

Asymptotic Statistics

This book is an introduction to the field of asymptotic statistics. The treatment is both practical and mathematically rigorous. In addition to most of the standard topics of an asymptotics course, including likelihood inference, M -estimation, asymptotic efficiency, U -statistics, and rank procedures, the book also presents recent research topics such as semiparametric models, the bootstrap, and empirical processes and their applications.

One of the unifying themes is the approximation by limit experiments. This entails mainly the local approximation of the classical i.i.d. set-up with smooth parameters by location experiments involving a single, normally distributed observation. Thus, even the standard subjects of asymptotic statistics are presented in a novel way.

Suitable as a text for a graduate or Master's level statistics course, this book also gives researchers in statistics, probability, and their applications an overview of the latest research in asymptotic statistics.

A.W. van der Vaart is Professor of Statistics in the Department of Mathematics and Computer Science at the Vrije Universiteit, Amsterdam.

CAMBRIDGE SERIES IN STATISTICAL AND PROBABILISTIC MATHEMATICS

Editorial Board:

- R. Gill, *Department of Mathematics, Utrecht University*
B.D. Ripley, *Department of Statistics, University of Oxford*
S. Ross, *Department of Industrial Engineering, University of California, Berkeley*
M. Stein, *Department of Statistics, University of Chicago*
D. Williams, *School of Mathematical Sciences, University of Bath*

This series of high-quality upper-division textbooks and expository monographs covers all aspects of stochastic applicable mathematics. The topics range from pure and applied statistics to probability theory, operations research, optimization, and mathematical programming. The books contain clear presentations of new developments in the field and also of the state of the art in classical methods. While emphasizing rigorous treatment of theoretical methods, the books also contain applications and discussions of new techniques made possible by advances in computational practice.

Already published

1. *Bootstrap Methods and Their Application*, by A.C. Davison and D.V. Hinkley
2. *Markov Chains*, by J. Norris

Asymptotic Statistics

A. W. VAN DER VAART



CAMBRIDGE
 UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom
 One Liberty Plaza, 20th Floor, New York, NY 10006, USA
 477 Williamstown Road, Port Melbourne, VIC 3207, Australia
 314–321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre,
 New Delhi – 110025, India
 103 Penang Road, #05–06/07, Visioncrest Commercial, Singapore 238467

Cambridge University Press is part of the University of Cambridge.
 It furthers the University's mission by disseminating knowledge in the pursuit of
 education, learning, and research at the highest international levels of excellence.

www.cambridge.org
 Information on this title: www.cambridge.org/9780521784504
 DOI: 10.1017/CBO9780511802256

© Cambridge University Press 1998

This publication is in copyright. Subject to statutory exception
 and to the provisions of relevant collective licensing agreements,
 no reproduction of any part may take place without the written
 permission of Cambridge University Press.

First published 1998
 First paperback edition 2000
 8th printing 2007

A catalogue record for this publication is available from the British Library

Library of Congress Cataloguing in Publication Data

Vaart, A. W. van der

Asymptotic statistics / A. W. van der Vaart.

p. cm. – (Cambridge series in statistical and probabilistic
 mathematics)

Includes bibliographical references.

1. Mathematical statistical – Asymptotic theory. I. Title.

II. Series: Cambridge series on statistical and probabilistic mathematics

CA2276. V22 1998

519.5–dc21

98-15176

ISBN 978-0-521-49603-2 Hardback

ISBN 978-0-521-78450-4 Paperback

Cambridge University Press has no responsibility for the persistence or accuracy of
 URLs for external or third-party internet websites referred to in this publication
 and does not guarantee that any content on such websites is, or will remain,
 accurate or appropriate.

To Maryse and Marianne

Contents

<i>Preface</i>	<i>page</i> xiii
<i>Notation</i>	<i>page</i> xv
1. Introduction	1
1.1. Approximate Statistical Procedures	1
1.2. Asymptotic Optimality Theory	2
1.3. Limitations	3
1.4. The Index n	4
2. Stochastic Convergence	5
2.1. Basic Theory	5
2.2. Stochastic o and O Symbols	12
*2.3. Characteristic Functions	13
*2.4. Almost-Sure Representations	17
*2.5. Convergence of Moments	17
*2.6. Convergence-Determining Classes	18
*2.7. Law of the Iterated Logarithm	19
*2.8. Lindeberg-Feller Theorem	20
*2.9. Convergence in Total Variation	22
Problems	24
3. Delta Method	25
3.1. Basic Result	25
3.2. Variance-Stabilizing Transformations	30
*3.3. Higher-Order Expansions	31
*3.4. Uniform Delta Method	32
*3.5. Moments	33
Problems	34
4. Moment Estimators	35
4.1. Method of Moments	35
*4.2. Exponential Families	37
Problems	40
5. M - and Z -Estimators	41
5.1. Introduction	41
5.2. Consistency	44
5.3. Asymptotic Normality	51

*5.4.	Estimated Parameters	60
5.5.	Maximum Likelihood Estimators	61
*5.6.	Classical Conditions	67
*5.7.	One-Step Estimators	71
*5.8.	Rates of Convergence	75
*5.9.	Argmax Theorem	79
	Problems	83
6.	Contiguity	85
6.1.	Likelihood Ratios	85
6.2.	Contiguity	87
	Problems	91
7.	Local Asymptotic Normality	92
7.1.	Introduction	92
7.2.	Expanding the Likelihood	93
7.3.	Convergence to a Normal Experiment	97
7.4.	Maximum Likelihood	100
*7.5.	Limit Distributions under Alternatives	103
*7.6.	Local Asymptotic Normality	103
	Problems	106
8.	Efficiency of Estimators	108
8.1.	Asymptotic Concentration	108
8.2.	Relative Efficiency	110
8.3.	Lower Bound for Experiments	111
8.4.	Estimating Normal Means	112
8.5.	Convolution Theorem	115
8.6.	Almost-Everywhere Convolution	
	Theorem	115
*8.7.	Local Asymptotic Minimax Theorem	117
*8.8.	Shrinkage Estimators	119
*8.9.	Achieving the Bound	120
*8.10.	Large Deviations	122
	Problems	123
9.	Limits of Experiments	125
9.1.	Introduction	125
9.2.	Asymptotic Representation Theorem	126
9.3.	Asymptotic Normality	127
9.4.	Uniform Distribution	129
9.5.	Pareto Distribution	130
9.6.	Asymptotic Mixed Normality	131
9.7.	Heuristics	136
	Problems	137
10.	Bayes Procedures	138
10.1.	Introduction	138
10.2.	Bernstein–von Mises Theorem	140

Contents

ix

10.3. Point Estimators	146
*10.4. Consistency	149
Problems	152
11. Projections	153
11.1. Projections	153
11.2. Conditional Expectation	155
11.3. Projection onto Sums	157
*11.4. Hoeffding Decomposition	157
Problems	160
12. U -Statistics	161
12.1. One-Sample U -Statistics	161
12.2. Two-Sample U -statistics	165
*12.3. Degenerate U -Statistics	167
Problems	171
13. Rank, Sign, and Permutation Statistics	173
13.1. Rank Statistics	173
13.2. Signed Rank Statistics	181
13.3. Rank Statistics for Independence	184
*13.4. Rank Statistics under Alternatives	184
13.5. Permutation Tests	188
*13.6. Rank Central Limit Theorem	190
Problems	190
14. Relative Efficiency of Tests	192
14.1. Asymptotic Power Functions	192
14.2. Consistency	199
14.3. Asymptotic Relative Efficiency	201
*14.4. Other Relative Efficiencies	202
*14.5. Rescaling Rates	211
Problems	213
15. Efficiency of Tests	215
15.1. Asymptotic Representation Theorem	215
15.2. Testing Normal Means	216
15.3. Local Asymptotic Normality	218
15.4. One-Sample Location	220
15.5. Two-Sample Problems	223
Problems	226
16. Likelihood Ratio Tests	227
16.1. Introduction	227
*16.2. Taylor Expansion	229
16.3. Using Local Asymptotic Normality	231
16.4. Asymptotic Power Functions	236

16.5. Bartlett Correction	238
*16.6. Bahadur Efficiency	238
Problems	241
17. Chi-Square Tests	242
17.1. Quadratic Forms in Normal Vectors	242
17.2. Pearson Statistic	242
17.3. Estimated Parameters	244
17.4. Testing Independence	247
*17.5. Goodness-of-Fit Tests	248
*17.6. Asymptotic Efficiency	251
Problems	253
18. Stochastic Convergence in Metric Spaces	255
18.1. Metric and Normed Spaces	255
18.2. Basic Properties	258
18.3. Bounded Stochastic Processes	260
Problems	263
19. Empirical Processes	265
19.1. Empirical Distribution Functions	265
19.2. Empirical Distributions	269
19.3. Goodness-of-Fit Statistics	277
19.4. Random Functions	279
19.5. Changing Classes	282
19.6. Maximal Inequalities	284
Problems	289
20. Functional Delta Method	291
20.1. von Mises Calculus	291
20.2. Hadamard-Differentiable Functions	296
20.3. Some Examples	298
Problems	303
21. Quantiles and Order Statistics	304
21.1. Weak Consistency	304
21.2. Asymptotic Normality	305
21.3. Median Absolute Deviation	310
21.4. Extreme Values	312
Problems	315
22. <i>L</i> -Statistics	316
22.1. Introduction	316
22.2. Hájek Projection	318
22.3. Delta Method	320
22.4. <i>L</i> -Estimators for Location	323
Problems	324
23. Bootstrap	326

Contents

xi

23.1. Introduction	326
23.2. Consistency	329
23.3. Higher-Order Correctness	334
Problems	339
24. Nonparametric Density Estimation	341
24.1 Introduction	341
24.2 Kernel Estimators	341
24.3 Rate Optimality	346
24.4 Estimating a Unimodal Density	349
Problems	356
25. Semiparametric Models	358
25.1 Introduction	358
25.2 Banach and Hilbert Spaces	360
25.3 Tangent Spaces and Information	362
25.4 Efficient Score Functions	368
25.5 Score and Information Operators	371
25.6 Testing	384
*25.7 Efficiency and the Delta Method	386
25.8 Efficient Score Equations	391
25.9 General Estimating Equations	400
25.10 Maximum Likelihood Estimators	402
25.11 Approximately Least-Favorable	
Submodels	408
25.12 Likelihood Equations	419
Problems	431
<i>References</i>	433
<i>Index</i>	439

Preface

This book grew out of courses that I gave at various places, including a graduate course in the Statistics Department of Texas A&M University, Master's level courses for mathematics students specializing in statistics at the Vrije Universiteit Amsterdam, a course in the DEA program (graduate level) of Université de Paris-sud, and courses in the Dutch AIO-netwerk (graduate level).

The mathematical level is mixed. Some parts I have used for second year courses for mathematics students (but they find it tough), other parts I would only recommend for a graduate program. The text is written both for students who know about the technical details of measure theory and probability, but little about statistics, and vice versa. This requires brief explanations of statistical methodology, for instance of what a rank test or the bootstrap is about, and there are similar excursions to introduce mathematical details. Familiarity with (higher-dimensional) calculus is necessary in all of the manuscript. Metric and normed spaces are briefly introduced in Chapter 18, when these concepts become necessary for Chapters 19, 20, 21 and 22, but I do not expect that this would be enough as a first introduction. For Chapter 25 basic knowledge of Hilbert spaces is extremely helpful, although the bare essentials are summarized at the beginning. Measure theory is implicitly assumed in the whole manuscript but can at most places be avoided by skipping proofs, by ignoring the word “measurable” or with a bit of handwaving. Because we deal mostly with i.i.d. observations, the simplest limit theorems from probability theory suffice. These are derived in Chapter 2, but prior exposure is helpful.

Sections, results or proofs that are preceded by asterisks are either of secondary importance or are out of line with the natural order of the chapters. As the chart in Figure 0.1 shows, many of the chapters are independent from one another, and the book can be used for several different courses.

A unifying theme is approximation by a limit experiment. The full theory is not developed (another writing project is on its way), but the material is limited to the “weak topology” on experiments, which in 90% of the book is exemplified by the case of smooth parameters of the distribution of i.i.d. observations. For this situation the theory can be developed by relatively simple, direct arguments. Limit experiments are used to explain efficiency properties, but also why certain procedures asymptotically take a certain form.

A second major theme is the application of results on abstract empirical processes. These already have benefits for deriving the usual theorems on M -estimators for Euclidean parameters but are indispensable if discussing more involved situations, such as M -estimators with nuisance parameters, chi-square statistics with data-dependent cells, or semiparametric models. The general theory is summarized in about 30 pages, and it is the applications

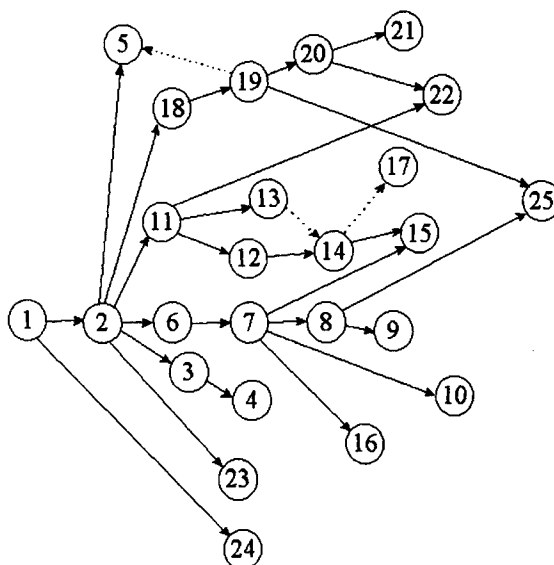


Figure 0.1. Dependence chart. A solid arrow means that a chapter is a prerequisite for a next chapter. A dotted arrow means a natural continuation. Vertical or horizontal position has no independent meaning.

that we focus on. In a sense, it would have been better to place this material (Chapters 18 and 19) earlier in the book, but instead we start with material of more direct statistical relevance and of a less abstract character. A drawback is that a few (starred) proofs point ahead to later chapters.

Almost every chapter ends with a “Notes” section. These are meant to give a rough historical sketch, and to provide entries in the literature for further reading. They certainly do not give sufficient credit to the original contributions by many authors and are not meant to serve as references in this way.

Mathematical statistics obtains its relevance from applications. The subjects of this book have been chosen accordingly. On the other hand, this is a mathematician’s book in that we have made some effort to present results in a nice way, without the (unnecessary) lists of “regularity conditions” that are sometimes found in statistics books. Occasionally, this means that the accompanying proof must be more involved. If this means that an idea could go lost, then an informal argument precedes the statement of a result.

This does not mean that I have strived after the greatest possible generality. A simple, clean presentation was the main aim.

Leiden, September 1997
 A.W. van der Vaart

Notation

A^*	adjoint operator
\mathbb{B}^*	dual space
$C_b(T), UC(T), C(T)$	(bounded, uniformly) continuous functions on T
$\ell^\infty(T)$	bounded functions on T
$\mathcal{L}_r(Q), L_r(Q)$	measurable functions whose r th powers are Q -integrable
$\ f\ _{Q,r}$	norm of $L_r(Q)$
$\ z\ _\infty, \ z\ _T$	uniform norm
lin	linear span
$\mathbb{C}, \mathbb{N}, \mathbb{Q}, \mathbb{R}, \mathbb{Z}$	number fields and sets
$EX, E^*X, \text{var } X, \text{sd } X, \text{Cov } X$	(outer) expectation, variance, standard deviation, covariance (matrix) of X
$\mathbb{P}_n, \mathbb{G}_n$	empirical measure and process
\mathbb{G}_P	P -Brownian bridge
$N(\mu, \Sigma), t_n, \chi_n^2$	normal, t and chisquare distribution
$z_\alpha, \chi_{n,\alpha}^2, t_{n,\alpha}$	upper α -quantiles of normal, chisquare and t distributions
\ll	absolutely continuous
$\triangleleft, \triangleleft \triangleright$	contiguous, mutually contiguous
\lesssim	smaller than up to a constant
\rightsquigarrow	convergence in distribution
\xrightarrow{P}	convergence in probability
$\xrightarrow{\text{as}}$	convergence almost surely
$N(\varepsilon, T, d), N_{[]}(\varepsilon, T, d)$	covering and bracketing number
$J(\varepsilon, T, d), J_{[]}(\varepsilon, T, d)$	entropy integral
$o_P(1), O_P(1)$	stochastic order symbols

