Empirical Processes Theory And Applications

This treatise by an acknowledged expert includes several topics not found in any previous book.
In a way, the world is made up of approximations, and surely there is no exception in the world of statistics. In fact, approximations, especially large sample approximations, are very important parts of both theoretical and - plied statistics.TheGaussiandistribution,alsoknownasthe normaldistri- tion,is merelyonesuchexample,dueto thewell-knowncentrallimittheorem. Large-sample techniques provide solutions to many practical problems; they simplify our solutions to di?cult, sometimes intractable problems; they j- tify our solutions; and they guide us to directions of improvements. On the other hand, just because large-sample approximations are used everywhere, and every day, it does not guarantee that they are used properly, and, when the techniques are misused, there may be serious consequences. 2 Example 1 (Asymptotic? distribution). Likelihood ratio test (LRT) is one of the fundamental techniques in statistics. It is well known that, in the 2 “standard” situation, the asymptotic null distribution of the LRT is?,with the degreesoffreedomequaltothe di?erencebetweenthedimensions,de?ned as the numbers of free parameters, of the two nested models being compared (e.g., Rice 1995, pp. 310). This might lead to a wrong impression that the 2 asymptotic (null) distribution of the LRT is always? . A similar mistake 2 might take place when dealing with Pearson’s? -test—the asymptotic distri- 2 2 bution of Pearson’s? -test is not always? (e.g., Moore 1978).
The theory of empirical processes provides valuable tools for the development of asymptotic theory in (nonparametric) statistical models, and makes it possible to give a unified treatment of various models. This book reveals the relation between the asymptotic behavior of M-estimators and the complexity of parameter space, using entropy as a measure of complexity, presenting tools and methods to analyze nonparametric, and in some cases, semiparametric methods. Graduate students and professionals in statistics, as well as those interested in applications, e.g. to econometrics, medical statistics, etc., will welcome this treatment.
An accessible account of the rich theory surrounding concentration inequalities in probability theory, with applications from machine learning and statistics to high-dimensional geometry. This book introduces key ideas and presents a detailed summary of the state-of-the-art in the area, making it ideal for independent learning and as a reference.
Empirical process techniques for independent data have been used for many years in statistics and probability theory. These techniques have proved very useful for studying asymptotic properties of parametric as well as non-parametric statistical procedures. Recently, the need to model the dependence structure in data sets from many different subject areas such as finance, insurance, and telecommunications has led to new developments concerning the empirical distribution function and the empirical process for dependent, mostly stationary sequences. This work gives an introduction to this new theory of empirical process techniques, which has so far been scattered in the statistical and probabilistic literature, and surveys the most recent developments in various related fields. Key features: A thorough and comprehensive introduction to the existing theory of empirical process techniques for dependent data * Accessible surveys by leading experts of the most recent developments in various related fields * Examines empirical process techniques for dependent data, useful for studying parametric and non-parametric statistical procedures * Comprehensive bibliographies * An overview of applications in various fields related to empirical processes: e.g., spectral analysis of time-series, the bootstrap for stationary sequences, extreme value theory, and the empirical process for mixing dependent observations, including the case of strong dependence. To date this book is the only comprehensive treatment of the topic in book literature. It is an ideal introductory text that will serve as a reference or resource for classroom use in the areas of statistics, time-series analysis, extreme value theory, point process theory, and applied probability theory. Contributors: P. Ango Nze, M.A. Arcones, I. Berkes, R. Dahlhaus, J. Dedecker, H.G. Dehling,
Self-normalized processes are of common occurrence in probabilistic and statistical studies. A prototypical example is Student's t-statistic introduced in 1908 by Gosset, whose portrait is on the front cover. Due to the highly non-linear nature of these processes, the theory experienced a long period of slow development. In recent years there have been a number of important advances in the theory and applications of self-normalized processes. Some of these developments are closely linked to the study of central limit theorems, which imply that self-normalized processes are approximate pivots for statistical inference. The present volume covers recent developments in the area, including self-normalized large and moderate deviations, and laws of the iterated logarithms for self-normalized martingales. This is the first book that systematically treats the theory and applications of self-normalization.
This volume highlights Prof. Hira Koul’s achievements in many areas of Statistics, including Asymptotic theory of statistical inference, Robustness, Weighted empirical processes and their applications, Survival Analysis, Nonlinear time series and Econometrics, among others. Chapters are all original papers that explore the frontiers of these areas and will assist researchers and graduate students working in Statistics, Econometrics and related areas. Prof. Hira Koul was the first Ph.D. student of Prof. Peter Bickel. His distinguished career in Statistics includes the receipt of many prestigious awards, including the Senior Humbolt award (1995), and dedicated service to the profession through editorial work for journals and through leadership roles in professional societies, notably as the past president of the International Indian Statistical Association. Prof. Hira Koul has graduated close to 30 Ph.D. students, and made several seminal contributions in about 125 innovative research papers. The long list of his distinguished collaborators is represented by the contributors to this volume.
Ces notes sont consacrées aux inégalités et aux théorèmes limites classiques pour les suites de variables aléatoires absolument régulières ou fortement mélangeantes au sens de Rosenblatt. Le but poursuivi est de donner des outils techniques pour l'étude des processus faiblement dépendants aux statisticiens ou aux probabilistes travaillant sur ces processus.
Originally published in 1986, this valuable reference provides a detailed treatment of limit theorems and inequalities for empirical processes of real-valued random variables. It also includes applications of the theory to censored data, spacings, rank statistics, quantiles, and many functionals of empirical processes, including a treatment of bootstrap methods, and a summary of inequalities that are useful for proving limit theorems. At the end of the Errata section, the authors have supplied references to solutions for 11 of the 19 Open Questions provided in the book's original edition.
The purpose of this book is to present results on the subject of weak convergence in function spaces to study invariance principles in statistical applications to dependent random variables, U-statistics, censor data analysis. Different techniques, formerly available only in a broad range of literature, are for the first time presented here in a self-contained fashion. Contents: Weak convergence of stochastic processes Weak convergence in metric spaces Weak convergence on C[0, 1] and D[0,?) Central limit theorem for semi-

martingales and applications Central limit theorems for dependent random variables Empirical process Bibliography
The fundamental question of characterizing continuity and boundedness of Gaussian processes goes back to Kolmogorov. After contributions by R. Dudley and X. Fernique, it was solved by the author. This book provides an overview of "generic chaining", a completely natural variation on the ideas of Kolmogorov. It takes the reader from the first principles to the edge of current knowledge and to the open problems that remain in this domain.
Empirical ProcessesTheory and ApplicationsIMSLectures on Empirical ProcessesTheory and Statistical ApplicationsTransaction PublishersWeak Convergence and Empirical ProcessesWith Applications to StatisticsSpringer Science & Business Media
Kosorok’s brilliant text provides a self-contained introduction to empirical processes and semiparametric inference. These powerful research techniques are surprisingly useful for developing methods of statistical inference for complex models and in understanding the properties of such methods. This is an authoritative text that covers all the bases, and also a friendly and gradual introduction to the area. The book can be used as research reference and textbook.
The purpose of these lecture notes is to provide an introduction to the general theory of empirical risk minimization with an emphasis on excess risk bounds and oracle inequalities in penalized problems. In recent years, there have been new developments in this area motivated by the study of new classes of methods in machine learning such as large margin classification methods (boosting, kernel machines). The main probabilistic tools involved in the analysis of these problems are concentration and deviation inequalities by Talagrand along with other methods of empirical processes theory (symmetrization inequalities, contraction inequality for Rademacher sums, entropy and generic chaining bounds). Sparse recovery based on l_1-type penalization and low rank matrix recovery based on the nuclear norm penalization are other active areas of research, where the main problems can be stated in the framework of penalized empirical risk minimization, and concentration inequalities and empirical processes tools have proved to be very useful.
High dimensional probability, in the sense that encompasses the topics rep resented in this volume, began about thirty years ago with research in two related areas: limit theorems for sums of independent Banach space valued random vectors and general Gaussian processes. An important feature in these past research studies has been the fact that they highlighted the es sential probabilistic nature of the problems considered. In part, this was because, by working on a general Banach space, one had to discard the extra, and often extraneous, structure imposed by random variables taking values in a Euclidean space, or by processes being indexed by sets in R or Rd. Doing this led to striking advances, particularly in Gaussian process theory. It also led to the creation or introduction of powerful new tools, such as randomization, decoupling, moment and exponential inequalities, chaining, isoperimetry and concentration of measure, which apply to areas well beyond those for which they were created. The general theory of em pirical processes, with its vast applications in statistics, the study of local times of Markov processes, certain problems in harmonic analysis, and the general theory of stochastic processes are just several of the broad areas in which Gaussian process techniques and techniques from probability in Banach spaces have made a substantial impact. Parallel to this work on probability in Banach spaces, classical proba bility and empirical process theory were enriched by the development of powerful results in strong approximations.
Advanced text; estimation methods in statistics, e.g. least squares; lots of examples; minimal abstraction.
This book compiles and presents new developments in statistical causal inference. The accompanying data and computer programs are publicly available so readers may replicate the model development and data analysis presented in each chapter. In this way, methodology is taught so that readers may implement it directly. The book brings together experts engaged in causal inference research to present and discuss recent issues in causal inference methodological development. This is also a timely look at causal inference applied to scenarios that range from clinical trials to mediation and public health research more broadly. In an academic setting, this book will serve as a reference and guide to a course in causal inference at the graduate level (Master's or Doctorate). It is particularly relevant for students pursuing degrees in statistics, biostatistics, and computational biology. Researchers and data analysts in public health and biomedical research will also find this book to be an important reference.
A preeminent expert in the field explores new and exciting methodologies in the ever-growing field of robust statistics Used to develop data analytical methods, which are resistant to outlying observations in the data, while capable of detecting outliers, robust statistics is extremely useful for solving an array of common problems, such as estimating location, scale, and regression parameters. Written by an internationally recognized expert in the field of robust statistics, this book addresses a range of well-established techniques while exploring, in depth, new and exciting methodologies. Local robustness and global robustness are discussed, and problems of non-identifiability and adaptive estimation are considered. Rather than attempt an exhaustive investigation of robustness, the author provides readers with a timely review of many of the most important problems in statistical inference involving robust estimation, along with a brief look at confidence intervals for location. Throughout, the author meticulously links research in maximum likelihood estimation with the more general M-estimation methodology. Specific applications and R and some MATLAB subroutines with accompanying data sets—available both in the text and online—are employed wherever appropriate. Providing invaluable insights and guidance, Robustness Theory and Application: Offers a balanced presentation of theory and applications within each topic-specific discussion Features solved examples throughout which help clarify complex and/or difficult concepts Meticulously links research in maximum likelihood type estimation with the more general M-estimation methodology Delves into new methodologies which have been developed over the past decade without stinting on coverage of “tried-and-true” methodologies Includes R and some MATLAB subroutines with accompanying data sets, which help

illustrate the power of the methods described Robustness Theory and Application is an important resource for all statisticians interested in the topic of robust statistics. This book encompasses both past and present research, making it a valuable supplemental text for graduate-level courses in robustness.

This book explores weak convergence theory and empirical processes and their applications to many applications in statistics. Part one reviews stochastic convergence in its various forms. Part two offers the theory of empirical processes in a form accessible to statisticians and probabilists. Part three covers a range of topics demonstrating the applicability of the theory to key questions such as measures of goodness of fit and the bootstrap.

How does a machine learn a new concept on the basis of examples? This second edition takes account of important new developments in the field. It also deals extensively with the theory of learning control systems, now comparably mature to learning of neural networks.

This classic introduction to probability theory for beginning graduate students covers laws of large numbers, central limit theorems, random walks, martingales, Markov chains, ergodic theorems, and Brownian motion. It is a comprehensive treatment concentrating on the results that are the most useful for applications. Its philosophy is that the best way to learn probability is to see it in action, so there are 200 examples and 450 problems. The fourth edition begins with a short chapter on measure theory to orient readers new to the subject.

Modern statistics deals with large and complex data sets, and consequently with models containing a large number of parameters. This book presents a detailed account of recently developed approaches, including the Lasso and versions of it for various models, boosting methods, undirected graphical modeling, and procedures controlling false positive selections. A special characteristic of the book is that it contains comprehensive mathematical theory on high-dimensional statistics combined with methodology, algorithms and illustrations with real data examples. This in-depth approach highlights the methods' great potential and practical applicability in a variety of settings. As such, it is a valuable resource for researchers, graduate students and experts in statistics, applied mathematics and computer science.

Sure to be influential, Watanabe's book lays the foundations for the use of algebraic geometry in statistical learning theory. Many models/machines are singular: mixture models, neural networks, HMMs, Bayesian networks, stochastic context-free grammars are major examples. The theory achieved here underpins accurate estimation techniques in the presence of singularities.

Functionals on stochastic processes; Uniform convergence of empirical measures; Convergence in distribution in euclidean spaces; Convergence in distribution in metric spaces; The uniform metric on space of cadlag functions; The skorohod metric on D [0, oo); Central limit teorems; Martingales.

Originally published in 1986, this valuable reference provides a detailed treatment of limit theorems and inequalities for empirical processes of real-valued random variables; applications of the theory to censored data, spacings, rank statistics, quantiles, and many functionals of empirical processes, including a treatment of bootstrap methods; and a summary of inequalities that are useful for proving limit theorems. At the end of the Errata section, the authors have supplied references to solutions for 11 of the 19 Open Questions provided in the book's original edition. Audience: researchers in statistical theory, probability theory, biostatistics, econometrics, and computer science.

This Proceedings volume contains a selection of invited and other papers by international scientists which were presented at the VIth International Vilnius Conference on Probability Theory and Mathematical Statistics, held in Vilnius, Lithuania, 28 June--3 July, 1993. The main topics of the conference were: limit theorems, stochastic analysis and stochastic physics, quantum probability theory, statistics, change detection in random processes, and probabilistic number theory.

The core of ths book presents a theory developed by the author to combine the recent insight into empirical data with mathematical models in freeway traffic research based on dynamical non-linear processes.

An integrated package of powerful probabilistic tools and key applications in modern mathematical data science.

This comprehensive Handbook presents the current state of art in the theory and methodology of macroeconomic data analysis. It is intended as a reference for graduate students and researchers interested in exploring new methodologies, but can also be employed as a graduate text. The Handbook concentrates on the most important issues, models and techniques for research in macroeconomics, and highlights the core methodologies and their empirical application in an accessible manner. Each chapter is largely self-contained, whilst the comprehensive introduction provides an overview of the key statistical concepts and methods. All of the chapters include the essential references for each topic and provide a sound guide for further reading. Topics covered include unit roots, non-linearities and structural breaks, time aggregation, forecasting, the Kalman filter, generalised method of moments, maximum likelihood and Bayesian estimation, vector autoregressive, dynamic stochastic general equilibrium and dynamic panel models. Presenting the most important models and techniques for empirical research, this Handbook will appeal to students, researchers and academics working in empirical macro and econometrics.

Benford's law states that the leading digits of many data sets are not uniformly distributed from one through nine, but rather exhibit a profound bias. This bias is evident in everything from electricity bills and street addresses to stock prices, population numbers, mortality rates, and the lengths of rivers. Here, Steven Miller brings together many of the world's leading experts on Benford's law to demonstrate the many useful techniques that arise from the law, show how truly multidisciplinary it is, and encourage collaboration. Beginning with the general theory, the contributors explain the prevalence of the bias, highlighting explanations for when systems should and should not follow Benford's law and how quickly such behavior sets in. They go on to discuss important applications in disciplines ranging from accounting and economics to psychology and the natural sciences. The contributors describe how Benford's law has been successfully used to expose fraud in elections, medical tests, tax filings, and financial reports. Additionally, numerous problems, background materials, and technical details are available online to help instructors create courses around the book. Emphasizing common challenges and techniques across the disciplines, this accessible book shows how Benford's law can serve as a productive meeting ground for researchers and practitioners in diverse fields.

As sintering applications march toward a $30 billion global business, the models for sintering have progressed, but generally follow behind observation. Documentation of the steps needed to build to a quantitative and predictive theory are often missed. Sintering: From Empirical Observations to Scientific Principles partitions sintering applications and observations to show critical turning points required to establish modern sintering as a predictive science. This book, written by the most cited author in his field, is laced with people, organizations, critical steps, and important formulations in a mixture of history, personalities, and applications. Exploring how insights in seemingly unrelated fields sparked progress, it is also a teaching tool to show where there is success, where there are problems, and how to organize teams to leapfrog to new applications or plateaus of use. Randall German's Sintering: From Empirical Observations to Scientific Principles is a platform for directly addressing the critical control parameters in

these new research and development efforts. Shows how the theories and understanding of sintering were developed and improved over time, and how different products were developed, ultimately leading to important knowledge and lessons for solving real sintering problems Covers all the necessary infrastructure of sintering theory and practice, such as atomic theory, surface energy, microstructure, and measurement and observation tools Introduces the history and development of such early sintered products as porcelain, tungsten lamp filaments, bronze bearings, steel automotive components, platinum crucibles and more