at the end of each chapter. Students need the opportunity to "try their hands" at practice problems and to get some immediate feed-

back as to their progress.

6. *A list of key points and names.* Each chapter contains a list of key points and names that also appears in the glossary. At the end of each chapter a convenient wrap-up summary is also provided.

7. *Computer program.* A statistical program (SPSS) is included that covers the statistical tests presented in the book. The program is totally menu driven and can be easily handled within an hour by the first-time user.

8. *Computer printouts.* A series of computer printouts, all containing errors of one kind or another, is presented in Chapter 18. Students may then use the "logic checks" found throughout the text to identify these errors. This should accomplish at least two important objectives: 1) remind us that the computer is not an infallible genius, but is instead, a fast idiot who needs a smart leader, and 2) reinforce the various statistical logic checks by making them "live" in the context of a computer printout.

9. *The binomial distribution.* A special section is included that covers the essentials of the binomial distribution and its relationship to the z distribution. This section includes methods for obtaining exact probabilities for the binomial distribution, as well as the z approximations. The coverage here also includes the z test for proportions and the t test for differences between proportions (and its relationship to chi square). Problems are worked out within the chapter, and a series of student problems presented at the end.

10. *Supplements.* A new and up-dated "Instructor's Manual" containing well over 1800 test items and problems is available. The Instructor's Manual also includes a number of suggestions for class activities, all designed to spark student interest.

Putting together a book of this type requires a lot of help. Special thanks must go to the "significant others" in my academic life, the professors who first initiated and then sustained my interest in statistical analysis: Greg Kimble, then of Brown University; the late Nathan (Mac) Maccoby, then of Boston University; and the late P. J. Rulon of Harvard University. Without them this book could never have happened. Also, I am grateful to the Literary Executor of the late Sir Ronald A. Fisher, F.R.S., to Dr. Frank Yates, F.R.S., and to the Longman Group Ltd., London, for permission to reprint Tables III, IV, and VI from their book *Statistical Tables for Biological, Agricultural, and Medical Research* (6th edition, 1974).

More recently I must thank my colleagues at American International College: Tom Fitzgerald for his creative chi square examples, Lee Sirois for his significant role in putting together Chapter 8, Gregory Schmutte for his valuable contributions to the research chapters, Gus Pesce for his "spec ed" examples, Pam Diamond for her z score matrix in Chapter 4, and Marty Lyman at the Hampden County House of Corrections for her research examples. I also wish to extend a special thanks to Barry Wadsworth at Mt. Holyoke College, Marjorie Marcotte at Springfield College and Norm Sprinthall at North Carolina State for their insightful comments on topic inclusions (and

exclusions), to Barbara Anderson Lounsbury at Rhode Island College for her help and encouragement during the early stages of the manuscript's development, to both Jim Vivian at Gordon College, Norm Berkowitz at Boston College, and Chris Hakala at the Western New England College for their ideas on sampling theory, to Steve Fisk at Bowdoin College for his expertise in showing the world how to put together a computer program that is really menu driven, to Mike Plant for his superb revision of the computer program (and to Peter Flournoy at the V.A. Medical Center in Portland, Oregon, for his many suggestions for the program revision). I must also thank Ken Weaver at Emporia State University in Kansas for his careful reading of the entire manuscript. Without Ken's help, many errors of omission and commission would have been left in the text. Thanks also to the following reviewers for their helpful comments: Ken Hobby, Harding University; James Johnson, University of North Carolina, Wilmington; Neva Sanders, Canisius College; and Laura Snodgrass, Muhlberg College.

Also a special note of gratitude must go to Dianne Ratcliffe Sprinthall for her illustrations, as well as her general help in artistic design, and to both Lou Conlin and Sidney Harris for their marvelously clever cartoons.

Contents

UNIT

I

Descriptive Statistics

Chapter

1

Introduction to Statistics

During the next hour, while you are reading this chapter, 250 Americans will die, and the chances are one in a million that you'll be one of them. Don't stop reading, however, since that won't reduce the probability. In fact, putting your book down might even increase the probability, especially if you decide to leave your room and go outside to engage in some high-risk activity.

According to the TV special "Against All Odds," if you go rock climbing, the probability of your getting killed is 200 in a million; or parachuting, 250 in a million; or hang gliding, 1140 in a million. So, sit safely still and read on, and while doing that let's look at some more of life's probabilities: the probability of having your car stolen this year, 1 out of 120; of a pregnant woman having twins, 1 out of 90 (or triplets, 1 out of 8000); of a young adult (18–22) being paroled from prison and then being rearrested for a serious crime, 7 out of 10; and of any single American baby becoming a genius (IQ of 135 or higher), less than 1 out of 100 (Krantz, 1992). Incidentally, by the time you finish reading Chapter 6, you'll be able to calculate that genius probability value—even if you're not a genius yourself.

As you probably know, most accidents occur at home, since typical Americans spend most of their time there. And 25% of all home accidents occur in the bathroom—falling in the tub, getting cut while shaving, and so forth. Don't assume, however, that you'll be safe if you decide to shave in the kitchen instead. Also, we can predict with a high degree of accuracy that during the next year 9000 pedestrians will be killed by a moving car. But this statistic does not tell us which 9000. Understanding probability situations is an important aspect of life, so maybe there are some good reasons for getting involved in statistical thinking.

Rather than continuing to list reasons why you should take a first course in statistics, let's assume that it is probably a required course and that you have to take it anyway. Perhaps you have put it off for quite a while, until there is no choice left but to "bite the bullet" and get it over with. This is not to say that all of you have been dragged, kicking and screaming, into this course; however, as statisticians would put it, the probability is high that this hypothesis is true for some of you.

STUMBLING BLOCKS TO STATISTICS

Let us look at some of the most common objections raised by students when confronted with this seemingly grim situation. Perhaps your feelings of intimidation arise because you know you have a math block. You're still being buffeted by lingering anxieties from some math course taken in the perhaps distant past. Or maybe it's that you have read or heard a research report and been totally confused by the seemingly endless and seemingly meaningless stream of jargon. Perhaps you're a person who simply does not trust statistical analysis. If this is the case, you're in good company. Benjamin Disraeli, Queen Victoria's prime minister, once said, "There are three kinds of liars: liars, damned liars, and statisticians." Disraeli obviously agreed with the sentiment expressed by many—that you can prove anything with statistics.

Before he died, Malcolm Forbes had been a hot-air balloon enthusiast, and one day the winds just took his balloon in so many directions that he became completely lost. Spotting what appeared to be a farmer down below tilling his field, Forbes lowered the balloon and called out to the man, "Please tell me where I am." The man called back, "You're up in a balloon, you goddamned fool." And Forbes answered, "You must be a statistician, since although your answer is complete, accurate, concise, and precise, it tells me absolutely nothing that I don't already know." And then there's the story of the three statisticians who went hunting and after a while spotted a solitary rabbit. The first statistician takes aim and overshoots. The second aims and undershoots. The third shouts out, "We must have got him."

Whatever their basis, your doubts about taking this course will probably prove unfounded. You may even, believe it or not, get to like it and voluntarily sign up for a more advanced course.

Math Block

First, although it is obvious that people do differ in math ability, a case of true math block is extremely rare and difficult to substantiate. It is true that some very fortunate people have a kind of perfect pitch for math. They take to math as gifted musicians take to harmony. (You remember the kid we all hated in high school, the one who completed calculus during his sophomore year and was angry because the school didn't offer any more advanced math courses.) At the other end of the continuum, we find those people who are definitely math phobics. To them, merely drawing a number on the chalkboard evokes strangulating feelings of sheer panic. They avoid any course or situation (even keeping a checkbook) that deals with those spine-chilling little inscriptions called numbers. If you're one of those who suffers from or borders on this condition, relax—this is not a math course. While numbers are involved and certain arithmetic procedures are required, thanks to the magic of electronics you won't have to do the arithmetic yourself.

Go to your friendly neighborhood discount store, and, for less than ten dollars (less than the price of a good slide rule), purchase a small electronic calculator. You