FUNDAMENTAL STATISTICS FOR BEHAVIORAL SCIENCES

Robert B. McCall

SIXTH EDITION

FUNDAMENTAL STATISTICS FOR BEHAVIORAL SCIENCES

SIXTH EDITION

Robert B. McCall
University of Pittsburgh

Under the General Editorship of

Jerome Kagan
Harvard University

Harcourt Brace College Publishers

Fort Worth Philadelphia San Diego New York Orlando Austin San Antonio
Toronto Montreal London Sydney Tokyo

Publisher **Ted Buchholz**
Acquisitions Editor **Christina N. Oldham**
Developmental Editor **Carol Jaramillo Hargrave**
Senior Project Editor **Steve Welch**
Senior Production Manager **Tad Gaither**
Art Director **Sue Hart**

ISBN: 0-15-500572-3
Library of Congress Catalog Card Number: 93-77653

Address for Editorial Correspondence: Harcourt Brace College Publishers, 301 Commerce Street, Suite 3700, Fort Worth, TX 76102.

Address for Orders: Harcourt Brace & Company, 6277 Sea Harbor Drive, Orlando, FL 32887-6777. 1-800-782-4479, or 1-800-433-0001 (in Florida).

Printed in the United States of America

3 4 5 6 7 8 9 0 1 2 039 9 8 7 6 5 4 3 2 1

FUNDAMENTAL STATISTICS FOR BEHAVIORAL SCIENCES

PREFACE

A veteran scholar of statistics once remarked that there are two ways to teach statistics—accurately or understandably. Disproving that "either-or" statement is the challenge that has guided my writing of this book.

As a textbook for the first course in applied statistics, *Fundamental Statistics for Behavioral Sciences* is used primarily by students majoring in psychology, education, and other behavioral sciences. In writing for this audience, my earliest and most basic decision was to emphasize the purpose, rationale, and application of important statistical concepts over rote memorization and the mechanical application of formulas. I believe that students at the introductory level, whether or not they plan to take advanced courses in statistics, are served better by a book that fosters an understanding of statistical logic than by one that stresses mechanics.

When understanding is emphasized, elementary statistics is neither dull nor mathematically difficult. *Fundamental Statistics for Behavioral Sciences* does not require much background in mathematics. The student need be familiar only with the thinking patterns learned in high school algebra and geometry; all relevant terms and operations are reviewed in Appendix 1. To be sure, the book contains many computations and problems to solve, but most statistical formulas rely heavily on simple arithmetic—addition, subtraction, multiplication, division, and the taking of square roots—and can be worked out quickly with the aid of a hand calculator. In addition, I have kept the data for computational problems simple so that the emphasis remains on the rationale and outcome of techniques instead of on calculation for its own sake.

The goal of understanding concepts is not hard to reach if students also understand as much of the mathematical reasoning as is within their grasp. Whenever possible, I have explained in mathematically simple terms the logic that undergirds the basic concepts and techniques, although a few items require advanced mathematics and must therefore be taken on faith at this level. Beginning in Chapter 3, optional tables that show full algebraic derivations and proofs supplement the text explanations. These tables can be omitted without loss of continuity.

For the sake of students, and contrary to the traditional practice of mathematical writers, I have included and explained every step in proofs, however "obvious." I have also avoided excessive use of symbols, since symbols require an extra mental translation and thus often confuse students. Derivation scores ($X_i - \bar{X}$), for instance, are *not* abbreviated by x. Further, each new symbol is carefully introduced and is frequently accompanied by its verbal equivalent.

Anyone who has analyzed his or her own data knows the anticipation that accompanies the final calculation of the r, t, or F hidden in a mass of numbers that took months to gather. Students, too—even though it is not their own data they are analyzing—can experience the excitement of seeing meaning emerge as they manipulate an apparently patternless collection of numbers. Yet they sometimes fail to see the fascination of statistical analysis because it is presented more as a numbers game played in a vacuum than as a crucial part of the scientific investigation of real phenomena.

For this reason, many end-of-chapter exercises and examples in the text are drawn from actual studies (modified for numerical simplicity). For example, the distinction between a correlation and a difference between means is demonstrated through findings that related the IQs of adopted children to those of their biological and adoptive mothers, and Freedman's work on the feeding behavior of dogs reared under indulged and disciplined conditions is used to illustrate interaction effects in the two-factor analysis of variance. Although most of these studies were performed by psychologists, many of them concern developmental and educational issues of interest to future teachers and school administrators.

As another means of showing that numbers can have real applications, I have tried to give students a feel for the behavior of a statistic by providing several data sets that display obvious contrasts. In Chapter 10, for example, I present a set of random numbers and calculate an analysis of variance on them, then I add a main effect to the data and recalculate, and finally I add an interaction effect and recalculate. This shows students how various effects in the data are reflected in the statistical quantities being studied. In addition, I ask in some exercises for the student to alter a set of scores in some way and observe the effect on the value of a statistic.

Despite these attempts to help students, some find this course more difficult than many others. Apart from differences in quantitative abilities, I suspect that much of the reason is that they try to study it in the same manner as a "reading course." A course in history or introduction to psychology, for example, typically uses a textbook, and, except for new vocabulary, everything one reads is quite understandable the first time through. Also, the material in tables and figures is largely illustrative and not crucial to understanding the main text. The task is largely one of memorizing vocabulary and concepts and, in more advanced courses, integrating material across chapters.

Statistics also has its vocabulary, some of which is in symbols, but once introduced, the symbol is often used repeatedly thereafter. While I have tried to repeat the definition or verbal label with the symbol an extra time or two to help this learning process, eventually the symbol will be used by itself. Therefore, if the student does not commit the symbol's meaning to memory when it is first encountered, everything subsequently that uses that symbol—and it is sometimes a good deal of material—will not be truly understood. So, students should pause to learn each vocabulary term and symbol whenever it is first introduced.

Also, statistics often requires the student to follow a sequence of logical steps or elementary mathematical operations. Some students should pause in the middle of the sequence to review after the first reading to be sure the entire sequence is understood. Also, the basic material of statistics is mathematical, symbolic, and conceptual, and it is quite abstract without the concrete examples that are given. But students need to read and especially **do** the examples, step by step, along with the text. Skipping them can be disastrous for some students, and simply reading but not studying them is insufficient for other students. They must actually go through the calculations or the steps in the proof to understand it. And some examples, proofs, and illustrations are presented in separate tables and figures. Students should stop reading the text when such material is cited, turn to that table or figure, and study it until it is understood. If it is a calculation, they should do it on a piece of scratch paper along with the text. These are not pictures in a newspaper, which typically can be skipped; they are integral to understanding what is being presented.

Another tip is to compute by hand at least one—but two or three would be better for some students—computational example or exercise of a given type before rushing to a fancy one-button calculator or computer program. Students often need to get a "computational feel" for a formula before they understand and can interpret it thoroughly. And, of course, it helps to do as many exercises as possible in the text and *Study Guide*. Some students assume that once they understand the presentation in the text, they will be able to work a problem without practice. Of course, a few can, but a very few, in my experience. Most of us need to work one or two or even more problems in the text and *Study Guide* to become proficient at them, even if we understand the text material quite well.

Lastly, the answers and even some of the intermediate steps for all the exercises are given in the backs of the text and *Study Guide*. This is so that students can check their work immediately and study more independently. But if students "peek" before finishing, they often deprive themselves of the opportunity to figure the problem out on their own. The exams will not have the answers on them, so it is best to practice without them, too.

This Sixth Edition incorporates a number of changes from previous editions, often suggested by users, reviewers, and students. Most important but least noticed is that many explanatory sections have been rewritten to be clearer. More obvious, exercises have been separated into those that help with a conceptual understanding of the material and those that help students learn to solve problems, and I have designated with an asterick those exercises that are a bit more complicated or difficult than the others. I have also added more conceptual or thinking exercises.

While I have always tried to explain concepts from both an algebraic and geometric perspective, with the belief that some students learn better from one than the other presentation, I have tried to make sure that more concepts are given this dual treatment. And I have added a few more "realistic" illustrations and exercises.

I have not added many new concepts, however, resisting the common urge to become encyclopedic. I prefer students to understand the basics more thoroughly at the introductory level than to have a more superficial acquaintance with many techniques. However, I added a few paragraphs about when to use a histogram versus a polygon and some material on using the standard error of estimate to make interval estimations.

Finally, the most visible change is in the *Study Guide* where Richard Sass and I have added boxed instructions and input-output samples for using MINITAB and SPSS to perform many of the calculations presented. This material is separated from the main presentation so that it can be skipped or another statistical program can be used by the instructor. Obviously, each instructor must provide instructions on how to access the programs on their systems. I urge instructors to consider carefully whether they want to teach computer usage at this stage. Many do, feeling it is almost as important in modern times for students to learn how to use the computerized statistical packages as it is to learn statistics. Other teachers feel that it is best for their students to understand the basic concepts, that most students in their course will not become researchers and will not need to actually analyze data, and that it takes more time for students to learn the computer packages than they save in computation time. At least now one has a more convenient choice in this matter.

The *Study Guide* in general is intended to be a workbook. All the basics are presented in stripped-down fashion and without much conceptual explanation. But the *Guide* is intended to be a supplement to the text and it assumes that the text has been read before the student begins the review and practice that the *Guide* offers. Some students find that reading the text, then doing the *Guide*, and then rereading the text is an effective strategy, because the *Guide* reviews the text, and then the concepts in the text are more understandable after the vocabulary and the concrete computational mechanics are mastered.

More specifically, the main section of each chapter in the *Guide* is a semiprogrammed unit that reviews the basic terms, concepts, and computational routines described in the corresponding chapter of the text. Often students are required to actually perform the computations in a step-by-step manner in an example, making it more difficult for them to skip the illustration. Also, they may be asked to compute a statistic for a small data set, then recompute it after the data are changed in some way to illustrate how the statistic reflects those changes in the data. Then a guided computational exercise is presented that shows students how to lay out the information and perform the calculation of a particular statistic. Of course, this is most useful when calculating a statistic by hand or with a hand calculator, but I feel it is helpful to most students to do at least one problem by hand before using a computer. Then a set of exercises with answers is presented, followed by the boxed instructions for using MINITAM and SPSS.

Individual instructors emphasize different aspects of elementary statistics, and many must select a subset of chapters that can be covered in the available time. The organization of the text is straightforward. Part 1 (Chapters 1–6) presents descriptive statistics, including central tendency, variability, relative position, regression, and correlation. Part 2 (Chapters 7–10) deals with elementary inferential statistics, including sampling distributions, the logic of hypothesis testing, elementary parametric tests, and simple analysis of variance. For most courses, these are the core chapters. Part 3 contains more advanced topics that instructors can select to supplement the core material. Most commonly, the chapter on non-parametrics is assigned, but those instructors teaching a more intense course may also add the chapters on research design, probability, and two-factor analysis of variance.

In addition to the colleagues and friends who contributed their advice to earlier editions, I want to thank Mark Appelbaum of the Department of Psychology at Vanderbilt University. He was my chief technical consultant, helping to blend my personal style with the demands of formal theory. I am indebted to the many whose commitment to the course and to the book led them to communicate directly with me over the years; their comments, criticisms,

and suggestions stimulated many of the improvements in this new edition.

Thanks also to the following statistics scholars, who reviewed the manuscript for this edition and provided helpful comments and suggestions: Nancy S. Anderson, University of Maryland, College Park; James Chumbley, University of Massachusetts, Amherst; Susan Donaldson, University of Southern Indiana; Tom Pyle, Eastern Washington University; and Lori Temple, University of Nevada, Las Vegas.

I also appreciate the assistance of Richard Sass with the *Study Guide* and *Instructor's Manual;* and Jeannine Maschak, Catherine Kelley, and Patricia Trainor for preparing the manuscript.

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 Longman Group Ltd., London, for permission to reprint Tables III, IV, VIII, and part of XXXIII from their book *Statistical Tables for Biological, Agricultural and Medical Research* (Sixth Edition, 1974).

Special thanks go to my wife, Rozanne, for her encouragement and for the absence of complaints while she was temporarily widowed for this cause.

Robert B. McCall

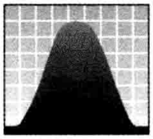
CONTENTS

PART 1

DESCRIPTIVE STATISTICS

One use of statistical methods is to organize, summarize, and describe quantitative information. Such techniques are called **descriptive statistics.** Baseball batting averages, the rate of inflation, and the degree of relationship between heredity and IQ test performance are three examples. This part of the text presents some elementary descriptive statistics.

CHAPTER 1



The Study of Statistics

THE NEWS IS FULL OF STATISTICS: "One-third of American seniors in high school read five or fewer pages per day for school or homework. . . . Only 36% of Americans are satisfied with the quality of the nation's health care and only one in nine are satisfied with the costs, which averaged $6,535 per family in 1991 and are expected to rise to $14,000 by the year 2000. . . . Last month inflation rose a seasonally adjusted .3% to an annual rate of 4.5%. . . . Pitcher Dwight Gooden's ERA is now 2.95. . . . The probability of rain today is 60%."

Such quantities are "statistics," and an elementary knowledge of them is useful to everyone. But the academic study of statistics is much more than these common percentages, rates, averages, or probabilities, and a detailed knowledge of this field is essential to anyone who wants to read or, especially, to conduct research in psychology, education, sociology, economics, or any of many other subjects.

As a field of study, statistics consists of a set of procedures for organizing, describing, and interpreting measurements and for drawing conclusions and making inferences about what is generally true for an entire group when only a few members of the group are actually measured. This chapter explains why the study of statistics is important, discusses the purpose of statistics, and presents the kinds of measurements typically made in the social and behavioral sciences.

WHY STUDY STATISTICS?

A knowledge of statistics is useful for two main purposes: to help understand the common use of statistics in the news, in our jobs, and elsewhere in our daily lives, and to be able to read and understand scientific articles and conduct research in the social and behavioral sciences.

Everyday Use

All of us have a basic understanding of simple percentages, rates, and averages. For example, if we read that the average household income in the United States in 1989 was $36,520, we assume that someone added up the incomes for all the households in the country and divided that total by the number of such families to give the average family income of $36,520.

But many more statistical quantities exist than averages. For example, suppose the same newspaper article reports that the *median* household income was only $28,906. What is the *median*? Why is it different than the average? Which value, the average or the median, is the best index of typical family income? (See Chapter 3.)

Sometimes two simple statistics seem to lead to opposite conclusions. For example, you might read in a magazine that the IQs of adopted children are more closely correlated with those of their biological parents than with those of the parents who reared them. This seems to suggest that heredity plays a more important role in the development of IQ than environment. But the magazine may also report that the same study using the same parents and children found that the IQs of the adoptive children were closer in average value to those of their rearing parents than to those of their biological parents, a fact that seems to suggest that environment is more important than heredity. How can both these facts be true, and what do they say about the contributions of heredity and environment to IQ? (See Chapter 6.)

Or you may wake up one morning and hear on the radio that the probability of rain is 40%, but you look outside and it is already raining. Or you might hear that the probability of rain that afternoon is 70%, so you decide not to use the free tickets you were given to the football game. But, in fact, it doesn't rain at all. In both cases you decide the weather forecaster is incompetent, but actually the forecaster has been correct in predicting the weather and in using the concept of probability. (See Chapters 7 and 12.)

So, even the statistics commonly used in everyday life are often more complicated than they first appear. An introduction to statistics will help you understand and interpret these statistics.

Scientific Use

While statistics are useful to everyone, they are crucial if you want to read a scientific paper, study a science, or conduct scientific research, and this is especially true if the science is behavioral. Undergraduate majors in psychology, for example, are required to take a course in statistics—which may be the only reason you are reading this—and many graduate programs in education, sociology, economics, and other sciences also require statistical training. Why?

First, as we will see below and throughout this text, the social and behavioral sciences are indeed sciences—they collect data about the topics of interest in systematic and objective ways. Those data must be summarized, described, and analyzed by statistical methods, and the conclusions reached form the knowledge base of these disciplines. So it is necessary for anyone wanting to study these disciplines to know some statistics to understand what is presented in textbooks and articles in these fields, even if one never conducts an experiment or makes a scientific observation. Quite simply, you will not be able to read with complete understanding in these areas much past an introductory level without some background in statistics, and the more advanced your studies the more knowledge of statistics you will need. This is why even those graduate programs in psychology that are designed to train professionals to diagnose and treat behavioral problems, not to conduct research, still typically require students to take one year of advanced statistics.

Second, if you are going to make any scientific observations in the social and behavioral sciences, even for an undergraduate thesis or research project, you will need to know and perform some statistical operations. This is so, because almost all numerical information, or data, in the social sciences contains variability. **Variability** refers to the fact that the scores or measurement values obtained in a study differ from one another, even when all the subjects in the study are assessed under the same circumstances. If your teacher were to give a test to all the students in your statistics class, you would not all obtain the same score. Similarly, some children diagnosed as hyperactive improve their classroom behavior while others do not after being placed on drug medication or after being put through an intensive behavioral regimen at home and at school. This dissimilarity in scores and outcomes is variability.

Variability exists in behavioral data for at least three reasons. First, the units (usually people) that behavioral scientists study are rarely identical to one another. A chemist or a physicist assumes that each of the units under study (molecules, atoms, electrons, and so forth) is identical in its composition and behavior to every other unit of its kind. But the behavioral scientist cannot assume that all people will respond identically in a given situation. In fact, the behavioral scientist can count on the fact that they will *not* respond in the same way. The first major source of variability, then, is **individual differences** in the behavior of different subjects.

Second, behavioral scientists cannot always measure the attribute or behavior they wish to study as accurately as they would like. Scores on a classroom examination are supposed to be a measure of how much students have learned, but even a good test has flaws that make it an imperfect measure of learning. Perhaps some questions are ambiguous, or maybe one part of the course material is emphasized more on the test than other parts, thereby favoring those students who happened to study more thoroughly the relevant parts. These and other factors that influence scores or performance and that are associated with how the measurements are made constitute **measurement error**. Almost all behavioral measurements contain some measurement error, this error favors some people more than others, and this fact contributes to variability.

Third, even a single unit (person, animal) usually will not respond exactly the same way on two different occasions. If you give a person several opportunities to rate the attractiveness of advertising displays or the degree of aggressiveness in filmed episodes of each of 15 preschool children, he or she will most likely not assign the same scores on each occasion. This source of variability is called **unreliability**.

Individual differences, measurement error, and unreliability are the major reasons that it is often difficult to draw accurate conclusions from behavioral science data. It is the task of statistics to quantify the variability in a set of measurements; to describe the data for a group of subjects, despite the variability inherent in the measurements; and to derive precisely stated and consistent decisions about the results by quantifying the uncertainty produced by variability. Below we consider the details of some of these functions of statistics.

DESCRIPTIVE AND INFERENTIAL STATISTICS

Statistics is the study of methods for describing and interpreting quantitative information, including techniques for organizing and summarizing data and techniques for making generalizations and inferences from data. The first of these two broad classes of methods is called descriptive statistics, and the second is called inferential statistics.

Descriptive statistics refers to procedures for organizing, summarizing, and describing quantitative information, which is called data. Most people are familiar with some statistics of this sort. The baseball fan is accustomed to checking over a favorite player's batting average; the sales manager relies on charts showing the sales distribution and cost efficiency of an enterprise; the head of a household may consult tables showing the average domestic expenditures of families of comparable size and income; and the actuary possesses charts outlining the life expectancies of people in various professions. These are relatively simple statistical tools for characterizing data, but additional techniques are available to describe such things as the extent to which measured values deviate from one another and the relationship between individual differences in performance of two different kinds. For example, one can describe statistically the degree of relationship between the scores of a group of students on a college entrance examination and their later grades in college.

The second major class of statistics, **inferential statistics**, includes methods for making inferences about a larger group of individuals on the basis of data actually collected on a much smaller group. For example, suppose a researcher wanted to know if administering the medication, Ritalin, helps children who have been diagnosed as having attention-deficit hyperactivity disorder (ADHD) to accomplish more of their schoolwork in class. Without medication, such children typically have difficulty paying attention to tasks over even moderate periods of time, and therefore they often do not finish assignments. They also have special difficulty with problems that require them to perform a sequence of steps to obtain the answer. Ritalin presumably slows them down so they can sustain their attention over longer periods, complete more work, and work more accurately.

Suppose, with their parents' permission, 40 third-grade children with the ADHD diagnosis are randomly divided into two groups. One group is given Ritalin treatment under the care of a physician while the other group is administered a *placebo,* a look-alike drug that actually has no effect on the children. Then, after a period of one month to adjust to the treatment, the number of correctly worked problems assigned in class is counted over a four-week period. The Ritalin group worked an average of 63 problems correctly while the placebo group worked an average of 56. The research question is: Does Ritalin improve school performance, at least under these circumstances?

Notice that the question is, "*Does* Ritalin improve performance?" not "*Did* Ritalin improve performance?" The word *does* is much more general than the word *did.* The results already show that the Ritalin group *did* better than the placebo group, 63 *vs.* 56. The scientific question is whether this difference is likely to happen again if two new groups of children are studied, and then again with two more groups, and again and again until all children with ADHD are studied. In short, *does* Ritalin have this effect for children in general? "All children with ADHD" constitutes the **population** of children with ADHD, while the specific group of 40 children studied constitutes a **sample** from this population. So the scientific question translates into whether Ritalin has an effect in the population, a conclusion that the scientist must *infer* from the results of the sample.

Of course, inferring what would be true for the population of all ADHD children on the basis of results from a small sample of 40 represents something of a guess, and we feel some uncertainty about drawing such a conclusion. For example, 63 problems is definitely more than 56, but we know that if we repeated the