

# Mathematical Statistics

JOHN E. FREUND

212  
E9022

3 22214

外文书库

JOHN E. FREUND

*Professor of Mathematics  
Arizona State University*



**MATHEMATICAL  
STATISTICS**



暨南大學  
數學系資料室

一九六四年 五月 六日

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

© 1962 by  
PRENTICE-HALL, INC.  
Englewood Cliffs, N. J.

All rights reserved. No part of this book  
may be reproduced in any form, by  
mimeograph or any other means, without  
permission in writing from the publisher.

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

<b>Chapter 1</b>	<b>INTRODUCTION</b>	<b>1</b>
1.1	Introduction	1
1.2	Fundamental problems of probability	3
1.3	Probabilities and sets	4
<b>Chapter 2</b>	<b>PROBABILITY—THE DISCRETE CASE</b>	<b>7</b>
2.1	Discrete sample spaces	7
2.1.1	<i>Subsets and events</i>	9
2.1.2	<i>Operations on sets</i>	10
2.1.3	<i>The algebra of sets</i>	13
2.2	Some combinatorial theory	17
2.2.1	<i>Permutations</i>	20
2.2.2	<i>Combinations</i>	22
2.2.3	<i>Binomial coefficients</i>	23
2.3	Probability	32
2.3.1	<i>The postulates of probability</i>	34
2.3.2	<i>Some elementary theorems of probability</i>	36
2.4	Conditional probability	43
2.5	Some further theorems of probability	52
<b>Chapter 3</b>	<b>PROBABILITY DISTRIBUTIONS</b>	<b>61</b>
3.1	Random variables	61
3.2	Special probability distributions	66
3.2.1	<i>The binomial distribution</i>	66
3.2.2	<i>The hypergeometric distribution</i>	70
3.2.3	<i>The Poisson distribution</i>	72
3.2.4	<i>Some applications</i>	78
3.3	Multivariate probability distributions	81
3.3.1	<i>The multinomial distribution</i>	85

<b>Chapter 4</b>	<b>MATHEMATICAL EXPECTATION: DISCRETE RANDOM VARIABLES</b>	<b>90</b>
4.1	Mathematical expectation	90
4.2	Moments	94
4.2.1	<i>Chebyshev's theorem</i>	96
4.3	Moments of special probability distributions	99
4.3.1	<i>Moments of the binomial distribution</i>	99
4.3.2	<i>Moments of the hypergeometric distribution</i>	102
4.3.3	<i>Moments of the Poisson distribution</i>	103
4.4	Moment generating functions	106
4.4.1	<i>The moment generating function of the binomial distribution</i>	108
4.4.2	<i>The moment generating function of the Poisson distribution</i>	108
4.4.3	<i>Moment generating functions and limiting distributions</i>	109
4.5	Product moments	112
4.6	Mathematical expectation and decision making	114
<b>Chapter 5</b>	<b>PROBABILITY DENSITIES</b>	<b>117</b>
5.1	Introduction	117
5.2	Probability densities and distribution functions	117
5.3	Special probability densities	125
5.3.1	<i>The uniform distribution</i>	126
5.3.2	<i>The exponential distribution</i>	126
5.3.3	<i>The gamma distribution</i>	127
5.3.4	<i>The normal distribution</i>	128
5.3.5	<i>Some applications</i>	131
5.4	Change of variable	132
5.5	Multivariate probability densities	137
<b>Chapter 6</b>	<b>MATHEMATICAL EXPECTATION: CONTINUOUS RANDOM VARIABLES</b>	<b>143</b>
6.1	Mathematical expectation	143
6.2	Moments	144
6.3	Moments of special probability densities	145
6.3.1	<i>Moments of the uniform distribution</i>	146
6.3.2	<i>Moments of the gamma distribution</i>	146
6.3.3	<i>Moments of the normal distribution</i>	147
6.4	Moment generating functions	152
6.4.1	<i>Some properties of moment generating functions</i>	153
6.4.2	<i>Moment generating functions of special distributions</i>	154
6.4.3	<i>Moment generating functions and limiting distributions</i>	157
6.5	Product moments	160
6.6	Mathematical expectation and decision making	161

Chapter 7	<b>SUMS OF RANDOM VARIABLES</b>	<b>164</b>
7.1	Introduction	164
7.2	Sums of random variables—convolutions	166
7.3	Sums of random variables—moment generating functions	171
7.4	Moments of linear combinations of random variables	173
	7.4.1 <i>The distribution of the mean</i>	175
	7.4.2 <i>Differences between means and differences between proportions</i>	179
	7.4.3 <i>Sampling from finite populations</i>	181
	7.4.4 <i>The distribution of rank sums</i>	183
7.5	The central limit theorem	185
Chapter 8	<b>SAMPLING DISTRIBUTIONS</b>	<b>189</b>
8.1	Introduction	189
8.2	Sampling from normal populations	190
	8.2.1 <i>The distribution of <math>\bar{X}</math></i>	191
	8.2.2 <i>The chi-square distribution and the distribution of <math>s^2</math></i>	193
	8.2.3 <i>The <math>F</math> distribution</i>	199
	8.2.4 <i>The <math>t</math> distribution</i>	201
8.3	Sampling distributions of order statistics	204
Chapter 9	<b>POINT ESTIMATION</b>	<b>209</b>
9.1	Statistical inference and decision theory	209
9.2	Point estimation and interval estimation	214
9.3	Properties of point estimators	215
	9.3.1 <i>Unbiased estimators</i>	215
	9.3.2 <i>Consistency</i>	217
	9.3.3 <i>Relative efficiency</i>	218
	9.3.4 <i>Sufficiency</i>	219
9.4	Methods of point estimation	222
	9.4.1 <i>The method of moments</i>	222
	9.4.2 <i>The method of maximum likelihood</i>	223
Chapter 10	<b>INTERVAL ESTIMATION</b>	<b>227</b>
10.1	Confidence intervals	227
10.2	Confidence intervals for the mean	230
10.3	Confidence intervals for proportions	232
10.4	Confidence intervals for variances	234



<b>Chapter 11</b>	<b>TESTS OF HYPOTHESES: THEORY</b>	<b>237</b>
11.1	Introduction	237
11.2	Simple hypotheses	238
11.2.1	<i>Type I and Type II errors</i>	239
11.2.2	<i>The Neyman-Pearson lemma</i>	240
11.3	Composite hypotheses	246
11.3.1	<i>The power function of a test</i>	247
11.3.2	<i>Likelihood ratio tests</i>	251
<b>Chapter 12</b>	<b>TESTS OF HYPOTHESES: APPLICATIONS</b>	<b>259</b>
12.1	Introduction	259
12.2	Tests concerning means	261
12.2.1	<i>One-sample tests concerning means</i>	262
12.2.2	<i>Differences between means</i>	266
12.3	Tests concerning variances	271
12.4	Tests based on count data	274
12.4.1	<i>Tests concerning proportions</i>	275
12.4.2	<i>Differences among <math>k</math> proportions</i>	276
12.4.3	<i>Contingency tables</i>	282
12.4.4	<i>Tests of goodness of fit</i>	285
12.5	Nonparametric tests	289
12.5.1	<i>The sign test</i>	289
12.5.2	<i>Tests based on rank sums</i>	290
<b>Chapter 13</b>	<b>REGRESSION AND CORRELATION</b>	<b>295</b>
13.1	The problem of regression	295
13.2	Regression curves	296
13.3	Linear regression	299
13.4	The bivariate normal distribution	303
13.4.1	<i>Some sampling theory: correlation analysis</i>	303
13.4.2	<i>Some sampling theory: regression analysis</i>	314
13.5	The method of least squares	321
<b>Chapter 14</b>	<b>INTRODUCTION TO ANALYSIS OF VARIANCE</b>	<b>328</b>
14.1	Introduction	328
14.2	One-way analysis of variance	331
14.3	Two-way analysis of variance	338
14.4	Some further considerations	349

**CONTENTS**

**xiii**

<b>APPENDIX SUMS AND PRODUCTS</b>	<b>351</b>
<b>STATISTICAL TABLES</b>	<b>353</b>
<b>ANSWERS</b>	<b>372</b>
<b>INDEX</b>	<b>385</b>

# 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*.

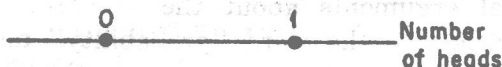


Figure 1.1

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.

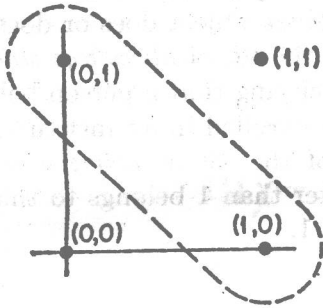


Figure 1.4

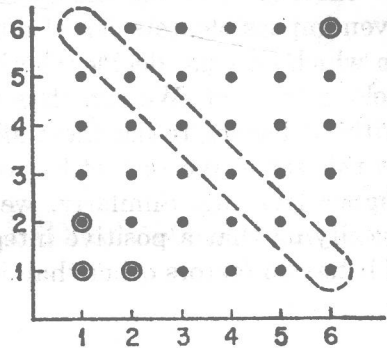


Figure 1.5

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  $\{\text{apple pie, ice cream, chocolate cake, rice pudding}\}$  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