



INTRODUCTION  
TO THE THEORY  
OF STATISTICS

**McGRAW-HILL**  
**BOOK COMPANY**  
New York  
St. Louis  
San Francisco  
Düsseldorf  
Johannesburg  
Kuala Lumpur  
London  
Mexico  
Montreal  
New Delhi  
Panama  
Rio de Janeiro  
Singapore  
Sydney  
Toronto

**ALEXANDER M. MOOD**

*Professor of Administration and  
Director of Public Policy Research Organization  
University of California, Irvine*

**FRANKLIN A. GRAYBILL**

*Department of Statistics  
Colorado State University  
Fort Collins, Colorado*

**DUANE C. BOES**

*Department of Statistics  
Colorado State University  
Fort Collins, Colorado*

# Introduction to the Theory of Statistics

**THIRD EDITION**

To HARRIET	A.M.M.
To my GRANDCHILDREN	F.A.G.
To JOAN, LISA, and KARIN	D.C.B.

#### **Library of Congress Cataloging in Publication Data**

Mood, Alexander McFarlane, 1913-  
Introduction to the theory of statistics.  
(McGraw-Hill series in probability and statistics)  
Bibliography: p.  
1. Mathematical statistics. I. Graybill,  
Franklin A., joint author. II. Boes, Duane C.,  
joint author. III. Title.  
QA276.M67 1974 519.5 73-292  
ISBN 0-07-042864-6

#### **INTRODUCTION TO THE THEORY OF STATISTICS**

Copyright © 1963, 1974 by McGraw-Hill, Inc. All rights reserved.  
Copyright 1950 by McGraw-Hill, Inc. All rights reserved.  
Printed in the United States of America. No part of this publication may be reproduced,  
stored in a retrieval system, or transmitted, in any form or by any means, electronic,  
mechanical, photocopying, recording, or otherwise, without the prior written  
permission of the publisher.

10 KPKP 83

This book was set in Times Roman.  
The editors were Brete C. Harrison and Madelaine Eichberg;  
the cover was designed by Nicholas Krenitsky;  
and the production supervisor was Ted Agrillo.  
The drawings were done by Oxford Illustrators Limited.  
The printer and binder was Kinsport Press, Inc.

## PREFACE TO THE THIRD EDITION

The purpose of the third edition of this book is to give a sound and self-contained (in the sense that the necessary probability theory is included) introduction to classical or mainstream statistical theory. It is not a statistical-methods-cookbook, nor a compendium of statistical theories, nor is it a mathematics book. The book is intended to be a textbook, aimed for use in the traditional full-year upper-division undergraduate course in probability and statistics, or for use as a text in a course designed for first-year graduate students. The latter course is often a “service course,” offered to a variety of disciplines.

No previous course in probability or statistics is needed in order to study the book. The mathematical preparation required is the conventional full-year calculus course which includes series expansion, multiple integration, and partial differentiation. Linear algebra is not required. An attempt has been made to talk to the reader. Also, we have retained the approach of presenting the theory with some connection to practical problems. The book is not mathematically rigorous. Proofs, and even exact statements of results, are often not given. Instead, we have tried to impart a “feel” for the theory.

The book is designed to be used in either the quarter system or the semester system. In a quarter system, Chaps. I through V could be covered in the first

quarter, Chaps. VI through part of VIII the second quarter, and the rest of the book the third quarter. In a semester system, Chaps. I through VI could be covered the first semester and the remaining chapters the second semester. Chapter VI is a “bridging” chapter; it can be considered to be a part of “probability” or a part of “statistics.” Several sections or subsections can be omitted without disrupting the continuity of presentation. For example, any of the following could be omitted: Subsec. 4.5 of Chap. II; Subsecs., 2.6, 3.5, 4.2, and 4.3 of Chap. III; Subsec. 5.3 of Chap. VI; Subsecs. 2.3, 3.4, 4.3 and Secs. 6 through 9 of Chap. VII; Secs. 5 and 6 of Chap. VIII; Secs. 6 and 7 of Chap. IX; and all or part of Chaps. X and XI. Subsection 5.3 of Chap VI on extreme-value theory is somewhat more difficult than the rest of that chapter. In Chap. VII, Subsec. 7.1 on Bayes estimation can be taught without Subsec. 3.4 on loss and risk functions but Subsec. 7.2 cannot. Parts of Sec. 8 of Chap. VII utilize matrix notation. The many problems are intended to be essential for learning the material in the book. Some of the more difficult problems have been starred.

ALEXANDER M. MOOD  
FRANKLIN A. GRAYBILL  
DUANE C. BOES

## EXCERPTS FROM THE FIRST AND SECOND EDITION PREFACES

This book developed from a set of notes which I prepared in 1945. At that time there was no modern text available specifically designed for beginning students of mathematical statistics. Since then the situation has been relieved considerably, and had I known in advance what books were in the making it is likely that I should not have embarked on this volume. However, it seemed sufficiently different from other presentations to give prospective teachers and students a useful alternative choice.

The aforementioned notes were used as text material for three years at Iowa State College in a course offered to senior and first-year graduate students. The only prerequisite for the course was one year of calculus, and this requirement indicates the level of the book. (The calculus class at Iowa State met four hours per week and included good coverage of Taylor series, partial differentiation, and multiple integration.) No previous knowledge of statistics is assumed.

This is a statistics book, not a mathematics book, as any mathematician will readily see. Little mathematical rigor is to be found in the derivations simply because it would be boring and largely a waste of time at this level. Of course rigorous thinking is quite essential to good statistics, and I have been at some pains to make a show of rigor and to instill an appreciation for rigor by pointing out various pitfalls of loose arguments.

While this text is primarily concerned with the theory of statistics, full cognizance has been taken of those students who fear that a moment may be wasted in mathematical frivolity. All new subjects are supplied with a little scenery from practical affairs, and, more important, a serious effort has been made in the problems to illustrate the variety of ways in which the theory may be applied.

The problems are an essential part of the book. They range from simple numerical examples to theorems needed in subsequent chapters. They include important subjects which could easily take precedence over material in the text; the relegation of subjects to problems was based rather on the feasibility of such a procedure than on the priority of the subject. For example, the matter of correlation is dealt with almost entirely in the problems. It seemed to me inefficient to cover multivariate situations twice in detail, i.e., with the regression model and with the correlation model. The emphasis in the text proper is on the more general regression model.

The author of a textbook is indebted to practically everyone who has touched the field, and I here bow to all statisticians. However, in giving credit to contributors one must draw the line somewhere, and I have simplified matters by drawing it very high; only the most eminent contributors are mentioned in the book.

I am indebted to Catherine Thompson and Maxine Merrington, and to E. S. Pearson, editor of *Biometrika*, for permission to include Tables III and V, which are abridged versions of tables published in *Biometrika*. I am also indebted to Professors R. A. Fisher and Frank Yates, and to Messrs. Oliver and Boyd, Ltd., Edinburgh, for permission to reprint Table IV from their book "Statistical Tables for Use in Biological, Agricultural and Medical Research."

Since the first edition of this book was published in 1950 many new statistical techniques have been made available and many techniques that were only in the domain of the mathematical statistician are now useful and demanded by the applied statistician. To include some of this material we have had to eliminate other material, else the book would have come to resemble a compendium. The general approach of presenting the theory with some connection to practical problems apparently contributed significantly to the success of the first edition and we have tried to maintain that feature in the present edition.

# CONTENTS

<b>Preface to the Third Edition</b>	<b>xiii</b>
<b>Excerpts from the First and Second Edition Prefaces</b>	<b>xv</b>
<b>I Probability</b>	<b>1</b>
1 Introduction and Summary	1
2 Kinds of Probability	2
2.1 Introduction	2
2.2 Classical or a Priori Probability	3
2.3 A Posteriori or Frequency Probability	5
3 Probability—Axiomatic	8
3.1 Probability Models	8
3.2 An Aside—Set Theory	9
3.3 Definitions of Sample Space and Event	14
3.4 Definition of Probability	19
3.5 Finite Sample Spaces	25
3.6 Conditional Probability and Independence	32



<b>II</b>	<b>Random Variables, Distribution Functions, and Expectation</b>	<b>51</b>
1	Introduction and Summary	51
2	Random Variable and Cumulative Distribution Function	52
2.1	Introduction	52
2.2	Definitions	53
3	Density Functions	57
3.1	Discrete Random Variables	57
3.2	Continuous Random Variables	60
3.3	Other Random Variables	62
4	Expectations and Moments	64
4.1	Mean	64
4.2	Variance	67
4.3	Expected Value of a Function of a Random Variable	69
4.4	Chebyshev Inequality	71
4.5	Jensen Inequality	72
4.6	Moments and Moment Generating Functions	72
<b>III</b>	<b>Special Parametric Families of Univariate Distributions</b>	<b>85</b>
1	Introduction and Summary	85
2	Discrete Distributions	86
2.1	Discrete Uniform Distribution	86
2.2	Bernoulli and Binomial Distributions	87
2.3	Hypergeometric Distribution	91
2.4	Poisson Distribution	93
2.5	Geometric and Negative Binomial Distributions	99
2.6	Other Discrete Distributions	103
3	Continuous Distributions	105
3.1	Uniform or Rectangular Distribution	105
3.2	Normal Distribution	107
3.3	Exponential and Gamma Distributions	111
3.4	Beta Distribution	115
3.5	Other Continuous Distributions	116
4	Comments	119
4.1	Approximations	119
4.2	Poisson and Exponential Relationship	121
4.3	Contagious Distributions and Truncated Distributions	122

<b>IV</b>	<b>Joint and Conditional Distributions, Stochastic Independence, More Expectation</b>	<b>129</b>
1	Introduction and Summary	129
2	Joint Distribution Functions	130
2.1	Cumulative Distribution Function	130
2.2	Joint Density Functions for Discrete Random Variables	133
2.3	Joint Density Functions for Continuous Random Variables	138
3	Conditional Distributions and Stochastic Independence	143
3.1	Conditional Distribution Functions for Discrete Random Variables	143
3.2	Conditional Distribution Functions for Continuous Random Variables	146
3.3	More on Conditional Distribution Functions	148
3.4	Independence	150
4	Expectation	153
4.1	Definition	153
4.2	Covariance and Correlation Coefficient	155
4.3	Conditional Expectations	157
4.4	Joint Moment Generating Function and Moments	159
4.5	Independence and Expectation	160
4.6	Cauchy-Schwarz Inequality	162
5	Bivariate Normal Distribution	162
5.1	Density Function	162
5.2	Moment Generating Function and Moments	164
5.3	Marginal and Conditional Densities	167
<b>V</b>	<b>Distributions of Functions of Random Variables</b>	<b>175</b>
1	Introduction and Summary	175
2	Expectations of Functions of Random Variables	176
2.1	Expectation Two Ways	176
2.2	Sums of Random Variables	178
2.3	Product and Quotient	180
3	Cumulative-distribution-function Technique	181
3.1	Description of Technique	181
3.2	Distribution of Minimum and Maximum	182
3.3	Distribution of Sum and Difference of Two Random Variables	185
3.4	Distribution of Product and Quotient	187

4	Moment-generating-function Technique	189
4.1	Description of Technique	189
4.2	Distribution of Sums of Independent Random Variables	192
5	The Transformation $Y = g(X)$	198
5.1	Distribution of $Y = g(X)$	198
5.2	Probability Integral Transform	202
6	Transformations	203
6.1	Discrete Random Variables	203
6.2	Continuous Random Variables	204
<b>VI</b>	<b>Sampling and Sampling Distributions</b>	<b>219</b>
1	Introduction and Summary	219
2	Sampling	220
2.1	Inductive Inference	220
2.2	Populations and Samples	222
2.3	Distribution of Sample	224
2.4	Statistic and Sample Moments	226
3	Sample Mean	230
3.1	Mean and Variance	231
3.2	Law of Large Numbers	231
3.3	Central-limit Theorem	233
3.4	Bernoulli and Poisson Distributions	236
3.5	Exponential Distribution	237
3.6	Uniform Distribution	238
3.7	Cauchy Distribution	238
4	Sampling from the Normal Distributions	239
4.1	Role of the Normal Distribution in Statistics	239
4.2	Sample Mean	240
4.3	The Chi-square Distribution	241
4.4	The $F$ Distribution	246
4.5	Student's $t$ Distribution	249
5	Order Statistics	251
5.1	Definition and Distributions	251
5.2	Distribution of Functions of Order Statistics	254
5.3	Asymptotic Distributions	256
5.4	Sample Cumulative Distribution Function	264

<b>VII Parametric Point Estimation</b>	<b>271</b>
1 Introduction and Summary	271
2 Methods of Finding Estimators	273
2.1 Methods of Moments	274
2.2 Maximum Likelihood	276
2.3 Other Methods	286
3 Properties of Point Estimators	288
3.1 Closeness	288
3.2 Mean-squared Error	291
3.3 Consistency and BAN	294
3.4 Loss and Risk Functions	297
4 Sufficiency	299
4.1 Sufficient Statistics	300
4.2 Factorization Criterion	307
4.3 Minimal Sufficient Statistics	311
4.4 Exponential Family	312
5 Unbiased Estimation	315
5.1 Lower Bound for Variance	315
5.2 Sufficiency and Completeness	321
6 Location or Scale Invariance	331
6.1 Location Invariance	332
6.2 Scale Invariance	336
7 Bayes Estimators	339
7.1 Posterior Distribution	340
7.2 Loss-function Approach	343
7.3 Minimax Estimator	350
8 Vector of Parameters	351
9 Optimum Properties of Maximum-likelihood Estimation	358
 <b>VIII Parametric Interval Estimation</b>	 <b>372</b>
1 Introduction and Summary	372
2 Confidence Intervals	373
2.1 An Introduction to Confidence Intervals	373
2.2 Definition of Confidence Interval	377
2.3 Pivotal Quantity	379

<b>3</b>	<b>Sampling from the Normal Distribution</b>	<b>381</b>
3.1	Confidence Interval for the Mean	381
3.2	Confidence Interval for the Variance	382
3.3	Simultaneous Confidence Region for the Mean and Variance	384
3.4	Confidence Interval for Difference in Means	386
<b>4</b>	<b>Methods of Finding Confidence Intervals</b>	<b>387</b>
4.1	Pivotal-quantity Method	387
4.2	Statistical Method	389
<b>5</b>	<b>Large-sample Confidence Intervals</b>	<b>393</b>
<b>6</b>	<b>Bayesian Interval Estimates</b>	<b>396</b>

## **IX Tests of Hypotheses 401**

<b>1</b>	<b>Introduction and Summary</b>	<b>401</b>
<b>2</b>	<b>Simple Hypothesis versus Simple Alternative</b>	<b>409</b>
2.1	Introduction	409
2.2	Most Powerful Test	410
2.3	Loss Function	414
<b>3</b>	<b>Composite Hypotheses</b>	<b>418</b>
3.1	Generalized Likelihood-ratio Test	419
3.2	Uniformly Most Powerful Tests	421
3.3	Unbiased Tests	425
3.4	Methods of Finding Tests	425
<b>4</b>	<b>Tests of Hypotheses—Sampling from the Normal Distribution</b>	<b>428</b>
4.1	Tests on the Mean	428
4.2	Tests on the Variance	431
4.3	Tests on Several Means	432
4.4	Tests on Several Variances	438
<b>5</b>	<b>Chi-square Tests</b>	<b>440</b>
5.1	Asymptotic Distribution of Generalized Likelihood-ratio	440
5.2	Chi-square Goodness-of-fit Test	442
5.3	Test of the Equality of Two Multinomial Distributions and Generalizations	448
5.4	Tests of Independence in Contingency Tables	452
<b>6</b>	<b>Tests of Hypotheses and Confidence Intervals</b>	<b>461</b>
<b>7</b>	<b>Sequential Tests of Hypotheses</b>	<b>464</b>
7.1	Introduction	464

7.2	Definition of Sequential Probability Ratio Test	466
7.3	Approximate Sequential Probability Ratio Test	468
7.4	Approximate Expected Sample Size of Sequential Probability Ratio Test	470
<b>X</b>	<b>Linear Models</b>	<b>482</b>
1	Introduction and Summary	482
2	Examples of the Linear Model	483
3	Definition of Linear Model	484
4	Point Estimation—Case A	487
5	Confidence Intervals—Case A	491
6	Tests of Hypotheses—Case A	494
7	Point Estimation—Case B	498
<b>XI</b>	<b>Nonparametric Methods</b>	<b>504</b>
1	Introduction and Summary	504
2	Inferences Concerning a Cumulative Distribution Function	506
2.1	Sample or Empirical Cumulative Distribution Function	506
2.2	Kolmogorov-Smirnov Goodness-of-fit Test	508
2.3	Confidence Bands for Cumulative Distribution Function	511
3	Inferences Concerning Quantiles	512
3.1	Point and Interval Estimates of a Quantile	512
3.2	Tests of Hypotheses Concerning Quantiles	514
4	Tolerance Limits	515
5	Equality of Two Distributions	518
5.1	Introduction	518
5.2	Two-sample Sign Test	519
5.3	Run Test	519
5.4	Median Test	521
5.5	Rank-sum Test	522
	<b>Appendix A. Mathematical Addendum</b>	<b>527</b>
1	Introduction	527

<b>2</b>	<b>Noncalculus</b>	<b>527</b>
2.1	Summation and Product Notation	527
2.2	Factorial and Combinatorial Symbols and Conventions	528
2.3	Stirling's Formula	530
2.4	The Binomial and Multinomial Theorems	530
<b>3</b>	<b>Calculus</b>	<b>531</b>
3.1	Preliminaries	531
3.2	Taylor Series	533
3.3	The Gamma and Beta Functions	534
	<b>Appendix B. Tabular Summary of Parametric Families of Distributions</b>	<b>537</b>
<b>1</b>	<b>Introduction</b>	<b>537</b>
	<i>Table 1.</i> Discrete Distributions	538
	<i>Table 2.</i> Continuous Distributions	540
	<b>Appendix C. References and Related Reading</b>	<b>544</b>
	Mathematics Books	544
	Probability Books	544
	Probability and Statistics Books	545
	Advanced (more advanced than MGB)	545
	Intermediate (about the same level as MGB)	545
	Elementary (less advanced than MGB, but calculus prerequisite)	546
	Special Books	546
	Papers	546
	Books of Tables	547
	<b>Appendix D. Tables</b>	<b>548</b>
<b>1</b>	<b>Description of Tables</b>	<b>548</b>
	<i>Table 1.</i> Ordinates of the Normal Density Function	548
	<i>Table 2.</i> Cumulative Normal Distribution	548
	<i>Table 3.</i> Cumulative Chi-square Distribution	549
	<i>Table 4.</i> Cumulative $F$ Distribution	549
	<i>Table 5.</i> Cumulative Student's $t$ Distribution	550
	<b>Index</b>	<b>557</b>

## 1 INTRODUCTION AND SUMMARY

The purpose of this chapter is to define *probability* and discuss some of its properties. Section 2 is a brief essay on some of the different meanings that have been attached to probability and may be omitted by those who are interested only in mathematical (axiomatic) probability, which is defined in Sec. 3 and used throughout the remainder of the text. Section 3 is subdivided into six subsections. The first, Subsec. 3.1, discusses the concept of probability models. It provides a real-world setting for the eventual mathematical definition of probability. A review of some of the set theoretical concepts that are relevant to probability is given in Subsec. 3.2. Sample space and event space are defined in Subsec. 3.3. Subsection 3.4 commences with a recall of the definition of a function. Such a definition is useful since many of the words to be defined in this and coming chapters (e.g., probability, random variable, distribution, etc.) are defined as particular functions. The indicator function, to be used extensively in later chapters, is defined here. The probability axioms are presented, and the probability function is defined. Several properties of this probability function are stated. The culmination of this subsection is the definition of a probability space. Subsection 3.5 is devoted to examples of probabilities



defined on finite sample spaces. The related concepts of independence of events and conditional probability are discussed in the sixth and final subsection. Bayes' theorem, the multiplication rule, and the theorem of total probabilities are proved or derived, and examples of each are given.

Of the three main sections included in this chapter, only Sec. 3, which is by far the longest, is vital. The definitions of probability, probability space, conditional probability, and independence, along with familiarity with the properties of probability, conditional and unconditional and related formulas, are the essence of this chapter. This chapter is a background chapter; it introduces the language of probability to be used in developing *distribution theory*, which is the backbone of the theory of statistics.

## 2 KINDS OF PROBABILITY

### 2.1 Introduction

One of the fundamental tools of statistics is probability, which had its formal beginnings with games of chance in the seventeenth century.

Games of chance, as the name implies, include such actions as spinning a roulette wheel, throwing dice, tossing a coin, drawing a card, etc., in which the outcome of a trial is uncertain. However, it is recognized that even though the outcome of any particular trial may be uncertain, there is a *predictable* long-term outcome. It is known, for example, that in many throws of an ideal (balanced, symmetrical) coin about one-half of the trials will result in heads. It is this long-term, predictable regularity that enables gaming houses to engage in the business.

A similar type of uncertainty and long-term regularity often occurs in experimental science. For example, in the science of genetics it is uncertain whether an offspring will be male or female, but in the long run it is known approximately what percent of offspring will be male and what percent will be female. A life insurance company cannot predict which persons in the United States will die at age 50, but it can predict quite satisfactorily *how many* people in the United States will die at that age.

First we shall discuss the classical, or *a priori*, theory of probability; then we shall discuss the frequency theory. Development of the axiomatic approach will be deferred until Sec. 3.