Mathematical Statistics

JOHN E. FREUND

JOHN E. FREUND

Professor of Mathematics
Arizona State University

MATHEMATICAL STATISTICS

PRENTICE-HALL, INC. Englewood Cliffs, N. J.

First printing............January, 1962
Second printing..........January, 1963

Library of Congress Catalog Card Number 62–9287

Printed in the United States of America

56219–C

Preface

This book has been written for an introductory two-semester or three-quarter course in mathematical statistics with the prerequisite of a standard undergraduate course in calculus. The first seven chapters, dealing with an introduction to probability, basic distribution theory, and some limit theorems, can also serve for a semester course in probability theory. The treatment of probability is rigorous in so far as discrete (finite or countably infinite) sample spaces are concerned. Some of the difficulties arising in the continuous case are pointed out in Section 5.1; it is felt that a rigorous treatment of general probability spaces is better left for a more advanced course.

One of the major problems in writing a text on mathematical statistics is to find a suitable balance between theory and application. Ultimately, of course, such a balance can only reflect the author's personal preference, as is evidenced by existing texts ranging from the purely theoretical to the largely applied. Although there is emphasis in this text on the mathematics of statistics, it is hoped that it will, nevertheless, aid the reader in developing an early appreciation for applications. It is for this reason that the author has included special sets of applied problems in Sections 3.2.4 and 5.3.5 and the material on decision making in Sections 4.6 and 6.6. Although the treatment of statistical inference is fairly traditional, an introduction to the fundamental concepts of decision theory and some simple illustrations are given in Section 9.1.

Since the language of statistics is not always acceptable to pure mathematicians, the author had to make some compromises while attempting to write this book in the language of modern mathematics. So far as symbolism is concerned, boldface type is used for random variables in order to distinguish between functions and the values which they assume.

The author would like to express his appreciation for the many helpful suggestions which he received from his students, colleagues, and friends. In particular, the author is indebted to Dr. Irwin Miller, who taught with a preliminary draft of this book. The author would also like to express his appreciation to the editorial staff of Prentice-Hall, Inc., for their courteous cooperation in the production of this book and, above all, to his wife for

putting up with the many inconveniences which are unavoidable while one is engaged in writing a book.

Finally, the author would like to express his appreciation and indebtedness to the McGraw-Hill Book Company for permission to reproduce the material in Tables I and II from their *Handbook of Probability and Statistics with Tables*; to Professor R. A. Fisher and Messrs. Oliver and Boyd, Ltd., Edinburg, for permission to reproduce the material in Table IV from their book *Statistical Methods for Research Workers*; and to Professor E. S. Pearson and the *Biometrika* trustees for permission to reproduce the material in Tables V and VI.

JOHN E. FREUND

Contents

Chapter 1

Introduction

1.1 Introduction

In recent years, the growth of statistics has made itself felt in practically every phase of human activity. Statistics no longer consists merely of the collection of data and their presentation in charts and tables, it is now considered to encompass the entire science of decision making in the face of uncertainty. This covers enormous ground since uncertainties are met when we flip a coin, when we experiment with a new drug, when we determine life insurance premiums, when we inspect manufactured products, when we compare the merits of different missiles, rate the abilities of human beings, make executive decisions, and so forth. It would be presumptuous to say that statistics, in its present state of development, can handle *all* situations involving uncertainties, but new methods are being developed constantly and modern mathematical statistics can, at least, provide the framework for looking at these situations in a logical and systematic fashion. In other words, probability and statistics provide the *models*, the underlying mathematical formulations, to study situations involving uncertainties just as calculus provides the *models*, the underlying mathematical formulations, to describe the concepts of Newtonian physics.

Historically speaking, the origin of probability theory dates back to the 17th century. It seems that the Chevalier de Méré, claimed to have been an ardent gambler, was baffled by some questions concerning a game of chance.* He consulted the French mathematician Blaise Pascal (1623–1662), who in turn wrote about this matter to Pierre Fermat (1601–1665); it is this correspondence which is generally considered the origin of modern probability theory.

The 18th century saw a rapid growth of the mathematics of probability as it applies to games of chance, but it was not until the work of Karl Gauss (1777–1855) and Pierre Laplace (1749–1827) that this theory found applications in other fields. Noting that the theory developed for "heads and tails" or "red and black" in games of chance applied also to situations where the outcomes are "life or death" or "boy or girl", probability theory was applied to actuarial mathematics and to some phases of social science. Later, statistical concepts were introduced into physics by L. Boltzmann, J. Gibbs, and J. Maxwell, and in this century the methods of probability and statistics have found applications in all phases of human endeavor which in some way involve an element of uncertainty or risk. The names which are connected most prominently with the growth of 20th century mathematical statistics are those of R. A. Fisher, J. Neyman, E. S. Pearson, and A. Wald. References to the particular contributions made by these statisticians will be made later in the text.

Since the subjects of probability and statistics can be presented at various levels of mathematical refinement and with various patterns of emphasis, let us point out briefly that the mathematical background expected of the reader is a basic course in differential and integral calculus, including some elementary material on partial differentiation, multiple integration, and series. So far as emphasis is concerned, Chapters 2 through 7 are devoted primarily to the mathematical concepts and techniques which are required to develop the statistical methods treated in Chapters 8 through 14. Chapters 2 and 5 provide a formal introduction to probability based on the concepts of sets, while the other chapters of the first half of the book deal with what might be called basic *distribution theory*. Chapters 8 through 14 contain an introduction to the most fundamental and the most widely used methods of statistics, with emphasis on their theoretical foundation.

* Essentially, his question was how to divide the stakes if two players start but fail to complete a game, in which the winner is the one who wins three matches out of five.

1.2 Fundamental Problems of Probability

Directly or indirectly, probability plays a role in all problems of science, business, and everyday life, which somehow involve an element of uncertainty. In view of this, it is unfortunate that the term "probability", itself, is difficult to define and controversial.

In the study of probability there are essentially three kinds of problems. First there is the question of what we *mean* when we say that a probability is 0.82, 0.25, and so forth; then there is the question of *how to obtain* numerical values of probabilities; and finally there is the question of how known or assumed values of probabilities can be used to determine others, namely, the formal *calculus or probability*. Most of the early chapters of this book will be devoted to the last kind of problem; the problem of how to obtain probabilities will be touched upon in Chapters 9 and 10, dealing with the general problem of estimation.

Philosophical arguments about the various meanings which have been attached to the word "probability" make interesting reading, but in view of our interests and objectives in this book, we shall limit our discussion to the so-called *objectivistic* view. Accordingly, we shall *interpret* probabilities in terms of relative frequencies, or to be more exact as *limits of relative frequencies*. When we say "the probability that a man aged 50 will live to be 65 is 0.72", we mean that if present conditions prevail, 72 per cent of all men aged 50 will live to be 65; when we say "the probability that it will rain tomorrow in Detroit is 0.27", we mean that in the long run it will rain there on that date 27 per cent of the time. The proportion of the time that an event takes place is called its relative frequency, and the relative frequency with which it takes place in the long run is interpreted as its probability.

When we say "the probability of getting heads with a balanced coin is 0.50", this means that in the long run we will get 50 per cent heads and 50 per cent tails; it does not mean that we must necessarily get 5 heads and 5 tails in 10 flips or 50 heads and 50 tails in 100.

It is important to note that the mathematical theory of probability and statistics, which is the subject matter of this book, does not depend on philosophical arguments concerning the meaning of "probability"; there is general agreement concerning the postulates and definitions given in Chapters 2 and 5. Questions of meaning arise only when mathematical theories are applied, and *when it comes to applications, the frequency interpretation of probability is the one that is held by many, probably most, statisticians.*

1.3 Probabilities and Sets

Probabilities invariably refer to the occurrence or non-occurrence of some event. We assign a probability to the *event* that a coin will come up heads, to the *event* that a given candidate will be elected governor of his state, to the *event* that it will not rain on the day of the company's picnic, to the *event* that a missile will hit its target, and so forth. To treat the subject of probability in a rigorous fashion, it will thus be necessary to explain what we mean by "event", and we shall do so by representing events with *sets*, usually *sets of points.*

Figure 1.1

For example, if we consider the flip of a coin, the outcomes, the *events* of getting head or tail, can be represented by the two points of Figure 1.1, where 0 stands for tail (0 heads) and 1 for head (1 head). Similarly, the outcomes of the roll of a die can be represented by the six points of Figure 1.2.

Figure 1.2

To consider a slightly more complicated example, the outcomes of an experiment consisting of two flips of a coin can be represented

Figure 1.3

by the points of Figure 1.3, provided one is interested only in the total number of heads. An alternate way of representing the outcomes by means of points is shown in Figure 1.4, where 0 and 1 again stand for tail and head, and the two coordinates represent the two flips of the coin. It should be noted that in Figure 1.4 the event of getting 1 head and 1 tail is represented by the set of two points inside the dotted line.

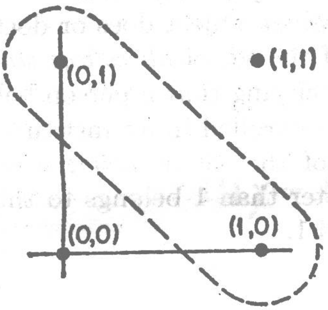

Figure 1.4

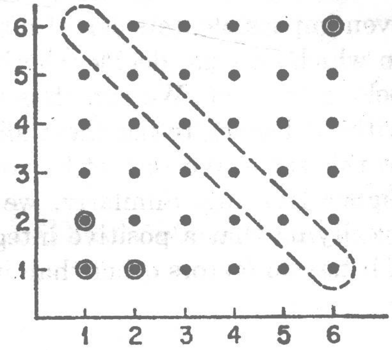

Figure 1.5

Referring to an experiment which consists of rolling a pair of dice (one red and one green), the event of rolling a 7 is represented by the set of six points inside the dotted line of Figure 1.5, and the event of rolling 2, 3, or 12 is represented by the set of four points which are circled in Figure 1.5.

Since sets play such an important role in the theory of probability, some of the basic concepts of set theory will be introduced in Chapter 2. Although the proofs of many theorems in probability theory require knowledge of the *Algebra of Sets*, it will be seen that they can also be justified quite readily by means of simple diagrams. Knowledge of the Algebra of Sets (or *Boolean Algebra* as it is also called) is, thus, desirable, but not absolutely essential.

Throughout this section we used the term "set" without giving it a definition, understanding tacitly that it stands for a collection, group, or class of points or other kinds of objects. Indeed, it is customary to leave this term undefined, subject to the qualification that sets must obey the rules set down in Section 2.1.3.

The objects that belong to a set are usually referred to as its *members* or its *elements*, and sets are sometimes specified by actually listing the individual elements. Thus, the set which consists of the different outcomes of a roll of a die may be indicated as {1, 2, 3, 4, 5, 6}; and similarly {apple pie, ice cream, chocolate cake, rice pudding} is a set of four desserts. It should be noted that when the elements of a set are thus listed, their order does not matter; $\{1, 2, 3\}$ and $\{2, 3, 1\}$ represent the same set consisting of the first three positive integers.

Instead of listing the elements, which is often impracticable or even impossible, sets can also be specified by giving a rule according to which one can decide whether any given object does or does not belong to a set. We can thus speak of the set of *all college students* without having to list them all, by specifying that a person belongs to this set if and only if he (or she) is enrolled in an institution of higher learning. Similarly, we speak of the set of *prime numbers*, specifying that a positive integer greater than 1 belongs to this set if it has no factors other than itself and 1.

Chapter 2

Probability—The Discrete Case

2.1 Discrete Sample Spaces

A set whose elements represent all possible outcomes of an experiment is generally called a *sample space* for the experiment and it will be denoted by the letter S. Using the word "experiment" rather loosely, we are referring here to any situation which permits a variety of outcomes involving somehow an element of chance. Instead of "sample space", the terms "universal set", "universe of discourse", and "possibilities space" are also used.

The fact that the sample space which represents a given experiment need not be unique is illustrated by Figures 1.3 and 1.4; both of these sample spaces represent the outcomes of an experiment consisting of two flips of a coin, the difference being in what we mean by "outcome". *Generally speaking, it is desirable to use sample spaces whose elements represent outcomes which do not permit further subdivision; that is, an individual element of the sample space should not represent two or more outcomes which are distinguishable in some fashion.* This is true for the sample space of Figure 1.4, but not for that of Figure 1.3. Following this convention avoids many of the difficulties and paradoxes which hampered the early development of the theory of probability.

Sample spaces are usually classified according to the *number of elements* which they contain and also according to the *dimension* of the geometrical configuration in which they are displayed. Thus, Figure 1.1 is a one-dimensional representation of the outcomes of a