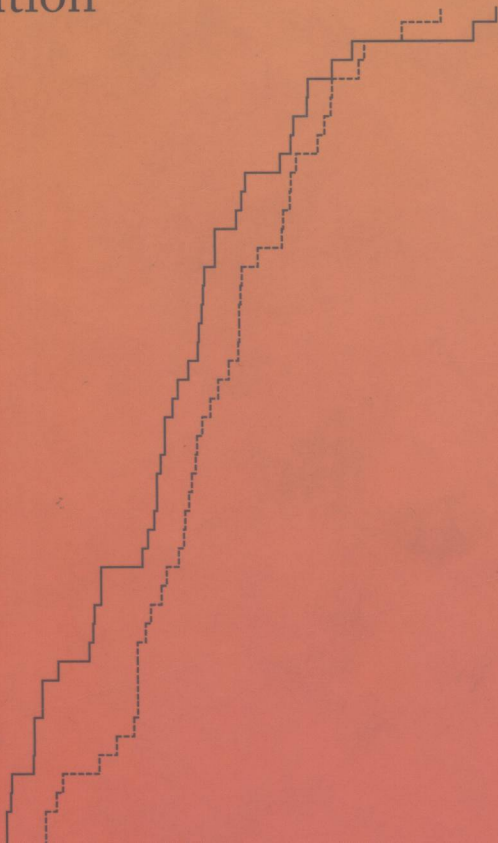


Frederick James

# Statistical Methods in Experimental Physics

2nd Edition



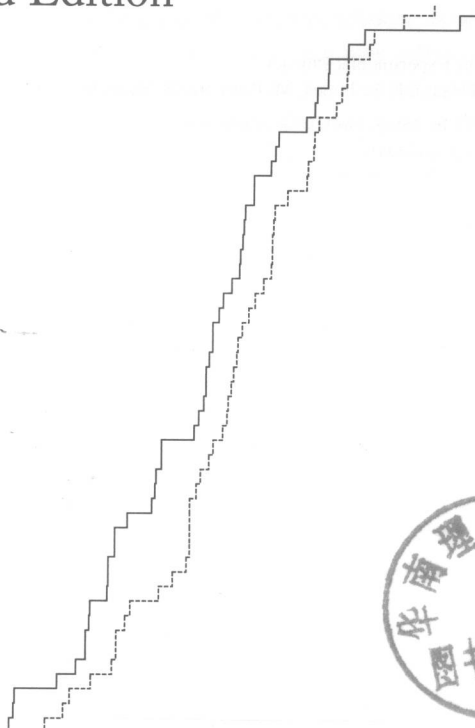
04-34  
J27  
E.2

Frederick James

*CERN, Switzerland*

# Statistical Methods in Experimental Physics

2nd Edition



E2009000997



**World Scientific**

NEW JERSEY • LONDON • SINGAPORE • BEIJING • SHANGHAI • HONG KONG • TAIPEI • CHENNAI

*Published by*

World Scientific Publishing Co. Pte. Ltd.

5 Toh Tuck Link, Singapore 596224

USA office: 27 Warren Street, Suite 401-402, Hackensack, NJ 07601

UK office: 57 Shelton Street, Covent Garden, London WC2H 9HE

**British Library Cataloguing-in-Publication Data**

A catalogue record for this book is available from the British Library.

The first edition:

Statistical Methods in Experimental Physics,

by W. T. Eadie, D. Drijard, F. E. James, M. Roos and B. Sadoulet

was published in 1971 by North-Holland Publishing Co.,

Amsterdam, New York and Oxford,

and reprinted in 1977, 1982 and 1988

**STATISTICAL METHODS IN EXPERIMENTAL PHYSICS (2nd Edition)**

Copyright © 2006 by World Scientific Publishing Co. Pte. Ltd.

*All rights reserved. This book, or parts thereof, may not be reproduced in any form or by any means, electronic or mechanical, including photocopying, recording or any information storage and retrieval system now known or to be invented, without written permission from the Publisher.*

For photocopying of material in this volume, please pay a copying fee through the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, USA. In this case permission to photocopy is not required from the publisher.

ISBN 981-256-795-X

ISBN 981-270-527-9 (pbk)

# Statistical Methods in Experimental Physics

2nd Edition

## PREFACE TO THE SECOND EDITION

Thirty years after the first publication of the first edition, it was decided that the continued demand justified the production of a new edition. The five original authors agreed that the new edition should reproduce as much as possible the complete text of the first edition, of course corrected for mistakes and modified to take into account recent developments. The original publisher ceded his rights to World Scientific Publishing Co. who kindly agreed to set the entire first edition in LaTeX to initiate the preparation of the new edition. Unfortunately, only one of the five original authors was ready to do this preparation, but the four others agreed to yield their rights in the interest of allowing the renewed availability of the book.

The new edition has required a considerable amount of work. For example, there are over 1000 formulae in the book, about half of which required modification, mostly for improved and consistent notation, but also to correct all the mistakes that have been reported over the years. In addition, the author of the second edition has had thirty years' additional experience in statistical data analysis, which necessarily translates into a better understanding of some problems and requires more than cosmetic changes in a few chapters. The overall result is that most of the text comes from the first edition, but the modifications are sufficiently important that the author of the second edition must take all the responsibility for the final text.

For the reader, the most striking difference between the editions will certainly be the improved typesetting. All the other benefits of computer preparation should make this edition much easier to read and more reliable than its predecessor.

F. E. James  
July 2006, Geneva

## PREFACE TO THE FIRST EDITION

This course in statistics, written by one statistician (W.T.E.) and four high-energy physicists, addresses itself to physicists (and experimenters in related sciences) in their task of extracting information from experimental data. Physicists often lack elementary knowledge of statistics, yet find themselves with problems requiring advanced methods, if adequate methods at all exist. To meet their needs, a sufficient course would have to be very long. Such courses do indeed exist [e.g. Kendall], only the physicists usually do not take the time to read them.

We attempt to give a course which is reasonably short, and yet sufficient for experimental physics. This obviously requires a compromise between theoretical rigour and amount of useful methods described.

Thus we are obliged to state many results without any rigorous proof (or with no proof at all); still we have the ambition to present more than just a cook-book of prescriptions and formulae. We omit the mention of many techniques which, in our judgement, seem to be of lesser importance to experimental physics.

On the other hand, we do introduce many theoretical concepts which may not seem immediately useful to the experimenter. This we think is necessary for two reasons. Firstly, the experimenter may need to know some theory or some “generalized methods” in order to design his own methods, experimental physics posing always novel questions. This is a justification for the stress on Information theory (Chapter 5), and for the attempt in Chapter 7 to define a “general” method of estimation. We hope that although the method the reader will arrive at may not be optimal, still it will be useful.

Secondly, the experimenter should be aware of the assumptions underlying a method, whether it be a standard method or his own. It is for this reason that we insist so much on the Central Limit Theorem, which is at the foundation of all “asymptotic” statistics (Chapters 3 and 7).

Quoting theorems, we also try to state their range of application, to avoid too careless use of some methods.

Among the underlying assumptions, especially important are the ones about the parent distributions of the data, since they will condition the results. In Chapter 4 we give a catalogue of useful ideal distributions; in real life they may have to be truncated (Sec. 4.3), experimental resolution may have to be folded in (Sec. 4.3), detection efficiency may have to be taken into account (Sec. 8.5). Moreover, the true distribution may not be known, in which case one is led to empirical distributions (Sec. 4.3), robust estimation (Sec. 8.7), and distribution-free tests (Chapter 11).

A very common tacit assumption in the everyday use of statistics is that the set of data is large enough for asymptotic conditions to apply. We try to distinguish clearly between asymptotic properties (usually simple whenever they are known) and finite sample properties (which are usually unknown). We also often give asymptotic expansions, in order to indicate how rapidly the asymptotic properties become true.

In general, we stress the various concepts of optimality. The justification for this is not only that this is the only way for a classical statistician to choose between different procedures, but also that experimental physicists handle ever increasing amounts of data, and therefore need increasingly optimal methods. However, there is an “optimal optimality”, because the last bit of optimality can often be achieved only at great cost. This introduces the aspect of economy, which we try to stress on many occasions.

Facing the controversy between Bayesians and Anti-Bayesians (“classical” statisticians), we tend to favour the classical approach (because of professional bias), however keeping the reader partly informed about the Bayesian approach throughout. This attitude we justify as follows. In Chapter 6 we show how taking a decision from a limited amount of information leads to a fundamental indeterminacy: any decision depends on *a priori* assumptions. These assumptions being largely subjective by nature, we think that it is not the role of an experimenter to take decisions. His aim should be to summarize the results of his experiment for the rest of the physics community in such a way as to convey a maximum of information about the unknowns measured. In a certain sense this leaves to the general consensus the task to take decisions.

This is our motivation to the Information theory approach to Estimation. Logically, Test theory should then be Bayesian (since testing really is a decision). Our excuse for not being Bayesian in Test theory (Chapters 10, 11) is that physicists, as a matter of general practice, consider a confidence level as an objective measure of the "distance" of the experiment from the hypothesis tested.

A minor consequence of our professional bias is that in contrast to most (if not all) books on probability and statistics, we avoid using examples from gambling. Physicists often find it frustrating trying to convert such examples into physics; therefore, our examples are taken from physics (mainly high-energy physics). The theory is, of course, the same and gamblers should not be discouraged from converting our examples back into card games, dice, etc!

Let us finally point out that we do not discuss numerical optimization techniques, very important e.g. in the methods of maximum likelihood and least squares. The reasons are that there exist in our opinion excellent treatises of optimization, should the experimenter want to know the details of optimum-searching algorithms, and most physicists do have powerful optimization programs at their disposal (e.g. in the CERN Computer Program Library), which save them one more worry.

W. T. Eadie

D. Drijard

F. E. James

M. Roos

B. Sadoulet

December 1970, Geneva



# CONTENTS

Preface to the Second Edition . . . . .	v
Preface to the First Edition . . . . .	vii
<b>Chapter 1. Introduction</b>	1
1.1. Outline . . . . .	1
1.2. Language . . . . .	2
1.3. Two Philosophies . . . . .	3
1.4. Notation . . . . .	4
<b>Chapter 2. Basic Concepts in Probability</b>	9
2.1. Definitions of Probability . . . . .	9
2.1.1. Mathematical probability . . . . .	10
2.1.2. Frequentist probability . . . . .	10
2.1.3. Bayesian probability . . . . .	11
2.2. Properties of Probability . . . . .	12
2.2.1. Addition law for sets of elementary events . . . . .	12
2.2.2. Conditional probability and independence . . . . .	13
2.2.3. Example of the addition law: scanning efficiency . . . . .	14
2.2.4. Bayes theorem for discrete events . . . . .	15
2.2.5. Bayesian use of Bayes theorem . . . . .	16
2.2.6. Random variable . . . . .	17
2.3. Continuous Random Variables . . . . .	18
2.3.1. Probability density function . . . . .	19
2.3.2. Change of variable . . . . .	20

2.3.3. Cumulative, marginal and conditional distributions . . . . .	21
2.3.4. Bayes theorem for continuous variables . . . . .	22
2.3.5. Bayesian use of Bayes theorem for continuous variables . . .	22
2.4. Properties of Distributions . . . . .	24
2.4.1. Expectation, mean and variance . . . . .	24
2.4.2. Covariance and correlation . . . . .	26
2.4.3. Linear functions of random variables . . . . .	28
2.4.4. Ratio of random variables . . . . .	30
2.4.5. Approximate variance formulae . . . . .	32
2.4.6. Moments . . . . .	33
2.5. Characteristic Function . . . . .	34
2.5.1. Definition and properties . . . . .	34
2.5.2. Cumulants . . . . .	37
2.5.3. Probability generating function . . . . .	38
2.5.4. Sums of a random number of random variables . . . . .	39
2.5.5. Invariant measures . . . . .	41
<b>Chapter 3. Convergence and the Law of Large Numbers</b>	<b>43</b>
3.1. The Tchebycheff Theorem and Its Corollary . . . . .	43
3.1.1. Tchebycheff theorem . . . . .	43
3.1.2. Bienaymé–Tchebycheff inequality . . . . .	44
3.2. Convergence . . . . .	45
3.2.1. Convergence in distribution . . . . .	45
3.2.2. The Paul Levy theorem . . . . .	46
3.2.3. Convergence in probability . . . . .	46
3.2.4. Stronger types of convergence . . . . .	47
3.3. The Law of Large Numbers . . . . .	47
3.3.1. Monte Carlo integration . . . . .	48
3.3.2. The Central Limit theorem . . . . .	49
3.3.3. Example: Gaussian (Normal) random number generator . .	51
<b>Chapter 4. Probability Distributions</b>	<b>53</b>
4.1. Discrete Distributions . . . . .	53
4.1.1. Binomial distribution . . . . .	53
4.1.2. Multinomial distribution . . . . .	56
4.1.3. Poisson distribution . . . . .	57
4.1.4. Compound Poisson distribution . . . . .	60
4.1.5. Geometric distribution . . . . .	62

4.1.6. Negative binomial distribution . . . . .	63
4.2. Continuous Distributions . . . . .	64
4.2.1. Normal one-dimensional (univariate Gaussian) . . . . .	64
4.2.2. Normal many-dimensional (multivariate Gaussian) . . . . .	67
4.2.3. Chi-square distribution . . . . .	70
4.2.4. Student's $t$ -distribution . . . . .	73
4.2.5. Fisher-Snedecor $F$ and $Z$ distributions . . . . .	77
4.2.6. Uniform distribution . . . . .	79
4.2.7. Triangular distribution . . . . .	79
4.2.8. Beta distribution . . . . .	80
4.2.9. Exponential distribution . . . . .	82
4.2.10. Gamma distribution . . . . .	83
4.2.11. Cauchy, or Breit-Wigner, distribution . . . . .	84
4.2.12. Log-Normal distribution . . . . .	85
4.2.13. Extreme value distribution . . . . .	87
4.2.14. Weibull distribution . . . . .	89
4.2.15. Double exponential distribution . . . . .	89
4.2.16. Asymptotic relationships between distributions . . . . .	90
4.3. Handling of Real Life Distributions . . . . .	91
4.3.1. General applicability of the Normal distribution . . . . .	91
4.3.2. Johnson empirical distributions . . . . .	92
4.3.3. Truncation . . . . .	93
4.3.4. Experimental resolution . . . . .	94
4.3.5. Examples of variable experimental resolution . . . . .	95
<b>Chapter 5. Information</b> . . . . .	<b>99</b>
5.1. Basic Concepts . . . . .	100
5.1.1. Likelihood function . . . . .	100
5.1.2. Statistic . . . . .	100
5.2. Information of R.A. Fisher . . . . .	101
5.2.1. Definition of information . . . . .	101
5.2.2. Properties of information . . . . .	101
5.3. Sufficient Statistics . . . . .	103
5.3.1. Sufficiency . . . . .	103
5.3.2. Examples . . . . .	104
5.3.3. Minimal sufficient statistics . . . . .	105
5.3.4. Darmais theorem . . . . .	106
5.4. Information and Sufficiency . . . . .	108

5.5. Example of Experimental Design . . . . .	109
<b>Chapter 6. Decision Theory</b>	111
6.1. Basic Concepts in Decision Theory . . . . .	112
6.1.1. Subjective probability, Bayesian approach . . . . .	112
6.1.2. Definitions and terminology . . . . .	113
6.2. Choice of Decision Rules . . . . .	114
6.2.1. Classical choice: pre-ordering rules . . . . .	114
6.2.2. Bayesian choice . . . . .	115
6.2.3. Minimax decisions . . . . .	116
6.3. Decision-theoretic Approach to Classical Problems . . . . .	117
6.3.1. Point estimation . . . . .	117
6.3.2. Interval estimation . . . . .	118
6.3.3. Tests of hypotheses . . . . .	118
6.4. Examples: Adjustment of an Apparatus . . . . .	121
6.4.1. Adjustment given an estimate of the apparatus performance	121
6.4.2. Adjustment with estimation of the optimum adjustment . .	123
6.5. Conclusion: Indeterminacy in Classical and Bayesian Decisions . .	124
<b>Chapter 7. Theory of Estimators</b>	127
7.1. Basic Concepts in Estimation . . . . .	127
7.1.1. Consistency and convergence . . . . .	128
7.1.2. Bias and consistency . . . . .	129
7.2. Usual Methods of Constructing Consistent Estimators . . . . .	130
7.2.1. The moments method . . . . .	131
7.2.2. Implicitly defined estimators . . . . .	132
7.2.3. The maximum likelihood method . . . . .	135
7.2.4. Least squares methods . . . . .	137
7.3. Asymptotic Distributions of Estimates . . . . .	139
7.3.1. Asymptotic Normality . . . . .	139
7.3.2. Asymptotic expansion of moments of estimates . . . . .	141
7.3.3. Asymptotic bias and variance of the usual estimators . . . .	144
7.4. Information and the Precision of an Estimator . . . . .	146
7.4.1. Lower bounds for the variance — Cramér–Rao inequality . .	147
7.4.2. Efficiency and minimum variance . . . . .	149
7.4.3. Cramér–Rao inequality for several parameters . . . . .	151
7.4.4. The Gauss–Markov theorem . . . . .	152
7.4.5. Asymptotic efficiency . . . . .	153

7.5. Bayesian Inference . . . . .	154
7.5.1. Choice of prior density . . . . .	154
7.5.2. Bayesian inference about the Poisson parameter . . . . .	156
7.5.3. Priors closed under sampling . . . . .	157
7.5.4. Bayesian inference about the mean, when the variance is known . . . . .	157
7.5.5. Bayesian inference about the variance, when the mean is known . . . . .	159
7.5.6. Bayesian inference about the mean and the variance . . . . .	161
7.5.7. Summary of Bayesian inference for Normal parameters . . . . .	162
<b>Chapter 8. Point Estimation in Practice</b> . . . . .	163
8.1. Choice of Estimator . . . . .	163
8.1.1. Desirable properties of estimators . . . . .	164
8.1.2. Compromise between statistical merits . . . . .	165
8.1.3. Cures to obtain simplicity . . . . .	166
8.1.4. Economic considerations . . . . .	168
8.2. The Method of Moments . . . . .	170
8.2.1. Orthogonal functions . . . . .	170
8.2.2. Comparison of likelihood and moments methods . . . . .	172
8.3. The Maximum Likelihood Method . . . . .	173
8.3.1. Summary of properties of maximum likelihood . . . . .	173
8.3.2. Example: determination of the lifetime of a particle in a restricted volume . . . . .	175
8.3.3. Academic example of a poor maximum likelihood estimate . . . . .	177
8.3.4. Constrained parameters in maximum likelihood . . . . .	179
8.4. The Least Squares Method (Chi-Square) . . . . .	182
8.4.1. The linear model . . . . .	183
8.4.2. The polynomial model . . . . .	185
8.4.3. Constrained parameters in the linear model . . . . .	186
8.4.4. Normally distributed data in nonlinear models . . . . .	190
8.4.5. Estimation from histograms; comparison of likelihood and least squares methods . . . . .	191
8.5. Weights and Detection Efficiency . . . . .	193
8.5.1. Ideal method maximum likelihood . . . . .	194
8.5.2. Approximate method for handling weights . . . . .	196
8.5.3. Exclusion of events with large weight . . . . .	199
8.5.4. Least squares method . . . . .	201

8.6. Reduction of Bias . . . . .	204
8.6.1. Exact distribution of the estimate known . . . . .	204
8.6.2. Exact distribution of the estimate unknown . . . . .	206
8.7. Robust (Distribution-free) Estimation . . . . .	207
8.7.1. Robust estimation of the centre of a distribution . . . . .	208
8.7.2. Trimming and Winsorization . . . . .	210
8.7.3. Generalized $p^{\text{th}}$ -power norms . . . . .	211
8.7.4. Estimates of location for asymmetric distributions . . . . .	213
<b>Chapter 9. Interval Estimation</b> . . . . .	<b>215</b>
9.1. Normally distributed data . . . . .	216
9.1.1. Confidence intervals for the mean . . . . .	216
9.1.2. Confidence intervals for several parameters . . . . .	218
9.1.3. Interpretation of the covariance matrix . . . . .	223
9.2. The General Case in One Dimension . . . . .	225
9.2.1. Confidence intervals and belts . . . . .	225
9.2.2. Upper limits, lower limits and flip-flopping . . . . .	227
9.2.3. Unphysical values and empty intervals . . . . .	229
9.2.4. The unified approach . . . . .	229
9.2.5. Confidence intervals for discrete data . . . . .	231
9.3. Use of the Likelihood Function . . . . .	233
9.3.1. Parabolic log-likelihood function . . . . .	233
9.3.2. Non-parabolic log-likelihood functions . . . . .	234
9.3.3. Profile likelihood regions in many parameters . . . . .	236
9.4. Use of Asymptotic Approximations . . . . .	238
9.4.1. Asymptotic Normality of the maximum likelihood estimate . . . . .	238
9.4.2. Asymptotic Normality of $\partial \ln L / \partial \theta$ . . . . .	238
9.4.3. $\partial L / \partial \theta$ confidence regions in many parameters . . . . .	240
9.4.4. Finite sample behaviour of three general methods of interval estimation . . . . .	240
9.5. Summary: Confidence Intervals and the Ensemble . . . . .	246
9.6. The Bayesian Approach . . . . .	248
9.6.1. Confidence intervals and credible intervals . . . . .	249
9.6.2. Summary: Bayesian or frequentist intervals? . . . . .	250
<b>Chapter 10. Test of Hypotheses</b> . . . . .	<b>253</b>
10.1. Formulation of a Test . . . . .	254
10.1.1. Basic concepts in testing . . . . .	254

10.1.2. Example: Separation of two classes of events . . . . .	255
10.2. Comparison of Tests . . . . .	257
10.2.1. Power . . . . .	257
10.2.2. Consistency . . . . .	259
10.2.3. Bias . . . . .	260
10.2.4. Choice of tests . . . . .	261
10.3. Test of Simple Hypotheses . . . . .	263
10.3.1. The Neyman–Pearson test . . . . .	263
10.3.2. Example: Normal theory test versus sign test . . . . .	264
10.4. Tests of Composite Hypotheses . . . . .	266
10.4.1. Existence of a uniformly most powerful test for the exponential family . . . . .	267
10.4.2. One- and two-sided tests . . . . .	268
10.4.3. Maximizing local power . . . . .	269
10.5. Likelihood Ratio Test . . . . .	270
10.5.1. Test statistic . . . . .	270
10.5.2. Asymptotic distribution for continuous families of hypotheses . . . . .	271
10.5.3. Asymptotic power for continuous families of hypotheses . .	273
10.5.4. Examples . . . . .	274
10.5.5. Small sample behaviour . . . . .	279
10.5.6. Example of separate families of hypotheses . . . . .	282
10.5.7. General methods for testing separate families . . . . .	285
10.6. Tests and Decision Theory . . . . .	287
10.6.1. Bayesian choice between families of distributions . . . . .	287
10.6.2. Sequential tests for optimum number of observations . . .	292
10.6.3. Sequential probability ratio test for a continuous family of hypotheses . . . . .	297
10.7. Summary of Optimal Tests . . . . .	298
<b>Chapter 11. Goodness-of-Fit Tests</b>	<b>299</b>
11.1. GOF Testing: From the Test Statistic to the P-value . . . . .	299
11.2. Pearson's Chi-square Test for Histograms . . . . .	301
11.2.1. Moments of the Pearson statistic . . . . .	302
11.2.2. Chi-square test with estimation of parameters . . . . .	303
11.2.3. Choosing optimal bin size . . . . .	304
11.3. Other Tests on Binned Data . . . . .	308
11.3.1. Runs test . . . . .	308

11.3.2. Empty cell test, order statistics . . . . .	309
11.3.3. Neyman–Barton smooth test . . . . .	311
11.4. Tests Free of Binning . . . . .	313
11.4.1. Smirnov–Cramér–von Mises test . . . . .	314
11.4.2. Kolmogorov test . . . . .	316
11.4.3. More refined tests based on the EDF . . . . .	317
11.4.4. Use of the likelihood function . . . . .	317
11.5. Applications . . . . .	318
11.5.1. Observation of a fine structure . . . . .	318
11.5.2. Combining independent estimates . . . . .	323
11.5.3. Comparing distributions . . . . .	327
11.6. Combining Independent Tests . . . . .	330
11.6.1. Independence of tests . . . . .	330
11.6.2. Significance level of the combined test . . . . .	331
References	335
Subject Index	341



## Chapter 1

# INTRODUCTION

### 1.1. Outline

The subject of the following ten chapters can be divided into two main parts:

- Theory of Probability: Chapters 2–4
- Statistics: Chapters 5–11

The theory of probability is needed only to provide the necessary tools for statistics, which forms the main body of the course.

Chapters 5 and 6 define two general approaches to the choice of estimators: the *information* approach and the *decision theory* approach. The former consists essentially in maximizing the amount of information in the estimate, whereas the latter is based on minimizing the loss involved in making the wrong decision about the parameter value. In the limit of large data samples, the two approaches are equivalent, but where they differ we will try to point out the distinction.

Estimation of parameters is divided into three chapters, 7 and 8 dealing with point estimation, theory and practice, and 9 dealing with interval estimation. Tests of hypotheses are divided into general testing, Chapter 10, and goodness-of-fit tests, Chapter 11.

Our reference policy is as follows. We quote literature when we have omitted the proof of an important result, or when we want to give hints for further