JOHN I. GRIFFIN

STATISTICS

METHODS AND APPLICATIONS

JOHN I. GRIFFIN

Professor of Economic Statistics
Bernard M. Baruch School of Business and Public Administration
The City College
of
The City University of New York

TISTICS

Methods and Applications

HOLT, RINEHART AND WINSTON · NEW YORK

Library of Congress Catalog Card Number: 62–8783
23284–0112
Printed in the United States of America

Preface

In 1869 Adolphe Quetelet, who merged the methods of political economy, mathematics, and government statistics into one powerful tool for investigating social phenomena, which we now call statistics, drew up a list of 180 different definitions of the term statistics. Even today, statisticians define the area of their professional interest and competence in different ways. The very extent of statistics makes it difficult to define. It is both a science and an art. It is a vital branch of scientific method. It is widely used in most sciences, business, government, and ordinary life. The early statisticians, who calculated the probabilities associated with games of chance and who developed the theory of errors of observation, could not visualize the extent to which statistical methods would become an indispensable part of the intellectual equipment of an educated man.

This text is designed to accomplish two major purposes, which complement each other. Within the established curriculums of American colleges and universities, formal training in statistics for most students is limited to one course, consisting of one or two semesters. The burden imposed upon the instructor in such a course is heavy, indeed. He has the responsibility not only of instruction in specific techniques but more particularly of conveying to the student a sense of the exciting new developments in statistics and the contributions which the science of statistics can make to the student's life as scientist, business manager, government administrator, and, not least of all, as citizen. This text seeks to show the student that statistics is a rapidly growing subject, fast developing a philosophy and body of techniques as far removed from the endless masses of descriptive statistics, "Die Tabellen Statistik," as the flintlock is from the I.C.B.M. The text provides the student with a body of specific techniques he can apply immediately to problems which he encounters in his college courses and which will be relevant to his subsequent life experience.

This text is divided into twenty chapters and seven appendixes, and provides sufficient material for a one-year course in statistics. The emphasis is general so that students majoring in business, economics, or any of the social sciences will find the illustrative material familiar. This is not a text in *business* statistics or *economic* statistics but in *statistics*. The distinction is an important one, since the objective of this text is to introduce the student to statistics rather than to particular subject-matter problems. The illustrative data, however, are real data taken from sources reflecting research in fields of direct subject-matter interest to the students. Thus, not only are techniques illustrated but also something is learned about the types of

data with which statisticians are concerned. In this manner, methods and applications are learned at the same time.

The twenty chapters of this text are so arranged that the major aspects of modern statistics are presented in a logical sequence, and the chapters should be studied in sequence. A one-semester course should provide sufficient time to cover about fourteen of the chapters, omitting all of Chapters 9, 12, 15, 16, 19, and 20. Depending upon the particular interests of the students, a choice may be made between the chapters dealing with Sampling Methods in Auditing, Statistical Quality Control in Production and Management, and Statistical Methods in Forecasting and Market Research. In all chapters the level of mathematical sophistication is such as to make a professional mathematician quite unhappy, but it is assumed, realistically I believe, that not much mathematical competence beyond secondary school algebra can be taken for granted. It is important that the student does not become self-conscious about mathematical formulas and realizes that mathematical notation is merely a shorthand and not an end in itself. Occupational specialization in the statistical field, as elsewhere, results in a distinction between those who are able to develop path-breaking new theories and those who are competent in the understanding and application of the existing body of scientific knowledge. This book is not addressed to those in the first category. Its objectives are much more modest.

Several features of this text may be pointed out, which should facilitate effective use by students. The chapters are divided into numbered sections, and cross references to material throughout the book are always made by reference to the section number. The statistical tables in Appendix F are keyed to the sections in which the proper use of the particular table is discussed and illustrated. Throughout the chapters **boldface** type has been employed to call attention to a new term, which is defined at that point in the book. Thus, a student may review the essential technical vocabulary by noting the terms in bold face. In each chapter a number of signals—  —have been inserted. These signals indicate that the student should turn to the self-reviewing workbook which is now available for use with this text. This separately available workbook not only performs the usual functions of a workbook but is also a self-instruction tool prepared on the basis of the new techniques of programmed learning. A kit of ten transparent slides, suitable for use on an overhead projector, has been prepared from the illustrations in this text, and instructors interested in this type of visual instruction should communicate with the publisher.

The frequent references to the great names in the history of statistical science are designed to give the student some acquaintance with the "founding fathers." It seems to me that the teaching of a concept can be made more effective if the student associates the concept with the scientist who developed it. On the new frontiers of the science, such as decision theory, my approach has been essentially conservative. That is, the newer methods of analysis are not presented as in conflict with the principles underlying the traditional theory of Neyman and Pearson. In a sense all statistics is concerned with decision making and not just special applications. The precise position to be occupied by some of the newer work in decision theory in the

entire body of general statistical methods is not completely clear at this time. The best preparation for further intensive work in decision theory in its applications to management and public policy is, in my opinion, a good understanding of accepted theory and practice as developed in the post Karl Pearsonian period.

The sources of data, tables, and concepts are given in the text and in the acknowledgments. If I have inadvertently adopted material without specific recognition of its source, I can only plead that every author of an introductory textbook is in debt to all his predecessors. My greatest obligation is to my students, who, during the past two decades, have challenged my ability to show them the power and excitement of modern statistics. For those to whom statistics has appealed as a professional career, I hope that the foundation has proven sound and that, in like fashion, many readers of this text will rapidly outgrow this book.

J. I. G.

New York, N. Y.
January, 1962

Acknowledgments

In the preparation of a general text in statistics intended for an introductory course for students with no special training in mathematics, it is necessary to draw upon many illustrations of the application of statistical methods made by government agencies and business firms. The generosity with which permission has been granted to reproduce illustrative material is acknowledged with gratitude. The many examples drawn from federal government documents are indicated by the source note identifying the department of issue. The individual statisticians in the several agencies have been most kind in providing materials. The general obligation of the author to the specialists in the field of statistics, as well as to previous writers of textbooks, is suggested, although not capable of detailed specification, by the list of references in Appendix E, which lists those books and articles that have been of major importance in shaping the author's approach to the subject of statistics.

Specific obligations for permission to reproduce materials are due to the Sanborn Map Company for the map which is used as the endpaper at the front of this book. Professor Herbert Arkin has given permission to reproduce certain new tables which he has prepared; these are Appendix Tables F-2, F-3, F-15, F-16, and F-17. Mr. Steve A. Demakopoulos, of the Remington Rand Division of Sperry Rand Corporation, ran on a UNIVAC the time series problem used in Chapter 18. Other machine companies that generously provided material are the International Business Machines Corporation, Royal McBee Corporation, and the Monroe Calculating Machine Company, Inc. Professor E. S. Pearson kindly granted permission to reproduce certain tables from *Biometrika Tables for Statisticians*, reproduced as Appendix Tables F-12 and F-19 and as Charts 10-2, 10-3, 14-4, and 14-5. The Biometrika trustees have granted permission to reproduce the illustration of the Quincunx in Chapter 8 and "Student's" illustrations in Section 6.5. Professor M. S. Bartlett has given permission to reproduce Appendix Table F-13 which shows selected pages from Tracts for Computers No. 26. Chart 13-5 is reproduced from *Introduction to the Theory of Statistics* by G. U. Yule and M. G. Kendall, Charles Griffin & Co., Ltd., by permission of author and publishers. Professor Maurice G. Kendall and Charles Griffin & Co., Ltd. have granted permission to reproduce Chart 17-7 from *The Advanced Theory of Statistics*. The author is indebted to Professor Sir Ronald A. Fisher, F.R.S., Cambridge, to Dr. Frank Yates, F.R.S., Rothamsted, and to Messrs. Oliver & Boyd, Ltd., Edinburgh, for permission to reprint Table VII from their book *Statistical Tables for Biological, Agricultural and Medical Research;* also to Professor Fisher to quote from his book *Design of Experiments* the extract in Section 2.6.

The American Management Association has given permission to use the material "What's Ahead? The Gentle Art of Business Forecasting," prepared by Al Hormel and Lydia Strong, which is reproduced as Chart 18-1. The American Society for Testing and Materials has given permission to reproduce Appendix Table F-21 as well as the Standard Recommended Practice statement, reprinted in full in Section 9.7. These ASTM Standards are under constant surveillance by technical committees and are subject to change, to keep abreast of changes in technology; the Standard Recommended Practice in respect to probability sampling was made Standard in 1961.

The author is indebted to the American Association of Railways for Table 12-1; to the American Telephone and Telegraph Company for Chart 9-7; to Professor Richard V. Andree for Chart 3-12; to the Bank of Montreal for permission to use Chart A-9; to the Cleveland Trust Company for permission to reproduce Charts 18-15 and 18-16; to the Codex Book Company, Inc., for permission to use certain of their graph papers; to the Curtis Publishing Company for Table 20-2; to D. Van Nostrand Company, Inc., and E. C. Molina for permission to abstract selected values of the Poisson distribution, which are shown in Appendix Table F-9; to Dover Publications, Inc., for Chart 16-2; to Francis I. duPont & Co. for Charts 19-1, 19-2, and A-3; to the Econometric Society for permission to extract from an article by the late Charles F. Roos, used in Section 20.2; to the Eno Foundation for permission to use the questionnaire shown as Table 20-3; to the Federal Reserve Bank of Richmond for permission to reproduce the diagram in Section 3.5; to the Free Press of Glencoe for Appendix Tables F-1 and F-10; to the Graduate School of the United States Department of Agriculture for Chart 10-1; to Harvard University Press for Appendix Table F-8; to the Editor of the Journal of the American Statistical Association for Chart 18-7 and Appendix Tables F-11, and F-25; to P. H. Kutschenreuter, of the United States Weather Bureau, for Chart 18-14; to McGraw-Hill Book Company, Inc., for permission to reproduce Chart 9-1 and Appendix Table F-20; to Professor Frederick C. Mills for permission to use Appendix Table F-4; to the National Bureau of Economic Research for Tables 18-5 and 18-6 and for the List of Sources of Current Data used in Section 18.6; to the National Industrial Conference Board for permission to reproduce Charts 18-4 and 18-5; to the National Planning Association for permission to reproduce Chart 20-1; to Fletcher Pratt for the tables of frequency of occurrence of letters used in Section 4.4; to the Railway Progress Institute for permission to use Chart 20-3; to John Wiley & Son, Inc., for Appendix Tables F-22 and F-24.

J. I. G.

When I consider a single state, I discern a vast number of things actually to be found therein. Among them are some which concern obviously its prosperity either in obstructing it or contributing to it. Such things we might call "Staatsmerkwürdigkeiten" (the remarkable things of the state). The totality of these "Staatsmerkwürdigkeiten" of a kingdom or a republic, makes up its constitution in the broadest sense, and the account of such constitutions of one or more states is "statistik."

GOTTFRIED ACHENWALL, 1749

Statistics can give us three kinds of knowledge: (1) Statistics give us knowledge which can be obtained by mere enumeration or by a succession of enumerations. (2) Statistics, in the course of these enumerations, often betray certain fixed relations which have the character of natural laws. (3) Statistics are sometimes able to trace relations of cause and effect which can be ascertained in no other way.

RICHMOND MAYO SMITH, 1888

The conception of statistics as the study of variation is the natural outcome of viewing the subject as the study of populations; for a population of individuals in all respects identical is completely described by a description of any one individual, together with the number in the group. The populations which are the object of statistical study always display variation in one or more respects. To speak of statistics as the study of variation also serves to emphasize the contrast between the aims of modern statisticians and those of their predecessors.

RONALD AYLMER FISHER, 1925

Contents

List of Tables

List of Charts