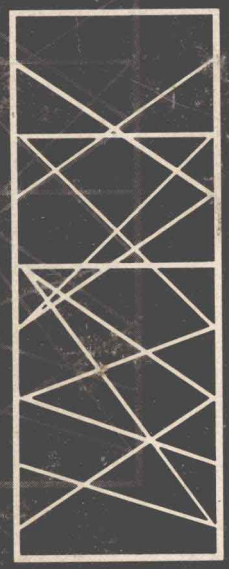
Dorothy C. Adkins



STATISTICS

An Introduction for Students in the Behavioral Sciences

PSYCHOLOGY SERIES

With unique self-testing feature that keeps students immediately and continuously informed of progress at strategic points.

STATISTICS

An Introduction for Students in the Behavioral Sciences

DOROTHY C. ADKINS
The University of North Carolina

CHARLES E. MERRILL BOOKS, INC.
Columbus, Ohio

Library of Congress Catalog Card Number: 64-17176

Printed in the United States of America

Preface

The principal aims of this book are described in some detail in the introductory chapter. In brief, it is intended primarily as a textbook for a 3-semester-hour (or 5-quarter-hour) course in elementary statistics for students in the behavioral sciences, especially psychology, education, and sociology. No prior training in statistics and no knowledge of mathematics beyond the equivalent of one high-school algebra course are assumed.

The entire book can be completed in one semester although some teachers may wish to modify particular topics or sections to their needs. It has been tried for an undergraduate course consisting of two hours of lecture and two hours of laboratory per week. In the laboratory, which was equipped with a variety of electric calculating machines, the student obtained supervised practice in working problems and otherwise applying statistical techniques. The book could also be used in a course entailing three hours of combined lecture and discussion per week with regularly assigned problems.

Perhaps the chief claim to novelty for this presentation is the incorporation into the textual material of multiple-choice questions. The reader, having moistened the dot that corresponds to the answer which he has selected, is at once informed of the correctness of his choice by a color change for right answers. If a choice is correct, the dot turns green. The idea for use of such a device came from some of the pioneer work of one of my former teachers at The Ohio State University, Dr. Sidney L. Pressey, who early gave counsel and encouragement to the present venture. Provision for questions at strategic places in the text, with immediate information as to whether or not the answers selected are correct, not only helps to overcome the initial resistance of many students to quantitative thinking but also serves as a positive motivating factor for students at all levels.

Although the book was written for use in conjunction with a course, the self-testing feature should also make it effective for self-instruction. In addition to the exercises within the textual material, there are review questions at the end of each chapter or every two chapters, which also use the color-changing device to inform the reader of his progress.

Appreciation is expressed to the hundred-odd students of The University of North Carolina who in the past year have been subjected to drafts of this book. They have freely pointed out all sorts of trouble-spots: ambiguities, inconsistencies, typographical errors, numerical mistakes, gaps, needless difficulty, awkwardness of style, and the like.

My thanks are owed to many others. Intermixing textbook material with multiple-choice questions intended to provide emphasis or provoke thought at the same time that they inform the reader of his progress is not so simple a matter as had been envisaged. The preparation required nearly twice as much time and effort as would be anticipated for an ordinary textbook. My special gratitude, therefore, is expressed to another Ohio State University professor, Dr. H. A. Toops, who had me compute 22,500 phi coefficients as a footnote to a Ph.D. thesis at a time in history when modern high-speed computing facilities were but a gleam in the inventors' eyes. After completion of that two-year chore, other tasks, by comparison, have never seemed unduly burdensome. Lacking such inurement, I doubt that I would have undertaken or completed the present work. Any reader who has been exposed to Dr. Toops' teaching will also be aware of his pervasive influence on ideas and modes of presentation throughout this book.

To Dr. B. J. Winer, now Professor of Psychology and Statistics at Purdue University and twice my former colleague, is owed a debt that I can scarcely hope to repay. In reviewing the first draft at the behest of the publisher and later versions of some sections in response to a

personal appeal, he went far beyond my expectations or hopes. Not confining himself to avoidance of lapses and some major errors, he also contributed innumerable positive suggestions. Any claim to erudition that this book may have, however modest, is largely attributable to Dr. Winer's painstaking and informed response to numerous entreaties.

I also want to express my thanks to Dr. Lyle V. Jones, currently a colleague at The University of North Carolina, who contributed helpful suggestions with respect to the treatment of statistical inference.

Another type of obligation, more indirect but nonetheless real, is felt toward my predecessors in the writing of textbooks intended for courses in statistics as applied to the behavioral sciences: the late Dr. L. L. Thurstone, whose *Fundamentals of Statistics* was the first book in the field that I read; several authors whose books I have assigned to classes in elementary statistics, including Dr. Helen M. Walker, Dr. J. P. Guilford, Dr. Allen Edwards, and Dr. George Ferguson; and Dr. Quinn McNemar, to whose work I have constantly referred.

Grateful acknowledgment is also made to an erstwhile friend who criticized the first draft of one chapter. The embryonic effort was thereafter dubbed the "Thus Book." The consequence was deletion of every "thus" I could see as well as many "now's," "here's," "then's," "hence's," "indeed's," "simply's," and "so's," and a few hundred miscellaneous words. The end product doubtless contains fewer "thuses" than other statistics books. Readers are invited to obliterate any that may have evaded me.

To Mrs. Robert Beaty, the secretary who readily became an expert at transcribing sometimes gruesome rough-draft copy, at typing statistical formulas and complicated tables, at drawing graphs on mimeograph stencils, and at uncomplainingly making what must have seemed endless and often unnecessary corrections—to her, too, I owe grateful thanks. In addition, appreciation is expressed to Mrs. Robert Boyce, who typed a sizeable portion of the original manuscript, and to Mrs. Carol Swaine, who performed many of the later typing chores.

Oliver and Boyd, Ltd., Edinburgh, graciously granted permission to adapt Table III and Table IV in Fisher, Ronald A., and Yates, Frank, *Statistical Tables for Biological, Agricultural and Medical Research*. To the publisher and the authors my debt is expressed.

DOROTHY C. ADKINS

Chapel Hill, North Carolina
February, 1964

Table of Contents

Chapter 6

Chapter 7

Chapter 8

Chapter 9

Chapter 10

Chapter 11

Chapter 12

Chapter 13

Chapter 14

Chapter 1

INTRODUCTION

NO ASSUMPTION OF PRIOR KNOWLEDGE OF STATISTICS

Most students of a behavioral science who venture into elementary statistics have had one or more courses entailing a casual introduction to it. They will have seen the need for some familiarity with statistical concepts and methods in order to understand the content of other courses, and they will have had varying amounts of instruction in statistics.

Since backgrounds differ and since some students have had little exposure to statistics or have been able to resist it successfully, this book does not assume any training in the field. Those who do retain knowledge acquired earlier should need to spend less time in studying this book than others who may remember little or who have eluded formal contact with statistics.

REQUIRED MATHEMATICAL BACKGROUND

You may have wondered what mathematical preparation you will be expected to have in order to use this book without undue discomfort. All college students, it should be safe to assume, have had a minimum

of the equivalent of a 1-semester course in high-school algebra. If the ability to understand and manipulate very elementary first-degree equations has been retained, you should be able to read this book with understanding, at least as far as any mathematical requirement is concerned.

Remedial Measures for Inadequate Preparation

If you have had no mathematics in college and have made little use of that learned in high school, or if you consider yourself very poor in quantitative ability or are immobilized by the mere sight of numbers and symbols, you will be wise to consult your instructor. He may suggest postponement of a statistics course until after you have studied elementary mathematics, or he may believe that a concurrent course would be beneficial.

Another possibility is to review early in the semester, perhaps with the help of such a book as Helen M. Walker's *Mathematics Essential for Elementary Statistics,* New York: Holt, Rinehart & Winston, Inc., 1934. If you are cursed with a disability that has strong emotional overtones, your instructor may go so far as to suggest enlisting the aid of a clinical psychologist or psychiatrist. Such an expedient would be indicated only rarely, however.

Try Chapter 2 in Any Case

You need not decide whether or not your mathematical background is grossly deficient, however, until after you have studied Chapter 2. The treatment of linear functions and their graphs contains much material that will not be new but that should serve to refresh your memory. If you are able to follow the presentation and the exercises in that chapter without an inordinate amount of difficulty, your present level of mathematical ability is adequate for the remainder of the book. Even if you experience some trouble with Chapter 2, you may find that a moderate amount of extra study will compensate for some deficiency in practiced algebraic skills.

WHY COURSES IN ELEMENTARY STATISTICS ARE GIVEN

Facilitation of Understanding of Literature

Students who have taken introductory courses in education or any of the behavioral sciences, as well as particular courses in many other

areas, will realize that some knowledge of statistics is germane to much of their reading. In order to understand fully descriptions of or allusions to research results in other courses, they find familiarity with statistical concepts and methods an asset. In this sense, statistics is a tool subject, to be applied in many different contexts.

Specific Preparation for Later Formal Education

Though certain other courses may not focus on statistical principles, rudimentary knowledge of them may be so essential that an introductory statistics course is prerequisite. This is often the case, for example, for rather elementary courses in educational or psychological measurement techniques, test construction and test theory, personnel administration, and experimental psychology. Training in the fundamentals of statistics is essential for certain advanced courses in testing and statistics as such, as well as in allied fields like psychophysics, attitude measurement, high-speed computer techniques, the design of experiments, and factor analysis, for example.

So relevant is statistics to the several different branches of psychology and sociology that many university graduate schools uniformly require at least a year of training in statistics beyond an introductory course, regardless of the particular area of specialization. This is true partly because statistics is a valuable tool subject, facilitating learning of the content of other courses, and partly because its mastery is essential for the conduct of many research enterprises. The graduate student who has adequately grasped certain statistical concepts and techniques is equipped to carry out the research that his educational program ordinarily includes. He is better able both to plan and to execute a research study and to analyze the data by statistical methods in case the design calls for such treatment. Moreover, he evidences fuller awareness of limits that should circumscribe interpretations of data, either his own or those of other investigators.

Statistics as a Cultural Subject

As the scope and depth of man's learning grow at a fantastically rapid rate and as research motivation and support reveal increasing acceleration, a modicum of statistical knowledge may come to be recognized as an indispensable ingredient of a liberal education. The college graduate who has added this achievement to his repertoire can more readily sift and absorb selected segments of the otherwise overwhelming array of knowledge that confronts him. From this point of view,

statistics should have high transfer value, adding significant leverage to an individual's educational endeavors. In other words, learning effort devoted to statistics should prove a sound investment that will continue to yield a regular return at the same time that it appreciates in value.

Enjoyment of Statistics for Its Own Sake

Doubtless the majority of teachers of statistics are habituated to the stereotype of their favorite subject as a legendary bugbear among students, some of whom seem to dare anyone to teach them anything about it. Yet all of us have encountered some students who approach it with gusto. A few read statistics books for relaxation, much as others bury themselves in mystery stories. These are among the persons whose earliest encounters with statistics were not so traumatic as to inflict relatively permanent mental disfigurement.

An Innovation That Should Further Enjoyment—Stemming from a conviction that almost all college students are capable of learning elementary statistics, an aim of this book has been to arrange the sequence of steps and gaps between them so as maximally to facilitate learning. At the same time, the textual material is interspersed with built-in testing devices. Through these the student can assure himself that he is making progress while he is engaged in study, as an alternative to anxious and uncertain anticipation of a post-mortem examination.

Although the major goal of this innovation in textbook writing has been to motivate learning, to the extent of its success it should add to the student's enjoyment of the study of statistics. The active reader who takes advantage of frequent opportunities to appraise his understanding is better able to diagnose weaknesses and adopt remedial measures. Being assured of mastery of what has gone before, he is in a position to approach new topics with confidence that he can cope with them, too, if he will but proceed at an optimal pace.

THE SUBJECT MATTER OF STATISTICS

Definitions

Statistics—An ordinary dictionary might define *statistics* in one sense as meaning a body of concepts and methods related to (or the science of) the collection, classification, and tabulation of facts or data that are expressed in numerical form or that have been quantified. In this

usage, the term "statistics" is construed as singular. In a plural interpretation, the same term also refers to the facts or data that have been collected, classified, and perhaps tabulated.

Statistic—The term *statistic* might be defined in the usual dictionary to mean a single bit of data or fact included in the collection of data called "statistics." The statistical lexicographers among us, however, ascribe a second meaning to the singular form, *statistic*. They use this word in an additional sense to refer to a summarizing index computed from several numerical observations or facts expressed in quantitative form. A familiar example of a statistic in this rather special sense is the *arithmetic average* or *mean* of a set of observations, which we see later to be their sum divided by their number.

Statistic Versus Parameter—The statistician goes further, to distinguish a *statistic*, based upon the particular data that may chance to be at hand, from a *parameter*. A parameter is the *true* value of the corresponding statistic in the larger assemblage of data from which the particular set can be considered to have been drawn. The statistic is computed from a *sample* of the larger *population* of observations which it is regarded as representing. Only by accident would the statistic be exactly equal to the parameter.

Statistics as a Means of Summarizing Data: Descriptive Statistics

This excursion into the lexicon of the statistician was made partly in the interests of defining his domain. He does, to be sure, assemble, classify, and tabulate numerical data. His primary concern is directed toward the group rather than the single individual instance or case. Once the data have been assembled and tabulated according to some useful categories, however, his interest by no means ceases. He typically proceeds to summarize the general trend of the data in terms of a measure of *average* or *central tendency* and the degree of individual differences among the observations or cases by an index of their *variability* or *dispersion*. The extent to which 1 characteristic of a group of individuals is associated with or *related* to other characteristics can be explored. If 2 variables are related, 1 can be used to *predict* the other; the goodness or precision of the prediction can be expressed by a statistical index.

The *differences* between different sets of data, with respect to general trends, variability, or relation to other variables, also can be explored. Knowledge of statistics assists the investigator in efficiently planning experiments designed to investigate matters of this kind and in learning how to analyze findings by suitable methods.