SOCIAL STATISTICS

SECOND EDITION

HUBERT M. BLALOCK

INTERNATIONAL STUDENT EDITION

Contents

Preface

This text is written primarily for those students of sociology, both advanced undergraduates and graduate students, who actually intend to engage in social research.

In the twelve years that have elapsed since the publication of the first edition, the level of training and the sophistication in applied statistics has undergone considerable improvement not only in sociology but also in political science, anthropology, geography, and social work. Nevertheless, the overwhelming majority of students and practitioners in these fields still lack the necessary mathematical backgrounds to take full advantage of the rapidly accumulating technical literature in mathematical statistics and econometrics. With these basic facts in mind, this text has been written so as to avoid mathematical derivations insofar as possible, and only a quick review of certain algebraic principles listed in Appendix 1 should therefore be sufficient preparation for the average student. Although it is not necessary in a first course in statistics to stress mathematical derivations, the author is convinced that certain basic and fundamental ideas underlying the principles of statistical inference must be thoroughly understood if one is to obtain more than a mere "cookbook" knowledge of statistics. Therefore, there is a relatively heavy emphasis on the underlying logic of statistical inference, including a chapter on probability, with relatively less attention being given to some of the more or less routine topics ordinarily discussed in elementary texts.

One of the most difficult problems encountered in the teaching of applied statistics is that of motivating students, both in enabling them to overcome their fears of mathematics and in learning to apply statistics to their own field of interest. It is for the latter reason that the author has not attempted to cover a wide range of applications but has selected examples of primary interest to sociologists. To some extent, examples

have also been chosen from disciplines which border on sociology: fields such as social psychology, social work, and political behavior. In most instances each new topic has been illustrated by a single example, under the assumption that most students will lose track of the basic line of thought if too many examples are used to illustrate the same point. Additional examples are therefore given in the form of exercises at the end of each chapter. In general, the author has tried to strike a reasonable compromise between the desirability of stating basic principles as clearly and concisely as possible and the necessity of repeating some of the more difficult ideas each time a new topic is discussed. Insofar as possible, new ideas have been introduced gradually, and, equally important, an effort has been made to relate each new topic to those which have preceded it. In so doing, the major goal has been to give an appreciation of the basic similarities underlying many of the most commonly used tests and measures.

Almost all the suggestions I have received from those wishing to help improve the first edition have implied additions to the book, rather than subtractions, and they have also implied that many of the topics originally treated should be discussed more technically. My own position is that sociologists and political scientists, in particular, need greater exposure to the more technical literature on experimental designs and on the use of simultaneous-equation procedures in connection with nonexperimental research. Yet, it became clear that if these materials were added to the original text, it would lose its appeal as an introductory text appropriate for advanced undergraduates majoring in the social sciences. It was therefore decided to treat experimental designs, factor analysis, and simultaneous-equation procedures, as well as other more advanced topics, in a separate text to be written with two of my colleagues, Lewis F. Carter and Krishnan Namboodiri.

Included in the text are a number of sections, paragraphs, and exercises which are either conceptually difficult or which presuppose that the student is reasonably familiar with topics ordinarily covered in courses on research methods. These portions of the text have been marked with an asterisk (*) and may be skimmed on first reading or omitted entirely. Instructors using the text for a one-semester course may wish to indicate that students should omit these materials.

There has been a slight upgrading in terms of technical coverage, but the basic structure of the text has not changed. There are a few modifications in the section on descriptive statistics. The author has added discussions of assumptions and basic concepts, hoping to clarify the relationship between the statistical models and the real world with which the social scientist must deal. In addition to these changes, this edition

contains discussions of several procedures, tests, and measures that have been increasingly used during the decade of the 1960s.

Chapter 9 on probability has been expanded to include discussions of permutations, tree diagrams, Bayes' theorem, calculations involving conditional probabilities, and the notion of expected values. And to Chapter 10, which includes a discussion on the binomial distribution, the author has added brief discussions of the multinomial distribution, the hypergeometric distribution, and the Poisson distribution. These additional topics should make the transition to texts specifically oriented to nonparametric statistics a much simpler task.

The author has also extended the coverage of nonparametric techniques to include the Friedman test for two-way analysis of variance with ranks; gamma and d_{yx} as measures of ordinal association; a test for interaction involving the difference of differences of proportions; and standardization in the case of nominal-scale procedures. There is also an expanded discussion of the properties of the several ordinal measures and partialling techniques with ordinal scales.

Similarly, the discussion of parametric approaches has been expanded to include discussion of the assumptions for the general linear model and coverage of the dummy variable approach as an alternative perspective on analysis of covariance. The basic essentials of the theory underlying the use of linear combinations have also been introduced and applied to the discussions of the standard error of the mean, difference of means, difference of differences of proportions, and the use of orthogonal comparisons in the case of multiple samples.

In an effort to help the reader to see the overall picture, the author has included a summary table of tests and measures, which appears on the front, inside cover, and extended summaries at the end of Chapters 2, 14, and 20.

Numerous persons have contributed to this revision, but the author would especially like to thank Richard G. Ames, Erica Borden, and Louis Goodman for providing extensive comments on the entire manuscript.

For assistance in the preparation of the first edition, I would especially like to thank those students and colleagues at the University of Michigan who read and suggested improvements in various drafts of the book. To Richard T. LaPiere, Sanford Dornbusch, Robert Ellis, Santo Camilleri, and Theodore Anderson my appreciation for reading and criticizing the original manuscript. For proofing, typing, and checking computations, I would like to thank Ann Blalock, Diane Etzel, Ann Laux, and Doris Slesinger. My thanks also to Daniel O. Price, who deserves the major credit for stimulating my interest in statistics.

I am indebted to Professor Sir Ronald A. Fisher, Cambridge, to Dr.

Frank Yates, Rothamsted, and to Messrs. Oliver and Boyd Ltd., Edinburgh, for permission to reprint Tables III, IV, and V from their book *Statistical Tables for Biological, Agricultural and Medical Research*. I am equally grateful to those other publishers and authors, acknowledged in the appropriate places, who kindly gave their permission for use of various tables and computing forms.

Hubert M. Blalock, Jr.

Social Statistics