STATISTICAL ANALYSIS

BY

EDMUND E. DAY

PROFESSOR OF ECONOMICS, AND DEAN OF THE
SCHOOL OF BUSINESS ADMINISTRATION
UNIVERSITY OF MICHIGAN

New York
THE MACMILLAN COMPANY
1925

PRINTED IN THE UNITED STATES OF AMERICA

Set up and electrotyped. Published November, 1925.

Norwood Press
J. S. Cushing Co. — Berwick & Smith Co.
Norwood, Mass., U.S.A.

TO MY COUNSELLOR AND FRIEND

ALLYN A. YOUNG

WHOSE WORK IN STATISTICS HAS COMBINED IN

RARE DEGREE MASTERY OF TECHNIQUE

BREADTH OF SCHOLARSHIP AND

SOUNDNESS OF JUDGMENT

PREFACE

THIS book affords a general introduction to statistical method. It is intended primarily for those who are making their first acquaintance with the subject. The requirements of classroom use have been kept constantly in mind, and every effort has been made to make the material teachable. At the same time, the treatment of the subject is comprehensive and designed to start the intelligent student on the road to a complete grasp of general statistical method.

The emphasis of the book is upon the analysis, rather than upon the collection and tabulation, of statistical material. It is the author's opinion that the great majority of students of statistics are not directly interested in the technique of collection and tabulation. They are concerned with these phases of statistical method only in so far as they affect the subsequent analysis of data. Furthermore, the more important aspects of collection and tabulation can be fully understood only after the nature and purposes of analysis have been mastered. (True, some acquaintance with the technicalities of collection and tabulation is desirable at an early stage of the study of analysis; consequently brief accounts of these are given in a series of appendices.) In general, therefore, attention may be profitably focused at the outset on the problems of analysis, and this is the plan of the present volume.

Special care has been taken throughout the book to make clear the logical relationships of the parts. This results in rather more formalism than would be desirable in any treatment other than that of an introductory text. But it is highly important at the beginning of the study of statistics to learn to place correctly in the general scheme of statistical method the processes that are to be undertaken in any particular analysis. The student should be made to see how the individual parts are related to the whole.

In more extended form, the title of this book might be given as "The Logic of Statistical Analysis." Several years of experience in teaching statistics have led the author to believe that it is much more important

for the beginning student to learn the logical setting of statistical analysis than it is for him to gain facility in any set of technical processes. Some command of technical procedure, of course, is necessary. But, after all, the objectives of statistical work lie in the interpretation of statistical results; and interpretation can never be wisely undertaken save with full recognition of the logical implications and limitations of the processes that have gone before. At the risk, perhaps, of being somewhat meticulous, if not pedantic, the present treatment has been written with the avowed intention of making perfectly clear the *logic* of statistical procedure.

In view of this underlying purpose, it has not been thought necessary to enter upon the more refined mathematical phases of statistical method. Nothing in the book should prove inconsistent with the findings of the most advanced mathematical statistics. But no command of advanced mathematics is expected of those for whom the book has been written. For those who have special interests or preparation along mathematical lines, supplementary studies can be readily undertaken. Furthermore, it is to be hoped that some of those who secure through this book their introduction to statistics will be encouraged to pursue the subject into its fascinating and important mathematical ramifications. A serviceable start in the understanding of statistical analysis may be made, however, with a modicum of purely mathematical procedure, and in the present exposition of the subject the mathematical requirements have been reduced to a minimum.

Attention should be called to one special feature of the book: the treatment of methods of graphic representation. It is customary in texts on statistical method to devote special chapters to statistical graphics. From a pedagogical point of view, this plan has certain advantages. It involves, however, the serious danger of divorcing the practices of statistical graphics from the processes of statistical analysis. Most of the abuses of statistical graphics have occurred because diagrams have not been made to exhibit accurately the results of analysis. A closer inter-relation of analytical and graphic practices is desirable. In the present text, therefore, questions of graphic method are dealt with wherever they arise in the explanation of the manifold phases of analysis.

It is quite impossible for the author to state explicitly the extent of his indebtedness to others. With other workers in the field, he has profited, in part perhaps unconsciously, from the labors of many scholars.

However, some of those whose contributions have been found especially valuable should be specifically mentioned. All readers who are at all familiar with the literature of the subject will quickly sense the author's heavy indebtedness to such eminent authorities as Bowley, Fisher, Mitchell, Pearson, Persons, and Yule. Numerous footnotes throughout the book indicate sources which have been explicitly utilized at particular points. The author's colleagues, Professors O. W. Blackett, J. P. Mitchell, and C. S. Yoakum, have examined the manuscript and offered numerous valuable comments. Former students — among whom perhaps should be particularly mentioned Professors W. A. Berridge and M. B. Hexter — have made many helpful suggestions. Finally, the author's secretary, Miss Carolyn E. Allen, has given invaluable and tireless assistance without which the book could hardly have been published at this time, if at all.

EDMUND E. DAY

ANN ARBOR, MICHIGAN
September 20, 1925

TABLE OF CONTENTS

INTRODUCTION

FORMULATION OF STATISTICAL DISTRIBUTIONS

ANALYSIS OF FREQUENCY DISTRIBUTIONS

ANALYSIS OF PAIRED VARIABLES: CORRELATION

ANALYSIS OF SPATIAL SERIES

ANALYSIS OF TIME SERIES

TABLES

CHARTS