Cambridge University Press
978-0-521-65002-1 - Applications of Empirical Process Theory
Sara A. van de Geer
Frontmatter
More information

# Applications of Empirical Process Theory

CAMBRIDGE SERIES IN STATISTICAL AND PROBABILISTIC MATHEMATICS



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*, A.C. Davison and D.V. Hinkley
2. *Markov Chains*, J. Norris
3. *Asymptotic Statistics*, A.W. van der Vaart

Cambridge University Press
978-0-521-65002-1 - Applications of Empirical Process Theory
Sara A. van de Geer
Frontmatter
More information

# Applications of Empirical Process Theory

*Sara A. van de Geer*

CAMBRIDGE
UNIVERSITY PRESS

PUBLISHED BY THE PRESS SYNDICATE OF THE UNIVERSITY OF CAMBRIDGE
The Pitt Building, Trumpington Street, Cambridge, United Kingdom

CAMBRIDGE UNIVERSITY PRESS
The Edinburgh Building, Cambridge CB2 2RU, UK http://www.cup.cam.ac.uk
40 West 20th Street, New York, NY 10011-4211, USA http://www.cup.org
10 Stamford Road, Oakleigh, Melbourne 3166, Australia
Ruiz de Alarcón 13, 28014 Madrid, Spain



First published 2000

Printed in the United States of America

Typeset in Monotype Times 11/13pt, in TEX [EPC]

*A catalog record for this book is available from the British Library*

*Library of Congress Cataloging in Publication data*
Geer, S. A. van de (Sara A.)
Applications of empirical process theory / Sara A. van de Geer.
p. cm.
Includes bibliographical references and indexes.
ISBN 0 521 65002 X
1. Nonparametric statistics. 2. Estimation theory.
3. Limit theorems (Probability theory). I. Title.
QA278.8.G44 1999
519.5′4 dc21 98-50544 CIP

ISBN 0 521 65002 X hardback

# Contents

Cambridge University Press
978-0-521-65002-1 - Applications of Empirical Process Theory
Sara A. van de Geer
Frontmatter
More information

# Preface

This book is an extended version of a set of lecture notes, written for the *AIO* course 'Applications of Empirical Process Theory', which was given in the spring of 1996 in Utrecht, The Netherlands. The abbreviation *AIO* stands for *Assistent in Opleiding*, which is the Dutch equivalent of PhD student. The course was intended for students with a master of sciences in mathematics or statistics.

Nonparametric (infinite-dimensional) models provide a good alternative to the more classical parametric models, because fewer, and often more natural assumptions are imposed. In practice, nonparametric methods can be computationally complex, but this is nowadays a minor drawback. This book investigates the theoretical (asymptotic) properties of nonparametric M-estimators, with the emphasis on maximum likelihood estimators and least squares estimators. It treats the different models and estimation procedures in a unifying way, by invoking the theory on empirical processes. The general theory is illustrated with numerous examples. We hope that the methods provided will show that nonparametric models are in fact as basic as the more classical parametric models.

Empirical process theory has turned out to be a very valuable tool in asymptotic statistics. Applications include the bootstrap, the delta-method, and goodness-of-fit testing. In this book, we consider its applications to M-estimation, with special focus on maximum likelihood and least squares. We treat the latter two methods in great detail, including penalties, sieves, and semiparametrics. The description of M-estimators in general is in fact deferred to the very last chapter, where we show that, basically, the methodology allows direct generalization.

We do not assume any a priori knowledge on empirical process theory, and treat the subject, not in an exhaustive way, but with the applications always in mind. Moreover, we do not state any measurability conditions, because the formulation of these would require too many digressions. For most of the results, a full proof (modulo measurability) is presented, except when it concerns the calculation of entropy. Here, only some illustrations are given, which are meant to provide the reader with some understanding of what entropy actually is.

I am most grateful to my husband Toon, for the many fruitful discussions we had on the subject of this book, and for his very valuable suggestions on how to distinguish signal from noise.

Cambridge University Press
978-0-521-65002-1 - Applications of Empirical Process Theory
Sara A. van de Geer
Frontmatter
More information

# Guide to the Reader

Chapters 2, 3, 5, 6 and 8 present empirical processes theory, and the other chapters concern applications. Some of the material in this book is added only for completeness and can certainly be skipped at first reading.

Chapter 1 provides an overview of the type of problems the book will address. Chapters 2 and 3 are essential for the rest of the book. Most of the results in Chapter 3 are along the lines of Pollard (1984, Chapter II).

Chapter 4 treats consistency of estimators, and can be seen as a preparation for the more complicated theory on rates of convergence.

Chapter 5 is one of the more technical chapters. The main result there is Theorem 5.11, which says that the increments of an empirical process behave as the integral of the (square root) entropy (with bracketing). Once this message is clear, most of what follows can be understood without knowing too many precise details.

Chapter 6 contains some of the fundamental results on empirical processes. This chapter could of course not be excluded from the book, but it plays only a small role in the subsequent chapters.

Chapter 7 derives rates of convergence for maximum likelihood estimators. Some of its examples are rather technical, whereas others are quite elegant. A reader might find some of the examples artificial, because certain constants that depend on the unknown parameter, are assumed to be known a priori. But one has to keep in mind that these models serve as a preparation for the more realistic models of Chapter 10.

Section 8.1 of Chapter 8 is needed to make the step to independent, but not identically distributed variables. The section supplies the remainder of the technical tools necessary for the statistical applications. The rest of Chapter 8 is about dependent variables, and can be skipped without being penalized for that later on in the book.

In Chapter 9, we consider rates of convergence for least squares estimators. Here, the results appear in their neatest form. A reader may choose to skip Chapters 6 and 7, and go immediately to Chapter 9 after consulting Section 8.1.

The rest of the book contains some selected topics which can be, more or less, read separately. Chapter 10 considers penalties and sieves. Again, the results for the regression estimators are more transparent than those for maximum likelihood estimators.

Chapter 11 consists of three independent parts. There are some links with the previous chapters here, because the results on rates of convergence play a prominent role in the derivation of asymptotic normality of certain functions of the curve estimators.

In Chapter 12, we first summarize the general recipe we used so far only for least squares and maximum likelihood estimators. The idea there is that the methods carry over immediately to general M-estimators. It is possible to read Section 12.1 at any stage. This might help the reader to understand what the common underlying structure actually is. The last section of the book completes the circle: we started with a parametric model, and also end with it.

Each chapter concludes with a problems section, where we sometimes also complement the theory with some auxiliary results. The level of the problems varies considerably.