

Preface

The purpose of this book is to present up-to-date theory and techniques of statistical inference in a logically integrated and practical form. Essentially, it incorporates the important developments in the subject that have taken place in the last three decades. It is written for readers with a background knowledge of mathematics and statistics at the undergraduate level.

Quantitative inference, if it were to retain its scientific character, could not be divested of its logical, mathematical, and probabilistic aspects. The main approach to statistical inference is inductive reasoning, by which we arrive at "statements of uncertainty." The rigorous expression that degrees of uncertainty require are furnished by the mathematical methods and probability concepts which form the foundations of modern statistical theory. It was my awareness that advanced mathematical methods and probability theory are indispensable accompaniments in a self-contained treatment of statistical inference that prompted me to devote the first chapter of this book to a detailed discussion of vector spaces and matrix methods and the second chapter to a measure-theoretic exposition of probability and development of probability tools and techniques.

Statistical inference techniques, if not applied to the real world, will lose their import and appear to be deductive exercises. Furthermore, it is my belief that in a statistical course emphasis should be given to both mathematical theory of statistics and to the application of the theory to practical problems. A detailed discussion on the application of a statistical technique facilitates better understanding of the theory behind the technique. To this end, in this book, live examples have been interwoven with mathematical results. In addition, a large number of problems are given at the end of each chapter. Some are intended to complement main results derived in the body of the chapter, whereas others are meant to serve as exercises for the reader to test his understanding of theory.

The selection and presentation of material to cover the wide field of

statistical inference have not been easy. I have been guided by my own experience in teaching undergraduate and graduate students, and in conducting and guiding research in statistics during the last twenty years. I have selected and presented the essential tools of statistics and discussed in detail their theoretical bases to enable the readers to equip themselves for consultation work or for pursuing specialized studies and research in statistics.

Why Chapter 1 provides a rather lengthy treatment of the algebra of vectors and matrices needs some explanation. First, the mathematical treatment of statistical techniques in subsequent chapters depends heavily on vector spaces and matrix methods; and second, vector and matrix algebra constitute a branch of mathematics widely used in modern treatises on natural, biological, and social sciences. The subject matter of the chapter is given a logical and rigorous treatment and is developed gradually to an advanced level. All the important theorems and derived results are presented in a form readily adaptable for use by research workers in different branches of science.

Chapter 2 contains a systematic development of the probability tools and techniques needed for dealing with statistical inference. Starting with the axioms of probability, the chapter proceeds to formulate the concepts of a random variable, distribution function, and conditional expectation and distributions. These are followed by a study of characteristic functions, probability distributions in infinite dimensional product spaces, and all the important limit theorems. Chapter 2 also provides numerous propositions, which find frequent use in some of the other chapters and also serve as good equipment for those who want to specialize in advanced probability theory.

Chapter 3 deals with continuous probability models and the sampling distributions needed for statistical inference. Some of the important distributions frequently used in practice, such as the normal, Gamma, Cauchy, and other distributions, are introduced through appropriate probability models on physical mechanisms generating the observations. A special feature of this chapter is a discussion of problems in statistical mechanics relating to the equilibrium distribution of particles.

Chapter 4 is devoted to inference through the technique of analysis of variance. The Gauss-Markoff linear model and the associated problems of estimation and testing are treated in their wide generality. The problem of variance-components is considered as a special case of the more general problem of estimating intraclass correlation coefficients. A unified treatment is provided of multiclassified data under different sampling schemes for classes within categories.

The different theories and methods of estimation form the subject matter of Chapter 5. Some of the controversies on the topic of estimation are examined; and to remove some of the existing inconsistencies, certain modifications are introduced in the criteria of estimation in large samples.

Problems of specification, and associated tests of homogeneity of parallel

samples and estimates, are dealt with in Chapter 6. The choice of a mathematical model from which the observations could be deemed to have arisen is of fundamental importance because subsequent statistical computations will be made on the framework of the chosen model. Appropriate tests have been developed to check the adequacy of proposed models on the basis of available facts.

Chapter 7 provides the theoretical background for the different aspects of statistical inference, such as testing of hypotheses, interval estimation, experimentation, the problem of identification, nonparametric inference, and so on.

Chapter 8, the last chapter, is concerned with inference from multivariate data. A special feature of this chapter is a study of the multivariate normal distribution through a simple characterization, instead of through the density function. The characterization simplifies the multivariate theory and enables suitable generalizations to be made from the univariate theory without further analysis. It also provides the necessary background for studying multivariate normal distributions in more general situations, such as distributions on Hilbert space.

Certain notations have been used throughout the book to indicate sections and other references. The following examples will help in their interpretation. A subsection such as 4f.3 means subsection 3 in section f of Chapter 4. Equation (4f.3.6) is the equation numbered 6 in subsection 4f.3 and Table 4f.3 β is the table numbered second in subsection 4f.3. The main propositions (or theorems) in each subsection are numbered: (i), (ii), etc. A back reference such as [(iii). 5d.2] indicates proposition (iii) in subsection 5d.2.

A substantial part of the book was written while I was a visiting professor at the Johns Hopkins University, Baltimore, in 1963-1964, under a Senior Scientist Fellowship scheme of the National Science Foundation, U.S.A. At the Johns Hopkins University, I had the constant advice of G. S. Watson, Professor of Statistics, who read the manuscript at the various stages of its preparation. Comments by Herman Chernoff on Chapters 7 and 8, by Rupert Miller and S. W. Dharmadhikari on Chapter 2, and by Ralph Bradley on Chapters 1 and 3, have been extremely helpful in the preparation of the final manuscript. I wish to express my thanks to all of them. The preparation and revision of the manuscript would not have been an easy task without the help of G. M. Das, who undertook the heavy burden of typing and organizing the manuscript for the press with great care and diligence.

Finally, I wish to express my gratitude to the late Sir Ronald A. Fisher and to Professor P. C. Mahalanobis under whose influence I have come to appreciate statistics as the new technology of the present century.