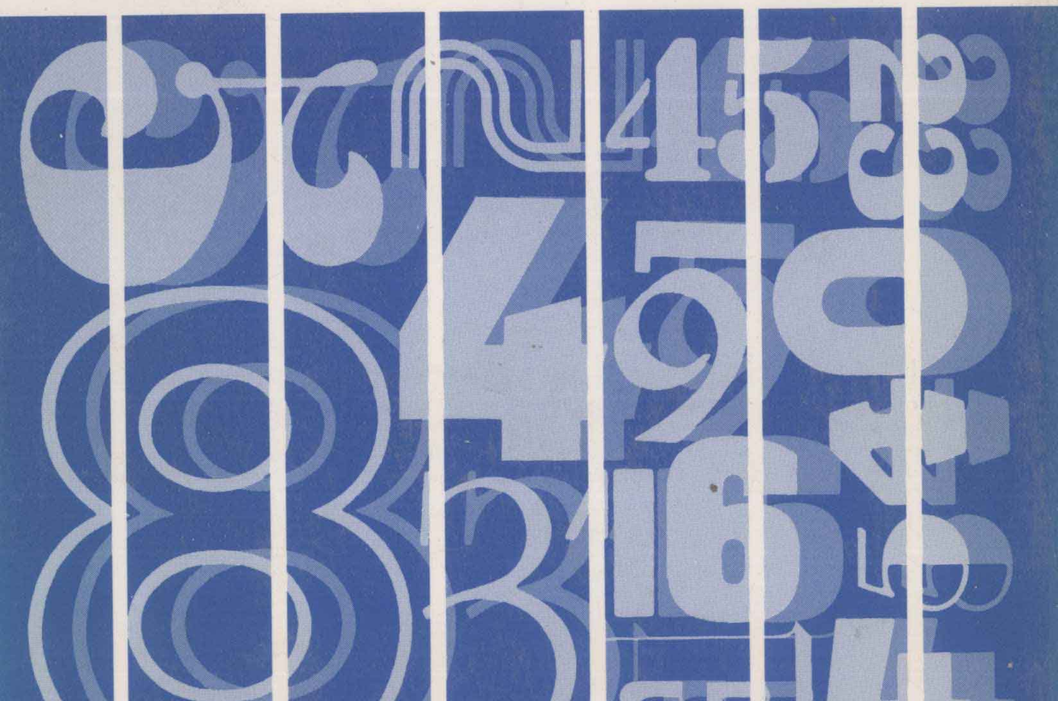
William L. Hays

Statistics

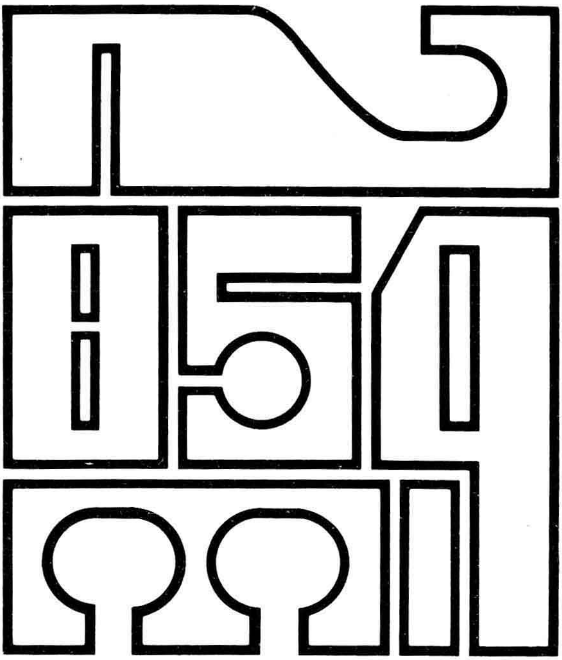
Third Edition



Holt-Saunders International Editions

Statistics

THIRD EDITION



William L. Hays
The University of Texas at Austin

Holt, Rinehar and Winston

New York Chicago San Francisco Philadelphia
Montreal Toronto London Sydney Tokyo
Mexico City Rio de Janeiro Madrid

Holt-Saunders Japan, Ltd.

Library of Congress Cataloging in Publication Data

Hays, William Lee, 1926–
Statistics.

Second ed. published in 1973 under title: Statistics for the social sciences.
Bibliography: p. 669
Includes index.
1. Social sciences—Statistical methods. 2. Statistics. 3. Psychometrics. I. Title.
HA29.H352 1981 519.5'0243 80-27220

ISBN 0-03-056706-8 (US College Edition)
ISBN 4-8337-0004-2 (Holt-Saunders International Edition)

Address correspondence to: 383 Madison Avenue, New York, N.Y. 10017

Printed in Japan, 1981

1 2 3 4 144 9 8 7 6 5 4 3 2 1

CBS COLLEGE PUBLISHING
Holt, Rinehart and Winston
The Dryden Press
Saunders College Publishing

Preface

In 1963, when this book was first published under the title, *Statistics for Psychologists,* its aims were set forth as follows:

> This book represents an attempt to give the elements of modern statistics in a relatively nonmathematical form, but in somewhat more detail than is customary in texts designed for psychologists, and with considerably more emphasis on the theoretical rather than the applied aspects of the subject. It is designed as a text of at least an intermediate level of difficulty. I have felt for some time that the serious student in experimental psychology both needs and wishes to know somewhat more of the language and concepts of theoretical statistics than is provided in the usual "cookbook" of statistical methods. Granted that a real understanding and appreciation of mathematical statistics requires a considerable degree of mathematical training and sophistication, a great deal of statistical theory can be got across to the serious student familiar only with elementary algebra, provided that a relatively detailed exposition of the concepts accompanies the mathematical treatment, and provided that the student is genuinely interested in acquiring a grasp of this subject.

The use that the book has received in the social sciences generally over the years apparently justifies my hope that there are students and teachers in many fields who find this approach worthwhile. Essentially, the aims of this third edition remain the same.

Nevertheless, when I set about preparing this new edition, my chief concern was to make it somewhat more concise. Among other things, having returned to regular teaching after a 15-year hiatus spent in university administration, I have recently grown much more sensitive to the need to make this text portable! To this end, I have shortened a great many sections, and eliminated a fair number, while trying to maintain the essential coverage of the earlier editions. Three entire chapters have been omitted: the chapter on set and function theory covered material that is now such a commonplace of school mathematics that it could hardly serve its original purpose; the chapter on joint variables and independence could be split up among other chapters; and the chapter on Bayesian methods, though still representing an approach I believe to be potentially important for the social and behavioral sciences, could not treat the subject fully in the limited space available. Furthermore, I am neither so sanguine nor quite so choleric about certain issues as I was a few years ago, and thus feel no need to waste further words on them. In all, then, a good many things have been shortened or omitted, with, I trust, little damage to the coverage of topics that this book attempts.

In addition, a good bit of reworking of the chapters on analysis of variance and on regression has been done. The ready availability of computer programs for multiple regression and for multivariate analysis generally is giving such methods a far more ubiquitous role in research than they formerly enjoyed. In order to understand and to take advantage of the many options these methods present, the student needs some early groundwork in the general linear model, and especially in the essential connections between multiple regression and analysis of variance. Again, because of the limitations of space and the level of preparation expected of students using this book, the coverage of these topics has to be rather cursory. Any adequate treatment of linear models and regression theory almost demands the use of matrix and vector theory, and I was tempted to go in this direction. Although I have included a small section on vector operations in Appendix C, in order to introduce the notions of orthogonalization, I resisted the impulse to go further on the grounds that a treatment by matrix theory belongs in another book, where the concepts can be introduced early and thoroughly.

I have also tried to include more, and simpler, problems in this edition than in the last. The 570 or so problems included represent an abosolute increase of 25 percent over the second edition, and, for the chapters included in both editions, nearly a 50 percent increase in number. Solutions to the odd-numbered exercises are printed at the end of the book. A separate answer key with solutions to the even-numbered exercises is available to instructors who adopt the third edition.

The almost universal use of hand-held calculators also had some bearing on my decisions about what to include. Thus, a lengthy table of squares and square roots now has about as much use in a statistics text as the multiplication tables. Although I have retained some other tables, such as factorials and powers of e, these are now such common operations on calculators that they probably are superfluous as well. A technique such as the log-linear analysis in Chapter 15 looks computationally formidable, but these computations are now relatively easy to do on a calculator with a natural log function, which a great many have. Details of the computation of means, variances, and even correlation coefficients become relatively less important in an age when these operations are preprogrammed on many inexpensive calculators.

As before, this book is aimed primarily at the first-year graduate student in one of the social or behavioral sciences. I have therefore assumed that the student probably will have had at least one undergraduate statistics course, and that the present course will be followed by more specialized advanced courses, such as experimental design. However, I believe that the level is generally elementary enough to be followed by an apt student without specific preparation in statistics, and yet advanced enough to give the student a start in relatively simple research and data analysis. Ideally, the text will be used in a two-semester course. However, many of the sections, especially in the early part of the book, are sufficiently self-contained that they can be omitted without serious loss of continuity. Thus, it is entirely possible for the teacher to cut and tailor the topics covered to fit the requirements of a one-semester course, especially if the students already have some background in this area.

I wish to indicate my indebtedness and to offer my thanks to the very many students and teachers who have contacted me through the years with comments and suggestions for improvement of this text. I am especially grateful to all of those who have identified and helped me to correct errors of various kinds. I hope that a new generation of students and teachers will be willing to give me the same assistance.

Once again I extend my sincere thanks to the late Professor E. S. Pearson and the trustees of *Biometrika* for their kind permission to use tables from the *Biometrika Tables for Statisticians*, (Vol. 1, 3rd ed.) and to Professor R. S. Burington and the McGraw-Hill Company for graciously allowing me to reprint the table of binomial probabilities from R. S. Burington and D. C. May, *Handbook of Probability and Statistics with Tables* (2nd ed.). I would also like to thank the reviewers of this text: Robert G. Malgady and Stanley A. Mulaik. Finally, to my wonderful, and wonderfully understanding, Palma, Leeann, and Scott, goes more appreciation than I can ever adequately express.

W. L. H.
Austin, Texas
December 1980

Contents

Chapter 10 The General Linear Model and the Analysis of Variance 325

Chapter 11 The Analysis of Variance: Models II and III, Random Effects and Mixed Models 376

Chapter 12 Comparisons Among Means 413

INTRODUCTION

ON THE NATURE AND THE ROLE OF INFERENTIAL STATISTICS

The word, "statistics," came into English from Latin and German, and ultimately derives from the same Indo-European root which gave us "standing," "status," "state," and even "understand." In the minds of most people, "statistics" has a lot in common with these related words, meaning roughly, a description of "how things are." It is, of course, true that a part of the theory of statistics concerns effective ways of summarizing and communicating masses of information which describe some situation. This part of the overall theory and set of methods is usually known as "descriptive statistics."

Although descriptive statistics form an important basis for dealing with data, a major part of the theory of statistics is concerned with another question: How does one go beyond a given set of data, and make general statements about the large body of potential observations, of which the data collected represent but a sample? This is the theory of *inferential statistics*, with which this book is mainly concerned.

Applications of inferential statistics occur in virtually all fields of research endeavor—the physical sciences, the biological sciences, the social sciences, engineering, market and consumer research, quality control in industry, and so on, almost without end. Although the actual methods differ somewhat in the different fields, the applications all rest on the same general theory of statistics. By examining what the fields have in common in their applications of statistics we can gain a picture of the basic problem studied in mathematical statistics. *The major applications of statistics in any field all rest on the possibility of repeated observations or experiments made under essentially the same conditions*. That is, either the researcher actually can observe the same process repeated many times, as in industrial quality control, or there is the *conceptual* possibility of repeated observation, as in a scientific experiment that might, in principle, be repeated under identical conditions. However, in any circumstance where repeated observations are made, even though every precaution is taken to make conditions exactly the same the results of observations will vary, or tend to be different, from trial to trial. The researcher has control over some, but not all, of the factors that make outcomes of observations tend to differ from each other.

When observations are made under the same conditions in one or more respects, but they give outcomes differing in other ways, then there is some *uncertainty* connected with observation of any given object or phenomenon. Even though some things are known to be true about that object in advance of the observation, the experimenter cannot predict with complete certainty what its other characteristics will be. Given enough repeated observations of the same object or kind of object a good bet may be formulated about what the other characteristics are likely to be, but one cannot be completely sure of the status of any given object.

This fact leads us to the central problem of inferential statistics: *in one sense, inferential statistics is a theory about uncertainty, the tendency of outcomes to vary when repeated observations are made under identical conditions*. Granted that certain conditions are fulfilled, theoretical statistics permits deductions about the *likelihood* of the various possible outcomes of observation. The essential concepts in statistics derive from the theory of probability, and the deductions made within the theory of statistics are, by and large, statements about the probability of particular kinds of outcomes, given that initial, mathematical, conditions are met.

Mathematical statistics is a formal mathematical system. Any mathematical system consists of these basic parts:

1. A collection of undefined **"things"** or **"elements,"** considered only as abstract entities;
2. A set of undefined **operations,** or possible relations among the abstract elements;
3. A set of **postulates** and **definitions,** each asserting that some specific relation holds among the various elements, the various operations, or both.

In any mathematical system the application of logic to combinations of the postulates and definitions leads to *new* statements, or theorems, about the undefined elements of the system. *Given* that the original postulates and definitions are true, then the new statements *must* be true. Mathematical systems are purely abstract, and essentially undefined, **deductive** structures. In the first chapter we will see that the abstract system known as the theory of probability has this character.

Mathematical systems are not really "about" anything in particular. They are systems of statements about "things" having the formal properties given by the postulates. No one may know what the original mathematician really had in mind to call these abstract elements. Indeed, they may represent absolutely nothing that exists in the real world of experience, and the sole concern may be in what one can derive about the other necessary relations among abstract elements given particular sets of postulates. It is perfectly true, of course, that many mathematical systems originated from attempts to describe real objects or phenomena and their interrelationships: historically, the abstract systems of geometry, school algebra, and the calculus grew out of problems where something very practical and concrete was in the back of the mathematician's mind. However, as *mathematics* these systems deal with completely abstract entities.

When a mathematical system is interpreted in terms of real objects or events, then the system is said to be a **mathematical model** for those objects or events. Somewhat more precisely, the undefined terms in the mathematical system are identified with particular, relevant, **properties** of objects or events; thus, in applications of arithmetic, the number symbols are identified with magnitudes or amounts of some particular property that objects possess, such as weight, or extent, or numerosity. The system of arithmetic need not apply to other characteristics of the same objects, as, for example, their colors. Once this identification can be made between the mathematical system and the relevant properties of objects, then anything that is a logical consequence in the system is a true statement about objects in the model, *provided,* of course, *that the formal characteristics of the system actually parallel the real characteristics of objects in terms of the particular properties considered.* In short, in order to be useful as a mathematical model, a mathematical system must have a formal structure that "fits" at least one aspect of a real situation.

Probability theory and statistics are each both mathematical systems *and* mathematical models. Probability theory deals with elements called "events," which are completely abstract. Furthermore, these abstract things are paired with numbers called "probabilities." The theory itself is the system of logical relations among these essentially undefined things. The experimenter uses this abstract system as a mathematical model: the experiment produces a real outcome, which is called an event, and the model of probability theory provides a value which is interpreted as the relative frequency of occurrence for that outcome. If the requirements of the model are met, this is a true, and perhaps useful result. If the experiment really does not fit the requirements of probability theory as a system, then the statement made about the actual result need not be true. (This point must not be overstressed, however. We will find that often a statistical method can yield practically useful results even when its requirements are not fully satisfied. Much of the art in applying statistical methods lies in understanding when and how this is true.)

Mathematical systems such as probability theory and the theory of statistics are, by their very nature, deductive. That is, formal assertions are postulated as true, and then by logical argument true conclusions are reached. All well-developed theories have this formal, logico-deductive character.

On the other hand, the problem of the empirical scientist is essentially different from that of the logician or mathematician. Scientists search for general relations among events; these general relations are those which can be expected to hold whenever the appropriate set of circumstances exists. The very name "empirical science" asserts that these laws shall be discovered and verified by the actual observation of what happens in the real world of experience. However, no mortal scientist ever observes all the phenomena about which a generalization must be made. Scientific conclusions about what would happen for *all* of a certain class of phenomena always come from observations of only a very few particular cases of that phenomenon.

The student acquainted with logic will recognize that this is a problem of **induction.** The rules of logical *deduction* are rules for arriving at true consequences from true premises. Scientific theories are, for the most part, systems of deductions from basic principles held to be true. If the basic principles are true, then the deductions must be true. However, how does one go about arriving at and checking the truth of the initial propositions? The answer is, for an empirical science, observation and inductive generalization—going from what is true of some observations to a statement that this is true for *all possible* observations made under the same conditions. Any empirical science begins with observation and generalization.

Furthermore, even after deductive theories exist in a science, experimentation is used to check on the truth of these theories. Observations that contradict deductions made within the theory are prima facie evidence against the truth of the theory itself. Yet, how does the scientist know that the results are not an accident, the product of some chance variation in procedure or conditions over which there is no control? Would the result be the same in the long run if the experiment could be repeated many times?

It takes only a little imagination to see that this process of going from the specific to the general is a very risky one. Each observation the scientist makes is different in some way from the next. Innumerable influences are at work altering—sometimes minutely, sometimes radically—the similarities and differences the scientist observes among events. Controlled experimentation in any science is an attempt to minimize at least part of the accidental variation or "error" in observation. Precise techniques of measurement are aids to scientists in sharpening their own rather dull powers of observation and comparison among events. So-called "exact sciences," such as physics and chemistry, have thus been able to remove a substantial amount of the unwanted variation among observations from time to time, place to place, observer to observer, and hence are often able to make general statements about physical phenomena with great assurance from the observation of quite limited numbers of events. Observations in these sciences can often be made in such a way that the generality of conclusions is not a major point at issue. Here, there is relatively little reliance on probability and statistics. (However, as even these scientists delve into the molecular, atomic, and subatomic domain, negligible differences turn into enormous unpredictabilities and statistical theories become an important adjunct to their work.)

In the biological, behavioral, and social sciences, however, the situation is radically different. In these sciences the variations between observations are not subject to the precise experimental controls that are possible in the physical sciences. Refined measurement techniques have not reached the stage of development that they have attained in physics and chemistry. Consequently, the drawing of general conclusions is a much more dangerous business in these fields, where the sources of variability among living things are extremely difficulty to identify, measure, and control. And yet the aim of the social or biological scientist is precisely the same as that of the physical scientist—arriving at general statements about the phenomena under study.

Faced with only a limited number of observations or with an experiment that can be conducted only once, the scientist can reach general conclusions only in the form of a "bet" about what the true, long run, situation actually is like. Given only sample evidence, the scientist is always unsure of the "goodness" of any assertion made about the true state of affairs. The theory of statistics provides ways to assess this uncertainty and to calculate the probability of being wrong in deciding in a particular way. *Provided that the experimenter can make some assumptions about what is true, then the deductive theory of statistics tells us how likely particular* *results should be*. Armed with this information, the experimenter is in a better position to decide what to say about the true situation. Regardless of what one decides from evidence, it *could* be wrong; but deductive statistical theory can at least determine the probabilities of error in a particular decision.

In recent years, a branch of mathematics has been developed around this problem of decision making under uncertain conditions. This is sometimes called "statistical decision theory." One of the main problems treated in decision theory is the choice of a decision rule, or "deciding how to decide" from evidence. Decision theory evaluates rules for deciding from evidence in the light of what the decision maker wants to accomplish. As we shall see in later chapters, mathematics can tell us wise ways to decide how to decide under some circumstances, but it can never tell the experimenter how a decision must be reached in any particular situation. The theory of statistics supplies one very important piece of information to the experimenter: the probability of sample results *given* certain conditions. Decision theory can supply another: optimal ways of using this and other information to accomplish certain ends. Nevertheless, neither theory tells the experimenter *exactly* how to decide—how to make the inductive leap from observation to what is true in general. This is the experimenter's problem, and the answer must be sought outside of deductive mathematics, and in the light of what the experimenter is trying to do.

ABOUT THIS BOOK

This book is addressed to upper division or graduate students in the social and behavioral sciences. Such students are the people who will produce the significant social and behavioral science research in years to come, and who will make up the audience for much of this research. As a part of their professional equipment, these students need to know statistics, at a level beyond an undergraduate course, and just short of the specialized research design and methodology courses needed to round out their graduate programs. Such students are the "you" in this book.

You will soon discover that the main concern in this book is with the theory underlying inferential methods, rather than with a detailed exposition of all the different methods social scientists and others find useful. The author had no intention of writing a "cookbook" that would equip students to meet every possible situation they might encounter. Many methods will be introduced, it is true, and we will, in fact, discuss most of the elementary techniques for statistical inference currently in use. However, in the past few years the concerns of the social scientist have begun to grow increasingly complicated. Theory is growing, and social scientists are turning their attention to new problems and techniques for data analysis that are becoming much more sophisticated than in the past. The statistical analyses required in many such studies are simply not in the "cookbooks." From all indications, this trend will continue, and by the time that you, the student, are in the midst of your professional career it may well be the case that entirely new statistical methods will be required, replacing many of the methods currently found useful.

Furthermore, a true revolution has occurred in the past two decades, deeply affecting the application and the teaching of statistical methodology. This has been brought about by the new generations of computers, which are faster, more flexible, and cheaper to use than anyone would have dreamed only a few years back. Large-scale statistical analysis is now done by computer in almost all research settings.