

Parametric Statistical Inference

Basic Theory
and Modern Approaches

by

SHELEMYAHU ZACKS

Parametric Statistical Inference

Basic Theory
and Modern Approaches

by

SHELEMYAHU ZACKS

State University of New York at Binghamton



PERGAMON PRESS

OXFORD · NEW YORK · TORONTO · SYDNEY · PARIS · FRANKFURT

U.K.	Pergamon Press Ltd., Headington Hill Hall, Oxford OX3 0BW, England
U.S.A.	Pergamon Press Inc., Maxwell House, Fairview Park, Elmsford, New York 10523, U.S.A.
CANADA	Pergamon of Canada, Suite 104, 150 Consumers Road, Willowdale, Ontario M2J 1P9, Canada
AUSTRALIA	Pergamon Press (Aust.) Pty. Ltd., P.O. Box 544, Potts Point, N.S.W. 2011, Australia
FRANCE	Pergamon Press SARL, 24 rue des Ecoles, 75240 Paris, Cedex 05, France
FEDERAL REPUBLIC OF GERMANY	Pergamon Press GmbH, 6242 Kronberg-Taunus, Hammerweg 6, Federal Republic of Germany

Copyright © 1981 S. Zacks

All Rights Reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means: electronic, electrostatic, magnetic tape, mechanical, photocopying, recording or otherwise, without permission in writing from the publishers.

First edition 1981

British Library Cataloguing in Publication Data

Zacks, Shelemyahu

Parametric statistical inference.

- (International series in nonlinear mathematics; vol.4). - (Pergamon international library).

1. Mathematical statistics

2. Probabilities

I. Title II. Series

519.5'4 QA276 80-41715

ISBN 0-08-026468-9 (Hardcover)

In order to make this volume available as economical-ly and as rapidly as possible the author's typescript has been reproduced in its original form. This method has its typographical limitations but it is hoped that they in no way distract the reader.

*Printed and bound in Great Britain by
William Clowes (Beccles) Limited, Beccles and London*

PREFACE

The present textbook is aimed at the population of senior undergraduate and graduate students in statistics, or mathematical sciences, who have previously taken an introductory course in probability and statistics. The objective of the present book is to present the basic theory and some of the recent developments in a framework of a two-semester intermediate level course on the theory of statistics. The mathematical prerequisites for this course are only advanced calculus and matrix algebra.

It has been a common experience that students taking an intermediate theory course often lack the proficiency required in the theory of statistical distributions, despite the fact that they have had a previous course in probability and statistics. For this reason the beginning of the book (Chapter 2) is devoted to the development of the required tools in the area of statistical theory. This chapter is followed by a discussion in Chapter 3 of the notions of statistical information, sufficient statistics, completeness, exponential families and the important information functions: the Fisher and the Kullback-Leibler. Chapter 4 is presenting the classical theory of testing hypotheses. Chapter 5 is devoted to estimation theory. A whole spectrum of estimation approaches is developed, from unbiased estimation to robust estimation, including jackknifing, maximum-likelihood, equivariant, moment-equations and pre-test estimators. In Chapter 6 the concept of relative efficiency of estimators in small and large samples is discussed. Chapter 7 is concerned with confidence and tolerance intervals. Finally, in Chapter 8, the theory of statistical inference is represented from the point of view of decision theory and the Bayesian approach. Many examples are completely developed in the text and numerous problems are given at the end of each chapter. The text was tried in various graduate courses at Case Western University, Tel Aviv University, The Technion and at Virginia Polytechnic Institute and State University. The experience at all of these institutions has been very encouraging, despite the fact that the students were heterogeneous in their background. The students who master the material in the present book are well prepared for further advanced studies and for research in applied statistics. The material in the present text is as important for the curriculum in applied statistics as it is for that in theoretical statistics.

The author wishes to acknowledge the help provided by many of his colleagues and graduate students. In particular, Professor Robert Berk provided excellent comments on an early draft, when we both spent a troubled year at Tel Aviv University. In addition, Professors Dan Anbar, Chris Tsokos, and Micha Yadin read

the text carefully and provided valuable comments. The author wishes to express his sincere thanks to Tel Aviv University, Case Western Reserve University, and Virginia Polytechnic Institute and State University for providing the conditions and assistance during the years in which the book was written. Special acknowledgment is given to the Office of Naval Research for continuing research contracts, resulting in many papers, from which examples were drawn. Finally, the author is grateful to Miss Judy Galliher who typed the final version of the text with devotion and a lot of intelligence.

Blacksburg, Virginia
July, 1980

Shelemyahu Zacks

CONTENTS

List of Illustrations	xiii
CHAPTER 1. GENERAL REVIEW	1
1.1 Introduction	1
1.2 Statistical Models, Distribution Functions and the Essence of Statistical Inference	2
1.3 The Information in Samples and Sufficient Statistics	4
1.4 Testing Statistical Hypotheses	5
1.5 Estimation Theory	6
1.6 The Efficiency of Estimators	8
1.7 Confidence and Tolerance Intervals	10
1.8 Decision Theoretic and Bayesian Approach in Testing and Estimation	11
CHAPTER 2 BASIC THEORY OF STATISTICAL DISTRIBUTIONS	15
2.1 Introductory Remarks	15
2.2 Elementary Properties of Distribution Functions	16
2.2.1 Discrete Distributions	16
2.2.2 Absolutely Continuous Distributions	17
2.2.3 Inverse Functions	17
2.2.4 Transformations	18
2.3 Some Families of Discrete Distributions	19
2.3.1 Binomial Distributions	19
2.3.2 Hypergeometric Distributions	20
2.3.3 Poisson Distributions	20
2.3.4 Geometric, Pascal and Negative Binomial	21
2.4 Some Families of Continuous Distributions	22
2.4.1 Rectangular Distributions	22
2.4.2 Beta Distributions	24
2.4.3 Gamma Distributions	24
2.4.4 Weibull and Extreme Value Distributions	25
2.4.5 Normal Distributions	26
2.4.6 Normal Approximations	27
2.5 Expectations, Moments and Generating Functions	29
2.6 Joint Distributions, Conditional Distributions and Independence	33
2.6.1 Joint Distributions	33
2.6.2 Conditional Distributions	36
2.6.3 Independence	39
2.6.4 Transformations	40
2.7 Moments and Covariances of Linear Functions	43

2.8	Discrete Multivariate Distributions	46
2.8.1	Multinomial Distributions	46
2.8.2	Multivariate Negative Binomial	47
2.8.3	Multivariate Hypergeometric	48
2.9	Multinormal Distributions	49
2.9.1	Basic Theory	49
2.9.2	Distributions of Subvectors and Distributions of Linear Forms	50
2.9.3	Independence of Linear Forms	53
2.9.4	Normal Probability Transformations	53
2.10	Distributions of Symmetric Quadratic Forms of Normal Variables	54
2.11	Independence of Linear and Quadratic Forms of Normal Variables	57
2.12	The Order Statistics	59
2.13	The t-Distributions	61
2.14	The F-Distributions	62
2.15	The Distribution of the Sample Correlation	65
2.16	Limit Theorems	67
2.17	Problems	70
CHAPTER 3.	SUFFICIENT STATISTICS AND THE INFORMATION IN SAMPLES	84
3.1	Introduction	84
3.2	Definitions and Characterization of Sufficient Statistics	85
3.3	Likelihood Functions and Minimal Sufficient Statistics	90
3.4	Sufficient Statistics and Exponential Type Families	95
3.5	Sufficiency and Completeness	99
3.6	Information Functions and Sufficiency	103
3.6.1	The Fisher Information	103
3.6.2	The Kullback-Leibler Information	106
3.7	Problems	109
CHAPTER 4.	TESTING STATISTICAL HYPOTHESES	113
4.1	The General Framework	113
4.2	The Neyman-Pearson Fundamental Lemma	118
4.3	Testing One-Sided Composite Hypotheses in MLR Models	123
4.4	Testing Two-Sided Hypotheses in One-Parameter Exponential Families	128
4.5	Testing Composite Hypotheses with Nuisance Parameters- Unbiased Tests	130
4.6	Likelihood Ratio Tests	140
4.6.1	Testing in Normal Regression Theory	141
4.6.2	Comparison of Normal Means: The Analysis of Variance	144
4.7	The Analysis of Contingency Tables	149
4.7.1	The Structure of Multi-Way Contingency Tables and the Statistical Model	149
4.7.2	Testing the Significance of Association	149
4.7.3	The Analysis of 2x2 Tables	151
4.7.4	Likelihood Ratio Tests	153
4.8	Sequential Testing of Hypotheses	154
4.8.1	The Wald Sequential Probability Ratio Test	155
4.8.2	Sequential Tests with Power 1	163
4.9	Problems	165
CHAPTER 5.	ESTIMATION THEORY	176
5.1	General Discussion	176
5.2	Unbiased Estimators	177
5.2.1	General Definition and Example	177
5.2.2	Minimum Variance Unbiased Estimators	179
5.2.3	Bias Reduction by Jackknifing	182
5.3	Best Linear Unbiased and Least Squares Estimator	184
5.3.1	Best Linear Unbiased Estimators of the Mean	185
5.3.2	Least Squares and Best Linear Unbiased Estimators for Linear Models	186

5.3.3	Best Linear Combination of Order Statistics	192
5.4	Stabilizing the Least Squares Estimators: Ridge Regression	194
5.5	Maximum Likelihood Estimation	198
5.5.1	Definition and Examples	198
5.5.2	Maximum Likelihood Estimators in Exponential Type Families	200
5.5.3	The Invariance Principle	201
5.5.4	Numerical Problems	202
5.5.5	Anomalous Cases	204
5.5.6	MLE of the Parameters of Tolerance Distributions	207
5.6	Equivariant Estimators	208
5.6.1	The Structure of Equivariant Estimators	208
5.6.2	Minimum MSE Equivariant Estimators	210
5.6.3	The Pitman Estimators	213
5.7	Moment-Equations Estimators	216
5.8	Pre-Test Estimators	220
5.9	Robust Estimators	222
5.10	Problems	225
CHAPTER 6.	THE EFFICIENCY OF ESTIMATORS	236
6.1	General Introduction	236
6.2	The Cramér-Rao Lower Bound in Regular One-Parameter Cases	237
6.3	Extension of the Cramér-Rao Inequality to Multiparameter Cases	240
6.4	General Inequalities of the Cramér-Rao Type	242
6.5	The Efficiency of Estimators in Small Samples	244
6.6	Asymptotic Properties of Estimators	246
6.6.1	The Consistency of MLE	247
6.6.2	Asymptotic Normality and Efficiency of MLE	248
6.7	Second-Order Asymptotic Efficiency	252
6.8	Maximum Probability Estimators	255
6.9	Problems	258
CHAPTER 7.	CONFIDENCE AND TOLERANCE INTERVALS	262
7.1	General Introduction	262
7.2	The Construction of Confidence Intervals	264
7.3	Optimal Confidence Intervals	266
7.4	Large Sample Approximations	272
7.5	Tolerance Intervals	276
7.6	Distribution-Free Confidence and Tolerance Intervals	278
7.7	Simultaneous Confidence Intervals	281
7.8	Two-Stage and Sequential Sampling for Fixed-Width Confidence Intervals	285
7.9	Problems	289
CHAPTER 8.	DECISION THEORETIC AND BAYESIAN APPROACH IN TESTING AND ESTIMATION	294
8.1	The Bayesian Framework	295
8.1.1	Prior, Posterior and Predictive Distributions	295
8.1.2	Bayesian Information Functions	297
8.1.3	Non-Informative and Improper Prior Distributions	299
8.1.4	Risk Functions, Bayes and Minimax Procedures	301
8.1.5	Bayes Sequential Decision Procedures	305
8.2	Bayesian Testing of Hypotheses	307
8.2.1	Testing Simple Hypotheses	307
8.2.2	Testing Composite Hypotheses	310
8.2.3	Bayes Sequential Testing of Hypotheses	315
8.3	Bayesian Confidence Intervals	317
8.4	Bayes and Minimax Estimators	321
8.4.1	General Discussion and Examples	321
8.4.2	Bayesian Estimates in Linear Models	323
8.4.3	Minimax Estimators	325

8.5	Minimax Risk and Bayes Equivariant, Formal Bayes and Structural Estimators	328
8.5.1	Minimum Risk and Bayes Equivariant Estimators	328
8.5.2	Formal Bayes Estimators for Invariant Priors	333
8.5.3	Equivariant Estimators Based on Structural Distributions	335
8.6	Empirical Bayes Estimators	339
8.7	The Admissibility of Estimators	342
8.7.1	Some Basic Results	342
8.7.2	The Inadmissibility of Some Commonly Used Estimators	347
8.7.3	Minimax and Admissible Estimators of the Location Parameter	354
8.8	Problems	356
REFERENCES		364
INDEXES		380
Author Index		380
Subject Index		384

LIST OF ILLUSTRATIONS

CHAPTER 2		
2.1	A Mixture of Negative-Exponential and Poisson Distributions	17
2.2	Densities Under Transformations	19
2.3	Moments of a Log-Normal Distribution	33
2.4	Covariance of a Joint Negative-Binomial and Negative-Exponential Distribution	36
2.5	Moment Generating Functions of Mixtures of Binomials and of Poissons	38
2.6	Distributions of Sums of Independent Random Variables	41
2.7	Distributions of Sums and Ratios of Independent Gamma Variables	42
2.8	The Mean and Variance of Random Samples From Symmetric Distributions are Uncorrelated (Proof)	45
2.9	Multivariate Negative-Binomial in an Inventory System	48
2.10	Some Characteristics of the Bivariate Normal Distribution	52
2.11	Expressing the Variance of a Sample as a Symmetric Idempotent Quadratic Form	57
2.12	The Independence of the Mean and Variance of a Random Sample From a Normal Distribution (Proof)	57
2.13	Normal Linear Models of Full Rank, Orthogonal Projections and Least-Squares	58
2.14	Distribution of Order Statistics of Continuous Uniform Variables	60
2.15	An Example of a t-Distribution	61
CHAPTER 3		
3.1	The Sufficiency of the Order-Statistic in a Random Sample	86
3.2	The Sufficiency of the Sum in a Random Sample From a Poisson Distribution	86
3.3	The Sufficiency of the Sum of a Random Sample From a Normal Distribution	87
3.4	Sufficient Statistics in the Bivariate Normal Case	89
3.5	Minimal Sufficient Statistics:	
A.	Binomial Distributions	92
B.	Hypergeometric Distributions	92
C.	Negative-Binomial Distributions	92
D.	Multinomial Distributions	93
E.	Beta Distributions	93
F.	Gamma Distributions	93
G.	Weibull Distributions	93
H.	Extreme-Value Distributions	94
I.	Normal Distributions	94

3.6	The Gammas as an Exponential Type Distribution	95
3.7	The Bivariate Normal as an Exponential Type Distribution	95
3.8	Minimal Sufficient Statistics Under the Random Effect Normal Model (Model II of ANOVA)	98
3.9	Completeness of the Families of	
	A. Binomial Distributions	100
	B. Rectangular Distributions	100
3.10	An Application of Basu's Theorem for Proving the Independence of Quadratic Forms in Model II of ANOVA	102
3.11	The Fisher Information in the Binomial Case	103
3.12	The Fisher Information in the Case of a Mixture of Two Normal Distributions	104
3.13	The Kullback-Leibler Information Functions in the Cases of	
	A. Normal Distributions	108
	B. Gamma Distributions	108
CHAPTER 4		
4.1	Testing Hypothesis in a Binomial Case	115
4.2	A Test Function and Its Power in the Rectangular Case	117
4.3	Testing a Simple Hypothesis in the Normal Case	121
4.4	Testing a Simple Hypothesis Concerning the Shape Parameter of a Weibull Distribution	122
4.5	Testing a One-Sided Composite Hypothesis Concerning the Variance of a Normal Distribution	123
4.6	The UMP Test of a One-Sided Hypothesis in the Binomial Case	127
4.7	UMP Unbiased Test of a Two-Sided Hypothesis in the Poisson Case	129
4.8	UMP Unbiased Test of a Two-Sided Hypothesis on the Mean of a Normal Distribution	130
4.9	Tests of Significance	
	A. Of Sample Means, Normal Distributions	132
	B. Of the Sample Correlation, Normal Distributions	132
4.10	UMP Unbiased Test of the Equality of the Means of Two Poisson Distributions	135
4.11	UMP Unbiased Test of the Interaction in Binomial Experiments	136
4.12	UMP Unbiasedness of the t-Test	139
4.13	UMP Unbiasedness of the ANOVA Test in Model II	139
4.14	Testing the Significance in a 2×2 Contingency Table	152
4.15	An SPRT of Hypothesis on the Mean of a Normal Distribution	161
CHAPTER 5		
5.1	Estimating the Probability of $\{X \geq \xi_0\}$ in Log-Normal Distribution	177
5.2	Unbiased Estimation of the End Point of a Rectangular Distribution	178
5.3	UMVU Estimator of the Tail Probability of a Normal Distribution, Mean and Variance Unknown	180
5.4	UMVU Estimators of Poisson Probabilities	181
5.5	Jackknifing Estimators for Negative-Exponential Distributions	183
5.6	Jackknifing Estimator of the Ratio of Means of Two Normal Distributions in the Correlated Case	183
5.7	BLUE of a Common Mean in a Bivariate Normal Distribution	185
5.8	LSE of the Multiple Regression Parameters in the Uncorrelated Case	187
5.9	BLUE of the Multiple Regression Parameters in the Correlated Case	191
5.10	BLUE of the Scale and Shape Parameters of Weibull Distributions	193
5.11	Ridge Estimators of the Multiple Regression Parameters	197
5.12	Maximum Likelihood Estimators in Cases of Rectangular Distributions	198
5.13	Maximum Likelihood Estimators of the Location and Scale Parameters of Laplace Distributions	199
5.14	Maximum Likelihood Estimators of the Parameters of Normal Distributions	200
5.15	Application of the Invariance Principle for the MLE of the Mean	

LIST OF ILLUSTRATIONS

xv

	and Variance of a Log-Normal Distribution	202
5.16	MLE of the Parameters of a Weibull Distribution	203
5.17	An Anomalous MLE	204
5.18	Neyman and Scott's Example of the MLE of the Common Variance of Normal Distributions Having Different Means	205
5.19	Minimum MSE Equivariant Estimators of the Parameters of Normal Distributions	210
5.20	Minimum MSE Equivariant Estimator of the Common Mean of Two Normal Distributions When the Ratio of Variances is Known	211
5.21	The Pitman Estimators of the Location and Scale Parameters of Negative Exponential Distributions	215
5.22	Moment Equations Estimators (MEE) of the Parameters of a Laplace Distribution	216
5.23	MEE of the Parameters of a Log-Normal Distribution	218
5.24	MEE of the Parameters of a Weibull Distribution	218
5.25	MEE of the Correlation Parameter of a Standard Bivariate Normal Distribution	219
5.26	Pre-Test Estimator of the Variance of a Normal Distribution	220
5.27	Pre-Test Estimator of the Common Mean of Two Normal Distributions With Ratio of Variances Unknown	221
CHAPTER 6		
6.1	The Cramér-Rao Lower Bound in Estimating Poisson Probabilities	238
6.2	The Bhattacharyaa Lower Bound of Order 2 in Estimating the Probability of 0 in the Poisson Case	239
6.3	Extended Cramér-Rao Lower Bounds for Estimating the Parameters of a Bivariate Normal Distribution	240
6.4	The Chapman-Robbins Lower Bound for Estimating the Translation Parameter of a Negative-Exponential Distribution	423
6.5	The Efficiency of the MLE in a Three-Parameter Normal Case	246
6.6	The Asymptotic Normal Distribution of the MLE of the Parameters of a Weibull Distribution	250
6.7	Second-Order Efficiency of the UMVU and MLE of a Reliability Function in the Poisson Case	253
6.8	Maximum Probability Estimators in the Normal Case	256
CHAPTER 7		
7.1	Lower Tolerance Limit for Log-Normal Distributions	263
7.2	Lower Confidence Limits for the Scale Parameter of Negative-Exponential Distributions	265
7.3	UMA Lower Confidence Limits for:	
	A. The Variance of a Normal Distribution	268
	B. The Probability of Success in Binomial Distributions	268
7.4	UMA Lower Confidence Limit for the Circular Probability of a Bivariate Normal Distribution; Ratio of Variances Known	268
7.5	UMPU Confidence Limits for the Mean of a Normal Distribution	270
7.6	UMPU Confidence Limits for the Interaction Parameter in Four Binomial Experiments	270
7.7	Asymptotic Confidence Limits for	
	A. The Parameter of Binomial Distributions	273
	B. The Correlation Parameter of Bivariate Distributions	273
7.8	Asymptotic Confidence Limits for the Interaction Parameter of Four Binomial Experiments	274
7.9	β -Content Tolerance Interval for Negative-Binomial Distributions	276
7.10	A $(p, 1-\alpha)$ Guaranteed Coverage Tolerance Limit for the Negative-Exponential Case Based on Truncated Life Testing	279
7.11	Simultaneous Confidence Intervals for the Mean and Variance of a Normal Distribution	282

CHAPTER 8

8.1	Bayesian Lower Prediction Limit for Binomial Distributions	297
8.2	Bayesian Adaptive Control of a Simple Inventory System	303
8.3	Bayesian Testing of Two Simple Hypotheses Concerning the Epoch of Shift Parameter in a Normal Case	308
8.4	Bayesian Detection of the Epoch of Shift Parameter for the Mean of Normal Distributions	311
8.5	Bayes Test for the Sign of the Mean of a Normal Distribution	312
8.6	Bayes Lower Confidence Limit of the Circular Probability of a Bivariate Normal Distribution	317
8.7	Bayesian Confidence Interval for the Inverse Regression	318
8.8	Bayes Estimator of the Mean of a Poisson Distribution	321
8.9	Bayes Estimators of the Mean of a Normal Distribution	322
8.10	Minimax Estimator of the Parameter of Binomial Distribution	326
8.11	Minimax Estimator of the Mean of a Normal Distribution	327
8.12	Minimum MSE Equivariant Estimators of the Parameters of a Weibull Distribution	329
8.13	Bayes Equivariant Estimators of Variance Components	332
8.14	Structural Estimators of the Location and Scale Parameters of Negative-Exponential Distributions	337
8.15	Empirical Bayes Estimators of the Proportion Defectives in Finite Lots	341

CHAPTER 1

General Review

1.1 INTRODUCTION

The theory of statistical inference is developing very fast, with thousands of research papers and dozens of monographs and textbooks published every year. It is impossible therefore to give in one short volume a complete or even partial account of all the important results. The objective of the present volume is to discuss some fundamental results in certain subfields of statistical inference, in a manner which does not require much previous preparation. The emphasis in the present monograph is on modern trends and developments. Certain subject matters could be only mentioned or briefly discussed, mainly because they require advanced treatment. In order to make the development accessible to a wide population of readers, not necessarily mathematicians, we have adopted a level of exposition which does not require much more than advanced calculus and some linear algebra. The theoretical discussion is accompanied with many examples, which illustrate the theory in special cases of interest. On the whole there are over one hundred such examples with additional short illustrations and many problems at the end of each chapter. Hundreds of references are cited and listed at the end. In order to avoid the prerequisite of proficiency in the theory of statistical distributions and to establish some required tools, the second chapter is devoted to a brief discussion of statistical distributions and their properties. Chapter 3 is devoted to the problem of sufficient statistics and the information in samples. Chapter 4 provides some basic results from the theory of testing statistical hypotheses. In Chapter 5 we develop the classical theory of estimation. The efficiency of estimators and some large sample properties are discussed in Chapter 6. Confidence intervals can be considered to be the bridge between testing hypotheses and point estimation. We study topics from this area in Chapter 7. The development of the

statistical theory of inference throughout Chapters 3-7 is from the classical point of view of the frequentist approach. Chapter 8 is devoted to the other point of view, namely the decision theoretical and Bayesian approach. In this chapter we again treat testing and estimation problems in the framework of statistical decision theory. Some results from the theory of sequential analysis are given in Chapter 4, 7 and 8. The treatment of this subject is, however, short and elementary since the subject of sequential analysis is strongly linked with the theory of optimal stopping. A good treatment of this theory requires advanced techniques. A short discussion of robust estimation procedures and adaptive techniques is provided in Chapter 5. There is no discussion, however, in the present book of design of experiments, multivariate analysis, time-series analysis, ranking and classification, stochastic approximation, statistical control theory, cluster analysis, pattern recognition and non-parametric techniques. All these are important subject areas of statistical research which are covered by specific monographs. The topics covered in the present volume are in a sense basic and common to many fields of statistical inference and thus provide a good preparation for further study. In the following sections we review the topics of the book chapter by chapter. This review is done in general terms, trying to avoid technicalities and to provide a general picture of the content of the book. To supplement the general discussion with more technical substance we provide at the beginning of each chapter a general introductory section, in which the topics of the chapter are introduced and discussed.

1.2 STATISTICAL MODELS, DISTRIBUTION FUNCTIONS AND THE ESSENCE OF STATISTICAL INFERENCE

Generally stating, statistical methods are needed when we study stochastic or random phenomena which are not completely predictable. More specifically, in almost all empirical studies there are uncontrollable elements of variability. For example, if we observe the average hourly wind velocity at a given meteorological station the variable under consideration is a stochastic, or random variable. We cannot generally predict with certainty what value this variable will assume in the next hour of observation. A complete characterization of the random variable being observed is given only if we can specify exactly the distribution function of this variable. The distribution function determines the probability that the observed random variable will assume values in specified intervals. Such a distribution function may depend on one or many parameters, which are related to the physical, biological or other type of system which governs the observed phenomenon. Such a mathematical description of an empirical phenomenon is always an ideal representation of the real world. We therefore call it a statistical or probabilistic model.

Statistical models depend generally on sets of assumptions. In the decision whether to apply a certain model to explain empirical observations we have first of all to question the relevance of the assumptions. If we do believe that the assumptions are reasonable we have to check whether the adopted model fits the observed data well. In Chapter 2 we study various types of distribution functions, set the assumptions needed for the applicability of these types as models and mention possible applications. The chapter is self-contained and can be studied without any prerequisite in probability or statistics. In addition to the basic properties of univariate and multivariate distributions the chapter contains material on conditional distributions, mixtures of distributions, the algebra of covariances, probability integral transforms, order statistics and limit theorems. The types of distributions considered are the Binomial, Hypergeometric, Poisson, Negative-Binomial, Rectangular, Beta, Gamma, Weibull, Extreme-Value, Normal, Multinomial, Multinormal, the order statistics, t -, F - and the distribution of the sample correlation. All these classes of distribution functions depend on one or several parameters. The parameters may vary over a specified range, called the parameter space. The collection of all the distribution functions of a given type, when the parameters vary over their possible range is called a family. Thus, we speak about the family of Binomial distributions, etc. Generally, in the statistical analysis of observed data, even if we apply a model specifying a certain family of distributions the actual values of the characterizing parameters are unknown. Assuming that the observations follow a distribution function belonging to the family specified by the model, the problems of statistical inference are those of deciding whether the actual parameters belong to a specified set in the parameter space or to estimate their actual values. These are problems of parametric inference known as testing hypotheses or estimation problems, respectively. One can consider also a different formulation of statistical inference problems, known as statistical decision problems. These are problems of choosing actions from specified sets of actions so that certain expected loss (or payoff) functions will be minimized (or maximized). The loss or payoff depends not only on the chosen action but also on the actual (true) distribution function. If this distribution is unknown observations are taken to supplement information on this distribution. The optimal choice of an action depends then on the observed values. The problem of statistical decision theory is to determine a proper function of the observations, called a decision function, according to which an optimal action can be taken.

In certain problems of statistical inference it is unwarranted to consider a model specifying a certain family of distributions of a known type or form. In other words, we may not have sufficient prior knowledge or evidence to adopt one family

of distributions and not another. In such cases we may resort to more general methods of inference known as distribution-free methods. Distribution-free methods, generally require larger number of observations to attain certain precision than parametric methods. On the other hand, the distribution-free methods are independent of assumptions about the particular functional type of the actual distribution function. For this reason the distribution-free methods are often considerably more robust and sturdy than the parametric methods. If the parametric model is wrong the consequences may lead sometimes to substantial loss or error. Most of the discussion in the present monograph relates to parametric inference. A few short sections are devoted to distribution-free estimation. There are many papers in the literature concerning robust but parametric methods of estimating location parameters of symmetric distributions. For a survey of the important results on this topic see Hampel (1973) and the monograph of Andrews, Bickel, Hampel, Huber, Rogers and Tukey (1972).

Another class of general models is the class of non-parametric models. Many of non-parametric methods are based on the ranks of the observed variables in the sample. We refer the reader to the books of Hollander and Wolfe (1973), Lehmann (1975) or Gibbons (1976) for a study of these methods and for references to the rich literature.

1.3 THE INFORMATION IN SAMPLES AND SUFFICIENT STATISTICS

In the classical approach the information in the sample is given only by the sample values of the observable random variables X_1, X_2, \dots, X_n . The Bayesian approach on the other hand allows also the addition of information extraneous to the sample (see Section 1.8). Let \underline{X} be the observable vector of the random variables to the sampled. Suppose that we consider a parametric model specified by a family F of distribution functions $F(x, \theta)$, where the parameter(s) θ belong(s) to a parameter space Θ . In relation to the family F a non-negative weight function, called the likelihood function, is defined over Θ . This is a function $L(\theta; \underline{X})$ of θ and of \underline{X} which assigns for each value of θ a value proportional to the probability density function of \underline{X} under θ . Two likelihood functions $L_1(\theta; \underline{X})$ and $L_2(\theta; \underline{X})$ are called equivalent if $L_1(\theta; \underline{X}) = C L_2(\theta; \underline{X})$ for all $\theta \in \Theta$. C is some positive constant that may depend on \underline{X} but not on θ . We denote this equivalence by \sim . If two different samples yield vectors $\underline{X}^{(1)}$ and $\underline{X}^{(2)}$ such that $L(\theta; \underline{X}^{(1)}) \sim L(\theta; \underline{X}^{(2)})$ for all θ in Θ , then the two samples are said to have the same amount of information on the model (Basu, 1975). The question is whether we can reduce the data to some function or functions of the vector \underline{X} without losing information. Functions of the observed random vector \underline{X} , which do not depend on unknown