

# Statistics For High Dimensional Data Methods Theory And Applications

Right here, we have countless ebook Statistics For High Dimensional Data Methods Theory And Applications and collections to check out. We additionally pay for variant types and plus type of the books to browse. The normal book, fiction, history, novel, scientific research, as well as various extra sorts of books are readily to hand here.

As this Statistics For High Dimensional Data Methods Theory And Applications, it ends occurring living thing one of the favored ebook Statistics For High Dimensional Data Methods Theory And Applications collections that we have. This is why you remain in the best website to look the incredible books to have.

Bayesian and High-Dimensional Global Optimization Anatoly Zhigljavsky 2021-03-02 Accessible to a variety of readers, this book is of interest to specialists, graduate students and researchers in mathematics, optimization, computer science, operations research, management science, engineering and other applied areas interested in solving optimization problems. Basic principles, potential and boundaries of applicability of stochastic global optimization techniques are examined in this book. A variety of issues that face specialists in global optimization are explored, such as multidimensional spaces which are frequently ignored by researchers. The importance of precise interpretation of the mathematical results in assessments of optimization methods is demonstrated through examples of convergence in probability of random search. Methodological issues concerning construction and applicability of stochastic global optimization methods are discussed, including the one-step optimal average improvement method based on a statistical model of the objective function. A significant portion of this book is devoted to an analysis of high-dimensional global optimization problems and the so-called 'curse of dimensionality'. An examination of the three different classes of high-dimensional optimization problems, the geometry of high-dimensional balls and cubes, very slow convergence of global random search algorithms in large-dimensional problems, and poor uniformity of the uniformly distributed sequences of points are included in this book.

Statistical Inference from High Dimensional Data Carlos Fernandez-Lozano 2021-04-28 • Real-world problems can be high-dimensional, complex, and noisy • More data does not imply more information • Different approaches deal with the so-called curse of dimensionality to reduce irrelevant information • A process with multidimensional information is not necessarily easy to interpret nor process • In some real-world applications, the number of elements of a class is clearly lower than the other. The models tend to assume that the importance of the analysis belongs to the majority class and this is not usually the truth • The analysis of complex diseases such as cancer are focused on more-than-one dimensional omic data • The increasing amount of data thanks to the reduction of cost of the high-throughput experiments opens up a new era for integrative data-driven approaches • Entropy-based approaches are of interest to reduce the dimensionality of high-dimensional data

Statistical Models David A. Freedman 2009-04-27 This lively and engaging book explains the things you have to know in order to read empirical papers in the social and health sciences, as well as the techniques you need to build statistical models of your own. The discussion in the book is organized around published studies, as are many of the exercises. Relevant journal articles are reprinted at the back of the book. Freedman makes a thorough appraisal of the statistical methods in these papers and in a variety of other examples. He illustrates the principles of modelling, and the pitfalls. The discussion shows you how to think about the critical issues - including the connection (or lack of it) between the statistical models and the real phenomena. The book is written for advanced undergraduates and beginning graduate students in statistics, as well as students and professionals in the social and health sciences.

Statistical Learning with Sparsity Trevor Hastie 2015-05-07 Discover New Methods for Dealing with High-Dimensional Data A sparse statistical model has only a small number of nonzero parameters or weights; therefore, it is much easier to estimate and interpret than a dense model. Statistical Learning with

**Sparsity: The Lasso and Generalizations** presents methods that exploit sparsity to help recover the underlying signal in a set of data. Top experts in this rapidly evolving field, the authors describe the lasso for linear regression and a simple coordinate descent algorithm for its computation. They discuss the application of  $l_1$  penalties to generalized linear models and support vector machines, cover generalized penalties such as the elastic net and group lasso, and review numerical methods for optimization. They also present statistical inference methods for fitted (lasso) models, including the bootstrap, Bayesian methods, and recently developed approaches. In addition, the book examines matrix decomposition, sparse multivariate analysis, graphical models, and compressed sensing. It concludes with a survey of theoretical results for the lasso. In this age of big data, the number of features measured on a person or object can be large and might be larger than the number of observations. This book shows how the sparsity assumption allows us to tackle these problems and extract useful and reproducible patterns from big datasets. Data analysts, computer scientists, and theorists will appreciate this thorough and up-to-date treatment of sparse statistical modeling.

**Multivariate Statistics** Yasunori Fujikoshi 2011-08-15 A comprehensive examination of high-dimensional analysis of multivariate methods and their real-world applications **Multivariate Statistics: High-Dimensional and Large-Sample Approximations** is the first book of its kind to explore how classical multivariate methods can be revised and used in place of conventional statistical tools. Written by prominent researchers in the field, the book focuses on high-dimensional and large-scale approximations and details the many basic multivariate methods used to achieve high levels of accuracy. The authors begin with a fundamental presentation of the basic tools and exact distributional results of multivariate statistics, and, in addition, the derivations of most distributional results are provided. Statistical methods for high-dimensional data, such as curve data, spectra, images, and DNA microarrays, are discussed. Bootstrap approximations from a methodological point of view, theoretical accuracies in MANOVA tests, and model selection criteria are also presented. Subsequent chapters feature additional topical coverage including: High-dimensional approximations of various statistics High-dimensional statistical methods Approximations with computable error bound Selection of variables based on model selection approach Statistics with error bounds and their appearance in discriminant analysis, growth curve models, generalized linear models, profile analysis, and multiple comparison Each chapter provides real-world applications and thorough analyses of the real data. In addition, approximation formulas found throughout the book are a useful tool for both practical and theoretical statisticians, and basic results on exact distributions in multivariate analysis are included in a comprehensive, yet accessible, format. **Multivariate Statistics** is an excellent book for courses on probability theory in statistics at the graduate level. It is also an essential reference for both practical and theoretical statisticians who are interested in multivariate analysis and who would benefit from learning the applications of analytical probabilistic methods in statistics.

**Applied Biclustering Methods for Big and High-Dimensional Data Using R** Adetayo Kasim 2016-10-03 **Proven Methods for Big Data Analysis** As big data has become standard in many application areas, challenges have arisen related to methodology and software development, including how to discover meaningful patterns in the vast amounts of data. Addressing these problems, **Applied Biclustering Methods for Big and High-Dimensional Data Using R** shows how to apply biclustering methods to find local patterns in a big data matrix. The book presents an overview of data analysis using biclustering methods from a practical point of view. Real case studies in drug discovery, genetics, marketing research, biology, toxicity, and sports illustrate the use of several biclustering methods. References to technical details of the methods are provided for readers who wish to investigate the full theoretical background. All the methods are accompanied with R examples that show how to conduct the analyses. The examples, software, and other materials are available on a supplementary website.

**Introduction to High-Dimensional Statistics** Christophe Giraud 2014-12-17 Ever-greater computing technologies have given rise to an exponentially growing volume of data. Today massive data sets (with potentially thousands of variables) play an important role in almost every branch of modern human activity, including networks, finance, and genetics. However, analyzing such data has presented a challenge for statisticians and data analysts and has required the development of new statistical methods capable of separating the signal from the noise. **Introduction to High-Dimensional Statistics** is a concise guide to state-of-the-art models, techniques, and approaches for handling high-dimensional data. The book is intended to expose the reader to the key concepts and ideas in the most simple settings possible while avoiding unnecessary technicalities. Offering a succinct presentation of the mathematical foundations of high-dimensional statistics, this highly accessible text: Describes the challenges related to

the analysis of high-dimensional data Covers cutting-edge statistical methods including model selection, sparsity and the lasso, aggregation, and learning theory Provides detailed exercises at the end of every chapter with collaborative solutions on a wikisite Illustrates concepts with simple but clear practical examples Introduction to High-Dimensional Statistics is suitable for graduate students and researchers interested in discovering modern statistics for massive data. It can be used as a graduate text or for self-study.

Statistical Analysis for High-Dimensional Data Arnaldo Frigessi 2016-02-16 This book features research contributions from The Abel Symposium on Statistical Analysis for High Dimensional Data, held in Nuvågar, Lofoten, Norway, in May 2014. The focus of the symposium was on statistical and machine learning methodologies specifically developed for inference in "big data" situations, with particular reference to genomic applications. The contributors, who are among the most prominent researchers on the theory of statistics for high dimensional inference, present new theories and methods, as well as challenging applications and computational solutions. Specific themes include, among others, variable selection and screening, penalised regression, sparsity, thresholding, low dimensional structures, computational challenges, non-convex situations, learning graphical models, sparse covariance and precision matrices, semi- and non-parametric formulations, multiple testing, classification, factor models, clustering, and preselection. Highlighting cutting-edge research and casting light on future research directions, the contributions will benefit graduate students and researchers in computational biology, statistics and the machine learning community.

High-dimensional Data Analysis Tianwen Tony Cai 2011 Over the last few years, significant developments have been taking place in high-dimensional data analysis, driven primarily by a wide range of applications in many fields such as genomics and signal processing. In particular, substantial advances have been made in the areas of feature selection, covariance estimation, classification and regression. This book intends to examine important issues arising from high-dimensional data analysis to explore key ideas for statistical inference and prediction. It is structured around topics on multiple hypothesis testing, feature selection, regression, classification, dimension reduction, as well as applications in survival analysis and biomedical research. The book will appeal to graduate students and new researchers interested in the plethora of opportunities available in high-dimensional data analysis.

Model-Based Clustering and Classification for Data Science Charles Bouveyron 2019-07-25 Colorful example-rich introduction to the state-of-the-art for students in data science, as well as researchers and practitioners.

High-dimensional Microarray Data Analysis Shuichi Shinmura 2019-05-24 This book shows how to decompose high-dimensional microarrays into small subspaces (Small Matryoshkas, SMs), statistically analyze them, and perform cancer gene diagnosis. The information is useful for genetic experts, anyone who analyzes genetic data, and students to use as practical textbooks. Discriminant analysis is the best approach for microarray consisting of normal and cancer classes. Microarrays are linearly separable data (LSD, Fact 3). However, because most linear discriminant function (LDF) cannot discriminate LSD theoretically and error rates are high, no one had discovered Fact 3 until now. Hard-margin SVM (H-SVM) and Revised IP-OLDF (RIP) can find Fact3 easily. LSD has the Matryoshka structure and is easily decomposed into many SMs (Fact 4). Because all SMs are small samples and LSD, statistical methods analyze SMs easily. However, useful results cannot be obtained. On the other hand, H-SVM and RIP can discriminate two classes in SM entirely. RatioSV is the ratio of SV distance and discriminant range. The maximum RatioSVs of six microarrays is over 11.67%. This fact shows that SV separates two classes by window width (11.67%). Such easy discrimination has been unresolved since 1970. The reason is revealed by facts presented here, so this book can be read and enjoyed like a mystery novel. Many studies point out that it is difficult to separate signal and noise in a high-dimensional gene space. However, the definition of the signal is not clear. Convincing evidence is presented that LSD is a signal. Statistical analysis of the genes contained in the SM cannot provide useful information, but it shows that the discriminant score (DS) discriminated by RIP or H-SVM is easily LSD. For example, the Alon microarray has 2,000 genes which can be divided into 66 SMs. If 66 DSs are used as variables, the result is a 66-dimensional data. These signal data can be analyzed to find malignancy indicators by principal component analysis and cluster analysis.

Introduction to High-Dimensional Statistics Christophe Giraud 2014-12-17 Ever-greater computing technologies have given rise to an exponentially growing volume of data. Today massive data sets (with potentially thousands of variables) play an important role in almost every branch of modern human activity, including networks, finance, and genetics. However, analyzing such data has presented a

challenge for statisticians and data analysts and has required the development of new statistical methods capable of separating the signal from the noise. Introduction to High-Dimensional Statistics is a concise guide to state-of-the-art models, techniques, and approaches for handling high-dimensional data. The book is intended to expose the reader to the key concepts and ideas in the most simple settings possible while avoiding unnecessary technicalities. Offering a succinct presentation of the mathematical foundations of high-dimensional statistics, this highly accessible text: Describes the challenges related to the analysis of high-dimensional data Covers cutting-edge statistical methods including model selection, sparsity and the lasso, aggregation, and learning theory Provides detailed exercises at the end of every chapter with collaborative solutions on a wikisite Illustrates concepts with simple but clear practical examples Introduction to High-Dimensional Statistics is suitable for graduate students and researchers interested in discovering modern statistics for massive data. It can be used as a graduate text or for self-study.

Statistics for High-Dimensional Data Peter Bühlmann 2011-06-08 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.

Fundamentals of High-Dimensional Statistics Johannes Lederer 2021-11-16 This textbook provides a step-by-step introduction to the tools and principles of high-dimensional statistics. Each chapter is complemented by numerous exercises, many of them with detailed solutions, and computer labs in R that convey valuable practical insights. The book covers the theory and practice of high-dimensional linear regression, graphical models, and inference, ensuring readers have a smooth start in the field. It also offers suggestions for further reading. Given its scope, the textbook is intended for beginning graduate and advanced undergraduate students in statistics, biostatistics, and bioinformatics, though it will be equally useful to a broader audience.

Large-dimensional Panel Data Econometrics: Testing, Estimation And Structural Changes Chihwa Kao 2020-08-24 This book aims to fill the gap between panel data econometrics textbooks, and the latest development on 'big data', especially large-dimensional panel data econometrics. It introduces important research questions in large panels, including testing for cross-sectional dependence, estimation of factor-augmented panel data models, structural breaks in panels and group patterns in panels. To tackle these high dimensional issues, some techniques used in Machine Learning approaches are also illustrated. Moreover, the Monte Carlo experiments, and empirical examples are also utilised to show how to implement these new inference methods. Large-Dimensional Panel Data Econometrics: Testing, Estimation and Structural Changes also introduces new research questions and results in recent literature in this field.

Functional and High-Dimensional Statistics and Related Fields Germán Aneiros 2020-06-19 This book presents the latest research on the statistical analysis of functional, high-dimensional and other complex data, addressing methodological and computational aspects, as well as real-world applications. It covers topics like classification, confidence bands, density estimation, depth, diagnostic tests, dimension reduction, estimation on manifolds, high- and infinite-dimensional statistics, inference on functional data, networks, operatorial statistics, prediction, regression, robustness, sequential learning, small-ball probability, smoothing, spatial data, testing, and topological object data analysis, and includes applications in automobile engineering, criminology, drawing recognition, economics, environmetrics, medicine, mobile phone data, spectrometrics and urban environments. The book gathers selected, refereed contributions presented at the Fifth International Workshop on Functional and Operatorial Statistics (IWFOS) in Brno, Czech Republic. The workshop was originally to be held on June 24-26, 2020, but had to be postponed as a consequence of the COVID-19 pandemic. Initiated by the Working Group on Functional and Operatorial Statistics at the University of Toulouse in 2008, the IWFOS workshops provide a forum to discuss the latest trends and advances in functional statistics and related fields, and foster the exchange of ideas and international collaboration in the field.

Statistical Methods for Mediation, Confounding and Moderation Analysis Using R and SAS Qingzhao Yu 2022-03-14 Third-variable effect refers to the effect transmitted by third-variables that intervene in the

relationship between an exposure and a response variable. Differentiating between the indirect effect of individual factors from multiple third-variables is a constant problem for modern researchers. *Statistical Methods for Mediation, Confounding and Moderation Analysis Using R and SAS* introduces general definitions of third-variable effects that are adaptable to all different types of response (categorical or continuous), exposure, or third-variables. Using this method, multiple third-variables of different types can be considered simultaneously, and the indirect effect carried by individual third-variables can be separated from the total effect. Readers of all disciplines familiar with introductory statistics will find this a valuable resource for analysis. Key Features: Parametric and nonparametric method in third variable analysis Multivariate and Multiple third-variable effect analysis Multilevel mediation/confounding analysis Third-variable effect analysis with high-dimensional data Moderation/Interaction effect analysis within the third-variable analysis R packages and SAS macros to implement methods proposed in the book *Analysis of Mixed Data* Alexander R. de Leon 2013-01-16 A comprehensive source on mixed data analysis, *Analysis of Mixed Data: Methods & Applications* summarizes the fundamental developments in the field. Case studies are used extensively throughout the book to illustrate interesting applications from economics, medicine and health, marketing, and genetics. Carefully edited for smooth readability and Principles and Theory for Data Mining and Machine Learning Bertrand Clarke 2009-07-21 Extensive treatment of the most up-to-date topics Provides the theory and concepts behind popular and emerging methods Range of topics drawn from Statistics, Computer Science, and Electrical Engineering High-Dimensional Covariance Matrix Estimation Aygul Zagidullina

*Object Oriented Data Analysis* J. S. Marron 2021-11-18 *Object Oriented Data Analysis* is a framework that facilitates inter-disciplinary research through new terminology for discussing the often many possible approaches to the analysis of complex data. Such data are naturally arising in a wide variety of areas. This book aims to provide ways of thinking that enable the making of sensible choices. The main points are illustrated with many real data examples, based on the authors' personal experiences, which have motivated the invention of a wide array of analytic methods. While the mathematics go far beyond the usual in statistics (including differential geometry and even topology), the book is aimed at accessibility by graduate students. There is deliberate focus on ideas over mathematical formulas. J. S. Marron is the Amos Hawley Distinguished Professor of Statistics, Professor of Biostatistics, Adjunct Professor of Computer Science, Faculty Member of the Bioinformatics and Computational Biology Curriculum and Research Member of the Lineberger Cancer Center and the Computational Medicine Program, at the University of North Carolina, Chapel Hill. Ian L. Dryden is a Professor in the Department of Mathematics and Statistics at Florida International University in Miami, has served as Head of School of Mathematical Sciences at the University of Nottingham, and is joint author of the acclaimed book *Statistical Shape Analysis*.

*Big and Complex Data Analysis* S. Ejaz Ahmed 2017-03-21 This volume conveys some of the surprises, puzzles and success stories in high-dimensional and complex data analysis and related fields. Its peer-reviewed contributions showcase recent advances in variable selection, estimation and prediction strategies for a host of useful models, as well as essential new developments in the field. The continued and rapid advancement of modern technology now allows scientists to collect data of increasingly unprecedented size and complexity. Examples include epigenomic data, genomic data, proteomic data, high-resolution image data, high-frequency financial data, functional and longitudinal data, and network data. Simultaneous variable selection and estimation is one of the key statistical problems involved in analyzing such big and complex data. The purpose of this book is to stimulate research and foster interaction between researchers in the area of high-dimensional data analysis. More concretely, its goals are to: 1) highlight and expand the breadth of existing methods in big data and high-dimensional data analysis and their potential for the advancement of both the mathematical and statistical sciences; 2) identify important directions for future research in the theory of regularization methods, in algorithmic development, and in methodologies for different application areas; and 3) facilitate collaboration between theoretical and subject-specific researchers.

*Statistical Foundations of Data Science* Jianqing Fan 2020-09-21 *Statistical Foundations of Data Science* gives a thorough introduction to commonly used statistical models, contemporary statistical machine learning techniques and algorithms, along with their mathematical insights and statistical theories. It aims to serve as a graduate-level textbook and a research monograph on high-dimensional statistics, sparsity and covariance learning, machine learning, and statistical inference. It includes ample exercises that involve both theoretical studies as well as empirical applications. The book begins with an introduction to the stylized features of big data and their impacts on statistical analysis. It then introduces multiple linear

regression and expands the techniques of model building via nonparametric regression and kernel tricks. It provides a comprehensive account on sparsity explorations and model selections for multiple regression, generalized linear models, quantile regression, robust regression, hazards regression, among others. High-dimensional inference is also thoroughly addressed and so is feature screening. The book also provides a comprehensive account on high-dimensional covariance estimation, learning latent factors and hidden structures, as well as their applications to statistical estimation, inference, prediction and machine learning problems. It also introduces thoroughly statistical machine learning theory and methods for classification, clustering, and prediction. These include CART, random forests, boosting, support vector machines, clustering algorithms, sparse PCA, and deep learning.

High-dimensional Statistics Johannes Lederer 2020 "An Introduction to Regularized Estimation in High Dimensions considers statistical theory, methods, and algorithms for large and complex data. The main focus is on regularized estimators, which are at the cusp of entering the statistical toolkits of almost all scientific disciplines. This book provides clear expositions, motivational introductions to each chapter, rigorous step-by-step proofs, and comprehensive exercise sets with fully worked out solutions. These features make this book ideal for graduate level courses. Moreover, the book also discusses cutting-edge topics, such as aspects of inference, robustness, and tuning parameters. The book also contains results and insights that are new altogether, including improvements of existing theories and novel connections among different methods, which are organized into special chapters for those wishing to advance their reading in the field."--

Introduction to High-Dimensional Statistics Christophe Giraud 2021-08-26 Praise for the first edition: "[This book] succeeds singularly at providing a structured introduction to this active field of research. ... it is arguably the most accessible overview yet published of the mathematical ideas and principles that one needs to master to enter the field of high-dimensional statistics. ... recommended to anyone interested in the main results of current research in high-dimensional statistics as well as anyone interested in acquiring the core mathematical skills to enter this area of research." —Journal of the American Statistical Association Introduction to High-Dimensional Statistics, Second Edition preserves the philosophy of the first edition: to be a concise guide for students and researchers discovering the area and interested in the mathematics involved. The main concepts and ideas are presented in simple settings, avoiding thereby unessential technicalities. High-dimensional statistics is a fast-evolving field, and much progress has been made on a large variety of topics, providing new insights and methods. Offering a succinct presentation of the mathematical foundations of high-dimensional statistics, this new edition: Offers revised chapters from the previous edition, with the inclusion of many additional materials on some important topics, including compress sensing, estimation with convex constraints, the slope estimator, simultaneously low-rank and row-sparse linear regression, or aggregation of a continuous set of estimators. Introduces three new chapters on iterative algorithms, clustering, and minimax lower bounds. Provides enhanced appendices, minimax lower-bounds mainly with the addition of the Davis-Kahan perturbation bound and of two simple versions of the Hanson-Wright concentration inequality. Covers cutting-edge statistical methods including model selection, sparsity and the Lasso, iterative hard thresholding, aggregation, support vector machines, and learning theory. Provides detailed exercises at the end of every chapter with collaborative solutions on a wiki site. Illustrates concepts with simple but clear practical examples.

Principles and Methods for Data Science 2020-05-28 Principles and Methods for Data Science, Volume 43 in the Handbook of Statistics series, highlights new advances in the field, with this updated volume presenting interesting and timely topics, including Competing risks, aims and methods, Data analysis and mining of microbial community dynamics, Support Vector Machines, a robust prediction method with applications in bioinformatics, Bayesian Model Selection for Data with High Dimension, High dimensional statistical inference: theoretical development to data analytics, Big data challenges in genomics, Analysis of microarray gene expression data using information theory and stochastic algorithm, Hybrid Models, Markov Chain Monte Carlo Methods: Theory and Practice, and more. Provides the authority and expertise of leading contributors from an international board of authors Presents the latest release in the Handbook of Statistics series Updated release includes the latest information on Principles and Methods for Data Science

Analysis of Multivariate and High-Dimensional Data Inge Koch 2013-12-02 This modern approach integrates classical and contemporary methods, fusing theory and practice and bridging the gap to statistical learning.

High-Dimensional Statistics Martin J. Wainwright 2019-02-21 A coherent introductory text from a

groundbreaking researcher, focusing on clarity and motivation to build intuition and understanding. Modern Directional Statistics Christophe Ley 2017-08-03 Modern Directional Statistics collects important advances in methodology and theory for directional statistics over the last two decades. It provides a detailed overview and analysis of recent results that can help both researchers and practitioners. Knowledge of multivariate statistics eases the reading but is not mandatory. The field of directional statistics has received a lot of attention over the past two decades, due to new demands from domains such as life sciences or machine learning, to the availability of massive data sets requiring adapted statistical techniques, and to technological advances. This book covers important progresses in distribution theory, high-dimensional statistics, kernel density estimation, efficient inference on directional supports, and computational and graphical methods. Christophe Ley is professor of mathematical statistics at Ghent University. His research interests include semi-parametrically efficient inference, flexible modeling, directional statistics and the study of asymptotic approximations via Stein's Method. His achievements include the Marie-Jeanne Laurent-Duhamel prize of the Société Française de Statistique and an elected membership at the International Statistical Institute. He is associate editor for the journals Computational Statistics & Data Analysis and Econometrics and Statistics. Thomas Verdebout is professor of mathematical statistics at Université libre de Bruxelles (ULB). His main research interests are semi-parametric statistics, high-dimensional statistics, directional statistics and rank-based procedures. He has won an annual prize of the Belgian Academy of Sciences and is an elected member of the International Statistical Institute. He is associate editor for the journals Statistics and Probability Letters and Journal of Multivariate Analysis.

Contributions in infinite-dimensional statistics and related topics Enea G. Bongiorno 2014-05-21 The interest towards Functional and Operatorial Statistics, and, more in general, towards infinite-dimensional statistics has dramatically increased in the statistical community and in many other applied scientific areas where people face functional data. This volume collects the works selected and presented at the Third Edition of the International Workshop on Functional and Operatorial Statistics held in Stresa, Italy, from the 19th to the 21st of June 2014 (IWFOS'2014). The meeting represents an opportunity of bringing together leading researchers active on these topics both for what concerns theoretical aspects and a wide range of applications in various fields. To promote collaborations with other important strictly related areas of infinite-dimensional Statistics, such as High Dimensional Statistics and Model Selection Procedures, this book hosts works in the latter research subjects too.

Inverse Problems and High-Dimensional Estimation Pierre Alquier 2011-06-07 The "Stats in the Château" summer school was held at the CRC château on the campus of HEC Paris, Jouy-en-Josas, France, from August 31 to September 4, 2009. This event was organized jointly by faculty members of three French academic institutions ? ENSAE ParisTech, the Ecole Polytechnique ParisTech, and HEC Paris ? which cooperate through a scientific foundation devoted to the decision sciences. The scientific content of the summer school was conveyed in two courses, one by Laurent Cavalier (Université Aix-Marseille I) on "Ill-posed Inverse Problems", and one by Victor Chernozhukov (Massachusetts Institute of Technology) on "High-dimensional Estimation with Applications to Economics". Ten invited researchers also presented either reviews of the state of the art in the field or of applications, or original research contributions. This volume contains the lecture notes of the two courses. Original research articles and a survey complement these lecture notes. Applications to economics are discussed in various contributions.

Mathematical Foundations of Infinite-Dimensional Statistical Models

High-Dimensional Covariance Estimation Mohsen Pourahmadi 2013-06-24 "Focusing on methodology and computation more than on theorems and proofs, this book provides computationally feasible and statistically efficient methods for estimating sparse and large covariance matrices of high-dimensional data. Extensive in breadth and scope, it features ample applications to a number of applied areas, including business and economics, computer science, engineering, and financial mathematics; recognizes the important and significant contributions of longitudinal and spatial data; and includes various computer codes in R throughout the text and on an author-maintained web site"--

Modern Multivariate Statistical Techniques Alan J. Izenman 2009-03-02 This is the first book on multivariate analysis to look at large data sets which describes the state of the art in analyzing such data. Material such as database management systems is included that has never appeared in statistics books before.

Foundations of Data Science Avrim Blum 2020-01-31 Covers mathematical and algorithmic foundations of data science: machine learning, high-dimensional geometry, and analysis of large networks.

Multiple Testing Procedures with Applications to Genomics Sandrine Dudoit 2007-12-18 This book

establishes the theoretical foundations of a general methodology for multiple hypothesis testing and discusses its software implementation in R and SAS. These are applied to a range of problems in biomedical and genomic research, including identification of differentially expressed and co-expressed genes in high-throughput gene expression experiments; tests of association between gene expression measures and biological annotation metadata; sequence analysis; and genetic mapping of complex traits using single nucleotide polymorphisms. The procedures are based on a test statistics joint null distribution and provide Type I error control in testing problems involving general data generating distributions, null hypotheses, and test statistics.

High-Dimensional Probability Roman Vershynin 2018-09-27 An integrated package of powerful probabilistic tools and key applications in modern mathematical data science.

Geometric Structure of High-Dimensional Data and Dimensionality Reduction Jianzhong Wang 2012-04-28 "Geometric Structure of High-Dimensional Data and Dimensionality Reduction" adopts data geometry as a framework to address various methods of dimensionality reduction. In addition to the introduction to well-known linear methods, the book moreover stresses the recently developed nonlinear methods and introduces the applications of dimensionality reduction in many areas, such as face recognition, image segmentation, data classification, data visualization, and hyperspectral imagery data analysis. Numerous tables and graphs are included to illustrate the ideas, effects, and shortcomings of the methods. MATLAB code of all dimensionality reduction algorithms is provided to aid the readers with the implementations on computers. The book will be useful for mathematicians, statisticians, computer scientists, and data analysts. It is also a valuable handbook for other practitioners who have a basic background in mathematics, statistics and/or computer algorithms, like internet search engine designers, physicists, geologists, electronic engineers, and economists. Jianzhong Wang is a Professor of Mathematics at Sam Houston State University, U.S.A.

Large Sample Covariance Matrices and High-Dimensional Data Analysis Jianfeng Yao 2015-03-26 High-dimensional data appear in many fields, and their analysis has become increasingly important in modern statistics. However, it has long been observed that several well-known methods in multivariate analysis become inefficient, or even misleading, when the data dimension  $p$  is larger than, say, several tens. A seminal example is the well-known inefficiency of Hotelling's  $T^2$ -test in such cases. This example shows that classical large sample limits may no longer hold for high-dimensional data; statisticians must seek new limiting theorems in these instances. Thus, the theory of random matrices (RMT) serves as a much-needed and welcome alternative framework. Based on the authors' own research, this book provides a first-hand introduction to new high-dimensional statistical methods derived from RMT. The book begins with a detailed introduction to useful tools from RMT, and then presents a series of high-dimensional problems with solutions provided by RMT methods.

High-Dimensional Data Analysis with Low-Dimensional Models John Wright 2021-12-31 Connects fundamental mathematical theory with real-world problems, through efficient and scalable optimization algorithms.