STATISTICAL METHODS FOR STUDENTS IN EDUCATION

BY

KARL J. HOLZINGER
ASSOCIATE PROFESSOR OF EDUCATION, THE UNIVERSITY OF CHICAGO

GINN AND COMPANY
BOSTON · NEW YORK · CHICAGO · LONDON
ATLANTA · DALLAS · COLUMBUS · SAN FRANCISCO

PRINTED IN THE UNITED STATES OF AMERICA

647.10

The Athenæum Press
GINN AND COMPANY • PRO-
PRIETORS • BOSTON • U.S.A.

PREFACE

This book is intended for students in education, who usually have had little training in mathematics. For those who have had considerable mathematics the theory of statistics is comparatively easy, but for students without such training the more advanced statistical methods offer many difficulties. The present volume supplements the mathematical preparation of the student by including sections on such topics as graphing, logarithms, and elementary theory of probability. The proofs of difficult theorems have been omitted throughout and demonstrations have been included only when experience has shown that they come within the grasp of the ordinary student and assist in a clear understanding of the method involved.

Although no attempt has been made to include all statistical methods now used in the field of education, the present text treats a somewhat larger number than will be found in most elementary books. The chief additions to the usual topics are the percentile method, application of the normal curve in correlating qualitative series, partial and multiple correlation, and elementary theory of curve fitting. The important subject of index numbers has been omitted entirely because a satisfactory treatment is beyond the scope of this book. The increasing need for index numbers in the field of school costs will probably lead to a separate volume on these methods.

In order to insure a clear understanding of the statistical arithmetic involved in the various methods presented, complete model problems have been worked out in the text. The experience of the writer has been that the ordinary student has considerable difficulty in formulating his plans for calculation and is greatly assisted by detailed arithmetical schemes

for computation, particularly in the early part of the course. A considerable number of exercises with answers have been added at the end of each chapter to clarify the methods discussed and to afford the student sufficient arithmetical practice to enable him to become accurate in his work. The amount of such practice needed varies greatly with students, and enough exercises are included to meet the needs of those requiring most drill.

The material in this volume will be found sufficient for an ordinary course of six months, but it may be condensed for a shorter course by the omission of certain topics and chapters. For an introductory course in a normal school or college, Chapters I to IX with selected topics from Chapters XII, XIII, and XIV are suggested. In case a second course is offered, the last seven or eight chapters with supplementary reading and term papers will usually be ample.

The writer is greatly indebted to Professor Karl Pearson, Dr. Leonard P. Ayres, and Professor Harold Rugg for ideas acquired while he was under their instruction. Valuable advice and suggestions have also been contributed by Dr. Egon Pearson, Professor C. H. Judd, Professor F. N. Freeman, Professor E. R. Breslich, Dr. Douglas Scates, and Dr. Ralph Hogan, all of whom read the manuscript while in preparation. Additional thanks are due to Mr. Lumir Brazda for preparation of the diagrams and to Mrs. Bryan Mitchell for assistance in checking the proof.

KARL J. HOLZINGER

CONTENTS

STATISTICAL METHODS FOR STUDENTS IN EDUCATION

CHAPTER I

INTRODUCTION

1. The Need for Statistical Method in Dealing with Educational Problems

In recent years the scientific movement in education has led to the wide use of quantitative methods. Problems in school administration and in educational theory and practice are now being studied chiefly by the application of experimental and statistical technique.

The increasing demand for school surveys and the generous appropriations made by the various foundations to promote these and other financial inquiries have created a need for statistical training for persons conducting such investigations. Some of the outstanding problems in such studies are the apportionment of school funds, school accounting, unit costs, and budgetary control, all of which involve careful accumulation of data and application of appropriate statistical method.

Another field in which adequate knowledge of statistics has become imperative is that of standardized tests. In modern educational science the old types of personal estimate and school examination are being replaced by intelligence tests and scales for the measuring of achievement in the various school subjects. Statistical methods are fundamental in the theory of test and scale construction and in the interpretation of the results obtained from such tests.

In the selection and organization of test material and the standardization and preparation in final form, elaborate technique is often required. Modern developments in test construction have led to the use of more and more refined methods, so that the test-maker of today needs to be a thorough student of statistics.

In the application of standardized tests to such problems as pupil classification, vocational guidance, diagnosis of special abilities, and evaluation of methods of instruction, a sound knowledge of statistical method is imperative, because all such studies involve the collection of appropriate data, summarization of the results, and correct inferences from the statistical findings.

The quantitative trend in school investigation has given rise to a tremendous bulk of literature. There are now hundreds of volumes on school surveys filled with tables and diagrams; there are books, monographs, theses, and reports likewise replete with statistics; there are scores of government, state, and institutional pamphlets; there are hundreds of standardized tests; and there is an ever-increasing amount of periodical literature reporting the findings of quantitative studies.

It is evident that if the school administrators and teachers for whom a large part of this great body of literature was written are to understand and apply it, they must have considerable familiarity with statistical method. It is impossible to keep up with the most recent developments in school research without some knowledge of the methods upon which such investigations are based.

Professional schools and departments in universities devoted to the training of teachers and administrators are meeting the demand by courses in experimental and statistical method. The purpose of such courses, in general, is to give the student sufficient information for intelligent reading of the present quantitative literature, and to furnish him with the technique necessary for carrying on his own investigations. This twofold aim has been kept in mind in preparing the present text.

2. Some General Requirements for Success in the Use of Statistical Method

In conducting a statistical study the investigator, survey expert, or classroom teacher should have in mind some definite problem or purpose, no matter how limited in scope. The mere gathering of masses of data or the haphazard calculation and plotting of diagrams are of little value unless they can be brought to bear upon a problem. While desirable lines of investigation are often discovered after the data have been collected and tabulated in a tentative way, it is much safer to decide upon the problem first and then proceed to collect the data necessary for its solution. The selection of a problem which is worth while, and which is sufficiently limited so that controls may be made and all necessary details carried out thoroughly and completely, is perhaps the most difficult part of the whole statistical procedure. It requires wide knowledge of the general field in which the problem lies, and a certain *constructive imagination* in foreseeing the various difficulties which are likely to arise.

Another requisite for a good statistical investigation is adequate data. No matter how excellent the problem or the plan of procedure, if the data employed are scanty the results will be of little value. Statistical method usually involves some generalization based upon summaries of the data. If the data are small in number, therefore, the conclusions drawn will not be reliable. This may be illustrated by some unpublished experiments in maze-learning based upon about twenty-five cases. Out of eight similar studies five showed a superiority for one method of learning, while the other three showed a difference in favor of another method. In all the experiments the number of cases was so small that none of the differences obtained proved to be significant, but could be readily accounted for by mere chance fluctuations in the samples of data chosen. While there is no fixed number of cases necessary for making

a statistical study, a desirable minimum for experimental work is about fifty, provided they are well chosen.

Data adequate as to number are not alone sufficient to insure satisfactory material. The facts gathered must be reliable and pertinent to the problem in hand. Questionnaire returns often fail in this respect because the intelligent replies of a number of persons to whom the blanks are sent are offset by careless or random answers on the part of others. Increasing the bulk of such data is not likely to increase its reliability, but the selecting of even a smaller number of persons who could be depended upon to give careful replies would yield better results. Thus if one wished to discover the most important aims in the teaching of high-school English, returns from a small well-selected group of experienced teachers would be preferable to those from a much larger group taken at random.

It frequently happens that the worker loses sight of the fact that his data are inadequate as to *quantity* and *quality* and applies elaborate statistical methods with the expectation that the final results will be of value. Such procedure, if followed intentionally, has been rightly described as "hiding behind a statistical smoke-screen," and is nothing less than a scientific crime. The limitations of the data employed should always be frankly recognized and the conclusions of the study made with them in mind. No amount of subsequent juggling by complicated formulas can give good results when they are based upon originally faulty data.

The successful statistician must have the capacity for careful, painstaking, and scientifically honest work. It is so easy to gather a few figures and tabulate them in such a way as to show a desired result or "prove" a certain theory that the temptations on the path of scientific rectitude are great. The untrained reader is often so bewildered by tables and diagrams that he is incapable of verifying the method or the inferences in a statistical article and either accepts the conclusions on the reputation of the writer or perhaps concludes that "anything can be proved

by statistics." Educational science would be greatly improved by the production of a smaller number of studies based upon better data and a more cautious use of statistical method.

A final requisite for the successful use of statistics is training in methodology. The investigator needs to become familiar with the various technical methods and processes of calculation. He needs much training in the application of these methods to data and problems in the particular field in which he expects to work. He also needs some knowledge of the difficult field of statistical inference. It is this general pedagogical requirement which the textbook and course in statistics are expected to fulfill. Such a course of study should familiarize the student with methods appropriate to educational problems, insure skill in statistical arithmetic, and provide opportunity for working out a worthwhile problem under careful guidance.

3. General Statistical Procedure in Dealing with a Problem

While there is no set order in which the steps in a statistical study must be carried out, experience has shown that a systematic procedure like the following is logical and economical of time and labor. Most of these steps will be discussed and fully illustrated in subsequent chapters.

(1) **Planning of the study.** When the student has some problem selected, his first concern will be with a rough plan for the whole study. It may not be possible to define the problem very specifically until the data have been gathered and examined, but the more definitely the limits of the inquiry can be set in advance the easier will be the subsequent steps. The usual mistake is to select a problem much too broad and too difficult for any one individual or even a small group of workers to undertake effectively. The availability, sources, accuracy, and methods of gathering data should all be considered in the preliminary plan.

(2) **Collection of the data.** With the problem defined and a general plan made, the next step is to collect the necessary data. This is accomplished by the use of questionnaires, by personal tabulation from data already available in records, or by the application of standardized tests, rating schemes, and other such measuring devices (Chapter II).

(3) **Preliminary analysis of the data.** If a questionnaire has been used in collecting the material, it is usually necessary to examine the returns very carefully before making tabulations. Incompleteness, inaccuracy, and ambiguity in the answers given should all be considered before the data are used. Similar analysis is often necessary with the results of standardized scales; unusual test conditions and errors in giving and in scoring the tests need to be checked up before tabulation is begun.

A preliminary analysis of the material will also be desirable in many cases to determine whether or not the data are adequate for the problem in hand. It may be that question-blank returns from a certain source are too scanty or fail to appear in such a form as to meet the requirements of the problem. In such cases a revised blank and more data will be required (Chapter II).

(4) **Tabulation for primary records.** After the preliminary analysis the data should be tabulated in such a way as to form both a permanent and a convenient working record. The permanent record may be kept in a bound volume with a page to each case or in the form of a master sheet with the names and records in parallel columns. The working record is usually in the form of small cards. One of these is made out for each case and the data entered in compact form so that the cards may be readily sorted and the resulting distributions easily checked (section 7, Chapter II).

(5) **Classification of the material.** Distributions, tables, and serial arrangements may next be made from the primary record.

These furnish the basis for calculations and graphical representations of the material.

(6) **Analysis of the classified data and planning of the calculations.** The particular statistical calculations to be employed are often not apparent until the data have been arranged in systematic form. The choice and right use of the proper analytical methods are extremely important, and at this point sound statistical judgment is required.

After the required calculations have been decided upon, they should be planned throughout before computation is begun. This is particularly advisable with data involving correlations (Chapters IX and X), the tables for which may be checked against one another and also used to furnish other statistical quantities such as the averages and measures of variability (Chapters VI and VII).

(7) **Calculation of the statistical constants.** The computations required may be made with the assistance of calculating tables and machines. It is desirable to have complete checks on the arithmetical accuracy of the work. Some of these are afforded by formulas, but the best check is to have two persons perform the calculations independently (Chapter V).

(8) **Interpretation of results.** The study has now reached the point where a careful scrutiny of results is required. These need to be interpreted in terms of the problem in hand. If the investigator is fortunate, the results may come out in such a way that the conclusions to be drawn are clear-cut and unambiguous. Very frequently, however, the findings are incomplete or inconclusive, so that it is necessary to make inferences with extreme caution. Careful application of the methods of statistical inference will then be necessary in order to guard against unwarranted generalizations (Chapter XIII).

(9) **Presentation of results in tables and diagrams.** Before writing the report most workers will find it desirable to prepare

rough sketches of the tables and diagrams to be used in the study. It is often convenient to cut these out and to pin them into the text as it is written (Chapter III).

(10) **Writing the report.** A satisfactory report will usually parallel in a general way the steps outlined above. It will contain a statement of the problem and its setting in the larger field; a description of the group studied; an account of the materials and methods employed; the results, inferences, and conclusions of the study; and a summary of the results obtained.

With this general plan in mind we may next turn to a detailed account of the various statistical methods.

CHAPTER II

COLLECTION AND CLASSIFICATION OF DATA

1. Primary and Secondary Data

The raw material employed in statistical studies consists in measurements or estimates known as *data*, which are *numerical statements of facts* in any department of inquiry, such as astronomy, economics, biology, psychology, and education. In the last field examples are furnished by the scores of pupils on standardized tests, physical measurements of children, salaries of teachers, attendance records, etc.

Data from whatever source may be described as *primary* or *secondary*. These terms are used in statistical method in much the same way as in historical research. In the latter field a fact taken from an ordinary text is considered as secondary material because it is removed at least one step from the original record. If the information were secured first-hand from documentary sources such as laws, original proceedings, letters, etc., it would be considered as primary historical data.

In the case of statistical method, primary data may be described as those secured from questionnaires, measurements, or estimates before the material has been combined or treated in any way so as to obscure the units or method of collection. Secondary data, on the other hand, are those which have already been collected and tabulated in some form available for use. They are usually removed one or more steps from the form of the original record, and hence comparison with similar material is of doubtful significance.

If the problem were to determine the academic training of teachers beyond four years of high school, primary records might consist of returns from a large sampling of individual