The Carus Mathematical Monographs

NUMBER THREE

MATHEMATICAL STATISTICS

By

HENRY LEWIS RIETZ

Professor of Mathematics, The University of Iowa

Published for

THE MATHEMATICAL ASSOCIATION OF AMERICA

by

THE OPEN COURT PUBLISHING COMPANY

CHICAGO · ILLINOIS

PREFACE

This book on mathematical statistics is the third of the series of Carus Mathematical Monographs. The purpose of the monographs, admirably expressed by Professor Bliss in the first book of the series, is "to make the essential features of various mathematical theories more accessible and attractive to as many persons as possible who have an interest in mathematics but who may not be specialists in the particular theory presented."

The problem of making statistical theory available has been changed considerably during the past two or three years by the appearance of a large number of textbooks on statistical methods. In the course of preparation of the manuscript of the present volume, the writer felt at one time that perhaps the recent books had covered the ground in such a way as to accomplish the main purposes of the monograph which was in process of preparation. But further consideration gave support to the view that although the recent books on statistical method will serve useful purposes in the teaching and standardization of statistical practice, they have not, in general, gone far toward exposing the nature of the underlying theory, and some of them may even give misleading impressions as to the place and importance of probability theory in statistical analysis.

It thus appears that an exposition of certain essential features of the theory involved in statistical analysis would conform to the purposes of the Carus Mathematical Monographs, particularly if the exposition could be made interesting to the general mathematical reader. It is not the intention in the above remarks to imply a criticism of the books in question. These books serve certain useful purposes. In them the emphasis has been very properly placed on the use of devices which facilitate the description and analysis of data.

The present monograph will accomplish its main purpose if it makes a slight contribution toward shifting the emphasis and point of view in the study of statistics in the direction of the consideration of the underlying theory involved in certain highly important methods of statistical analysis, and if it introduces some of the recent advances in mathematical statistics to a wider range of readers. With this as our main purpose it is natural that no great effort is being made to present a well-balanced discussion of all the many available topics. This will be fairly obvious from omissions which will be noted in the following pages. For example, the very important elementary methods of description and analysis of data by purely graphic methods and by the use of various kinds of averages and measures of dispersion are for the most part omitted owing to the fact that these methods are so available in recent elementary books that it seems unnecessary to deal with them in this monograph. On the other hand, topics which suggest making the underlying theories more available are emphasized.

For the purpose of reaching a relatively large number of readers, we are fortunate in that considerable portions of the present monograph can be read by those who have relatively little knowledge of college mathematics. However, the exposition is designed, in general, for readers of a certain degree of mathematical maturity, and presupposes an acquaintance with elementary differential and integral calculus, and with the elementary principles of probability as presented in various books on college algebra for freshmen.

A brief list of references is given at the end of Chapter VII. This is not a bibliography but simply includes books and papers to which attention has been directed in the course of the text by the use of superscripts.

The author desires to express his special indebtedness to Professor Burton H. Camp who read critically the entire manuscript and made many valuable suggestions that resulted in improvements. The author is also indebted to Professor A. R. Crathorne for suggestions on Chapter I and to Professor E. W. Chittenden for certain suggestions on Chapters II and III. Lastly, the author is deeply indebted to Professor Bliss and to Professor Curtiss of the Publication Committee for important criticisms and suggestions, many of which were made with special reference to the purposes of the Carus Mathematical Monographs.

HENRY L. RIETZ

THE UNIVERSITY OF IOWA
December, 1926

TABLE OF CONTENTS

CHAPTER I

THE NATURE OF THE PROBLEMS AND UNDERLYING CONCEPTS OF MATHEMATICAL STATISTICS

1. **The scope of mathematical statistics.** The bounds of mathematical statistics are not sharply defined. It is not uncommon to include under mathematical statistics such topics as interpolation theory, approximate integration, periodogram analysis, index numbers, actuarial theory, and various other topics from the calculus of observations. In fact, it seems that mathematical statistics in its most extended meaning may be regarded as including all the mathematics applied to the analysis of quantitative data obtained from observation. On the other hand, a number of mathematicians and statisticians have implied by their writings a limitation of mathematical statistics to the consideration of such questions of frequency, probability, averages, mathematical expectation, and dispersion as are likely to arise in the characterization and analysis of masses of quantitative data. Borel has expressed this somewhat restricted point of view in his statement[1] that the general problem of mathematical statistics is to determine a system of drawings carried out with urns of fixed composition, in such a way that the results of a series of drawings lead, with a very high degree of probability, to a table of values identical with the table of observed values.

[1] For footnote references, see pp. 173–77.

On account of the different views concerning the boundaries of the field of mathematical statistics there arose early in the preparation of this monograph questions of some difficulty in the selection of topics to be included. Although no attempt will be made here to answer the question as to the appropriate boundaries of the field for all purposes, nevertheless it will be convenient, partly because of limitations of space, to adopt a somewhat restricted view with respect to the topics to be included. To be more specific, the exposition of mathematical statistics here given will be limited to certain methods and theories which, in their inception, center around the names of Bernoulli, De Moivre, Laplace, Lexis, Tchebycheff, Gram, Pearson, Edgeworth, and Charlier, and which have been much developed by other contributors. These methods and theories are much concerned with such concepts as frequency, probability, averages, mathematical expectation, dispersion, and correlation.

2. **Historical remarks.** While we are currently experiencing a period of special activity in mathematical statistics which dates back only about forty years, some of the concepts of mathematical statistics are by no means of recent origin. The word "statistics" is itself a comparatively new word as shown by the fact that its first occurrence in English thus far noted seems to have been in J. F. von Bielfeld, *The Elements of Universal Erudition*, translated by W. Hooper, London, 1770. Notwithstanding the comparatively recent introduction of the word, certain fundamental concepts of mathematical statistics to which attention is directed in this monograph date back to the first publication relating to Bernoulli's theorem in 1713. The line of development started by Bernoulli was carried

forward by Stirling (1730), De Moivre (1733), Euler (1738), and Maclaurin (1742), and culminated in the formulation of the probability theory of Laplace. The *Théorie Analytique des Probabilités* of Laplace published in 1812 is the most significant publication underlying mathematical statistics. For a period of approximately fifty years following the publication of this monumental work there was relatively little of importance contributed to the subject. While we should not overlook Poisson's extension of the Bernoulli theory to cases where the probability is not constant, Gauss's development of methods for the adjustment of observations, Bravais's extension of the normal law to functions of two and three variables, Quetelet's activities as a popularizer of social statistics, nevertheless there was on the whole in this period of fifty years little progress.

The lack of progress in this period may be attributed to at least three factors: (1) Laplace left many of his results in the form of approximations that would not readily form the basis for further development; (2) the followers of Gauss retarded progress in the generalization of frequency theory by overpromoting the idea that deviations from the normal law of frequency are due to lack of data; (3) Quetelet overpopularized the idea of the stability of certain striking forms of social statistics, for example, the stability of the number of suicides per year, with the natural result that his activities cast upon statistics a suspicion of quackery which exists even to some extent at present.

An important step in advance was taken in 1877 in the publication of the contributions of Lexis to the classification of statistical distributions with respect to normal, supernormal, and subnormal dispersion. This theory will receive attention in the present monograph.

The development of generalized frequency curves and the contributions to a theory of correlation from 1885 to 1900 started the period of activity in mathematical statistics in which we find ourselves at present. The present monograph deals largely with the progress in this period, and with the earlier underlying theory which facilitated relatively recent progress.

3. **Two general types of problems.** For purposes of description it seems convenient to recognize two general classes of problems with which we are concerned in mathematical statistics. In the problems of the first class our concern is largely with the characterization of a set of numerical measurements or estimates of some attribute or attributes of a given set of individuals. For example, we may establish the facts about the heights of 1,000 men by finding averages, measures of dispersion, and various statistical indexes. Our problem may be limited to a characterization of the heights of these 1,000 men.

In the problems of the second class we regard the data obtained from observation and measurement as a random sample drawn from a well-defined class of items which may include either a limited or an unlimited supply. Such a well-defined class of items may be called the "population" or universe of discourse. We are in this case concerned with using the properties of a random sample of variates for the purpose of drawing inferences about the larger population from which the sample was drawn. For example, in this class of problems involving the heights of the 1,000 men we would be concerned with the question: What approximate or probable inferences may be drawn about the statures of a whole race of men from an analysis of the heights of a sample of 1,000 men drawn at random from the men of the race? In dealing with such questions, we should in the first place consider the difficulties involved in drawing a sample that is truly random, and in the next place the problem of developing certain parts of the theory of probability involved in statistical inference.

The two classes of problems to which we have directed attention are not, however, entirely distinct with regard to their treatment. For example, the conceptions of probable and standard error may be used both in describing the facts about a sample and in indicating the probable degree of precision of inferences which go beyond the observed sample by dealing with certain properties of the population from which we conceive the sample to be drawn. Moreover, a satisfactory description of a sample is not likely to be so purely descriptive as wholly to prevent the mind from dwelling on the inner meaning of the facts in relation to the population from which the sample is drawn.

As a preliminary to dealing in later chapters with certain of the problems falling under these two general classes we shall attempt in the present chapter to discuss briefly the nature of certain underlying concepts. We shall find it convenient to consider these concepts in pairs as follows: relative frequency and probability; observed and theoretical frequency distributions; arithmetic mean and mathematical expectation; mode and most probable value; moments and mathematical expectations of a power of a variable.

4. Relative frequency and probability. The frequency f of the occurrence of a character or event among s possible occurrences is one of the simplest items of statistical information. For example, any one of the following items illustrates such statistical information: Five deaths in a year among 1,000 persons aged 30, nearest birthday; 610 boys among the last 1,200 children born in a city; 400 married men out of a total of 1,000 men of age 23; twelve cases of 7 heads in throwing 7 coins 1,536 times.

The determination of the numerical values of the relative frequencies f/s corresponding to such items is one of the simplest problems of statistics. This simple problem suggests a fundamental problem concerning the probable or expected values of such relative frequencies if s were a very large number. When s is a large number, the relative frequency f/s is very commonly accepted in applied statistics as an approximate measure of the probability of occurrence of the event or character on a given occasion.

To take an illustration from an important statistical problem, let us assume that among l persons equally likely to live a year we find d observed deaths during the year. That is, we assume that d represents the frequency of deaths per year among the l persons each exposed for one year to the hazards of death. If l is fairly large, the relative frequency d/l is often regarded as an approximation to what is to be defined as the probability of death of one such person within a year. In fact, it is a fundamental assumption of actuarial science that we may regard such a relative frequency as an approximation to the probability of death when a sufficiently large number of persons are exposed to the hazards of death. For a numerical illustration, suppose there are 600 deaths among 100,000 persons exposed for a year at age 30. We accept .006 as an approximation to the probability in question at age 30. In the method of finding such an approximation we decide on a population which constitutes an appropriate class for investigation and in which individuals satisfy certain conditions as to likeness. Then we depend on observation to obtain the items which lead to the relative frequency which we may regard as an approximation to the probability.

For an ideal population, let us conceive an urn containing white and black balls alike except as to color and thoroughly mixed. Suppose further for the present that we do not know the ratio of the number of white balls to the total number in this urn which we may conceive to contain either any finite number or an indefinitely large number of balls. This ratio is often called the probability of drawing a white ball. When the number in the urn is finite, we make drawings at random consisting of s balls taken one at a time with replacements to keep the ratio of the numbers of white and black balls constant. If we may assume the number in the urn to be infinite, the drawings may under certain conditions be made without replacements. Suppose we obtain f white balls as a result of thus drawing s balls, then we say that f/s is the relative frequency with which we drew white balls. When s is large, this relative frequency would ordinarily give us an approximate value of the probability of drawing a white ball in one trial, that is, an approximate value of the ratio of white balls to the total number of balls in the urn.

Thus far we have not defined probability, but have presented illustrations of approximations to probabilities. While these illustrations seem to suggest a definition, it is nevertheless difficult to frame a definition that is satisfactory and includes all forms of probability. The need for the concepts of relative frequency and probability in statistics arises when we are associating two events such that the first may be regarded as a *trial* and the second may be regarded as a *success* or a *failure* depending on the result of the trial. The relative frequency of success is then the ratio of the number of successes to the total number of trials.

If the relative frequency of success approaches a limit when the trial is repeated indefinitely under the same set of circumstances, this limit is called the probability of success in one trial.

There are some objections to this definition of probability as well as to any other that we could propose. One objection is concerned with questioning the validity of the assumption that a limit of the relative frequency exists, and another relates to the meaning of the expression, "the same set of circumstances." That the limit exists is an empirical assumption whose validity cannot be proved, but experience with data in many fields has given much support to the reasonableness and usefulness of the assumption. The objection based on the difficulty of controlling conditions so as to repeat the trial under the same set of circumstances is an objection that could be brought against experimental science in general with respect to the difficulties of repeating experiments under the same circumstances. The experiments are repeated as nearly as circumstance permits.

It seems fairly obvious that the development of statistical concepts is approached more naturally from this limit definition than from the familiar definitions suggested by games of chance. However, we shall at certain points in our treatment (for example, see § 11) give attention to the fact that various definitions of probability exist in which the assumptions differ from those involved in the above definition. The meaning of probability in statistics is fairly well expressed for some purposes by any one of the expressions, *theoretical relative frequency*, *presumptive relative frequency*, or *expected value* of a relative frequency. Indeed, we sometimes express the fact that the relative frequency f/s is assumed to have the probability p as a limit when $s\rightarrow\infty$ in abbreviated form by writing $E(f/s)=p$, where $E(f/s)$ is read, "expected value of f/s." It is fairly clear that in our definition of probability we simply idealize actual experience by assuming the existence of a limit of the relative frequency. This idealization, for purposes of definition, is in some respects analogous to the idealization of the chalk mark into the straight line of geometry.

In certain cases, notably in games of chance or urn schemata, the probability may be obtained without collecting statistical data on frequencies. Such cases arise when we have urn schemata of which we know the ratio of the number of white balls to the total number. For example, suppose an urn contains 7 white and 3 black balls and that we are to inquire into the probability that a ball to be drawn will be white. We could experiment by drawing one ball at a time with replacements until we had made a very large number of drawings and then estimate the probability from the ratio of the number of white balls to the total number of balls drawn. It would however in this case ordinarily be much more convenient and satisfying to examine the balls to note that they are alike except as to color and then make certain assumptions that would give us the probability without actually making the trials.

Thus, when all the possible ways of drawing the balls one at a time may be analyzed into 10 equally likely ways, and when 7 of these 10 ways give white balls, we assume that 7/10 is the probability that the ball to be drawn in one trial will be white. This simple case illustrates the following process of arriving at a probability:

If all of an aggregate of ways of obtaining successes and failures can be analyzed into s' *possible mutually exclusive ways each of which is equally likely; and if* f' *of these ways give successes, the probability of a success in a single trial may be taken to be* f'/s'.

Thus in throwing a single die, what is the probability of obtaining an ace? We assume that there are 6 equally likely ways in which the die may fall. One of these ways gives an ace. Hence, we say 1/6 is the probability of throwing an ace. A probability whose value is thus obtained from an analysis of ways of occurrence into sets of equally likely cases and a segregation of the cases in which a success would occur is sometimes called an *a priori probability*, while a probability whose approximate value is obtained from actual statistical data on repeated trials is called an *a posteriori* or *statistical probability*.

In making an analysis to study probabilities, difficult questions arise both as to the meaning and fulfilment of the condition that the ways are to be "equally likely." These questions have been the subject of lively debates