Bayesian Analysis of Rank Data with Covariates and Heterogeneous Rankers
Xinran Li, Dingdong Yi and Jun S. Liu

The Dependent Dirichlet Process and Related Models
Fernando A. Quintana, Peter Müller, Alejandro Jara and Steven N. MacEachern

A Comparative Tour through the Simulation Algorithms for Max-Stable Processes
Marco Oesting and Kirstin Strokorb

Analyzing Stochastic Computer Models: A Review with Opportunities
Evan Baker, Pierre Barbillon, Arindam Fadikar, Robert B. Gramacy, Radu Herbei, David Higdon,
Jiangeng Huang, Leah R. Johnson, Pulong Ma, Anirban Mondal, Bianica Pires,
Jerome Sacks and Vadim Sokolov

Statistical Dependence: Beyond Pearson’s $\rho$
Dag Tjøstheim, Håkon Otneim and Bård Støve

Some Perspectives on Inference in High Dimensions
H. S. Battey and D. R. Cox

Diffusion Smoothing for Spatial Point Patterns
Adrian Baddeley, Tilman M. Davies,
Suman Rakshit, Gopal Nair and Greg McSwiggan

A Conversation with Ross Prentice
Li Hsu and Charles Kooperberg

Statistical Science [ISSN 0883-4237 (print); ISSN 2168-8745 (online)], Volume 37, Number 1, February 2022. Published quarterly by the Institute of Mathematical Statistics, 9760 Smith Road, Waite Hill, Ohio 44094, USA. Periodicals postage paid at Cleveland, Ohio and at additional mailing offices.

POSTMASTER: Send address changes to Statistical Science, Institute of Mathematical Statistics, Dues and Subscriptions Office, PO Box 729, Middletown, Maryland 21769, USA.

Copyright © 2022 by the Institute of Mathematical Statistics
Printed in the United States of America
Bayesian Analysis of Rank Data with Covariates and Heterogeneous Rankers
Xinran Li, Dingdong Yi and Jun S. Liu

The Dependent Dirichlet Process and Related Models
Fernando A. Quintana, Peter Müller, Alejandro Jara and Steven N. MacEachern

A Comparative Tour through the Simulation Algorithms for Max-Stable Processes
Marco Oesting and Kirstin Strokorb

Analyzing Stochastic Computer Models: A Review with Opportunities
Evan Baker, Pierre Barbillon, Arindam Fadikar, Robert B. Gramacy, Radu Herbei, David Higdon, Jiangeng Huang, Leah R. Johnson, Pulong Ma, Anirban Mondal, Bianica Pires, Jerome Sacks and Vadim Sokolov

Statistical Dependence: Beyond Pearson's $\rho$
Dag Tjøstheim, Håkon Otneim and Bård Steine

Some Perspectives on Inference in High Dimensions
H. S. Battey and D. R. Cox

Diffusion Smoothing for Spatial Point Patterns
Adrian Baddeley, Tilman M. Davies, Suman Rakshit, Gopalan Nair and Greg McSwiggan

A Conversation with Ross Prentice
Li Hsu and Charles Kooperberg
Bayesian Analysis of Rank Data with Covariates and Heterogeneous Rankers

Xinran Li, Dingdong Yi and Jun S. Liu

Abstract. Data in the form of ranking lists are frequently encountered, and combining ranking results from different sources can potentially generate a better ranking list and help understand behaviors of the rankers. Of interest here are the rank data under the following settings: (i) covariate information available for the ranked entities; (ii) rankers of varying qualities or having different opinions; and (iii) incomplete ranking lists for nonoverlapping subgroups. We review some key ideas built around the Thurstone model family by researchers in the past few decades and provide a unifying approach for Bayesian Analysis of Rank data with Covariates (BARC) and its extensions in handling heterogeneous rankers. With this Bayesian framework, we can study rankers’ varying quality, cluster rankers’ heterogeneous opinions, and measure the corresponding uncertainties. To enable an efficient Bayesian inference, we advocate a parameter-expanded Gibbs sampler to sample from the target posterior distribution. The posterior samples also result in a Bayesian aggregated ranking list, with credible intervals quantifying its uncertainty. We investigate and compare performances of the proposed methods and other rank aggregation methods in both simulation studies and two real-data examples.

Key words and phrases: Thurstone model, rank aggregation, heterogeneous rankers, infinite mixture model, parameter-expanded data augmentation.

REFERENCES


The Dependent Dirichlet Process and Related Models

Fernando A. Quintana, Peter Müller, Alejandro Jara and Steven N. MacEachern

Abstract. Standard regression approaches assume that some finite number of the response distribution characteristics, such as location and scale, change as a (parametric or nonparametric) function of predictors. However, it is not always appropriate to assume a location/scale representation, where the error distribution has unchanging shape over the predictor space. In fact, it often happens in applied research that the distribution of responses under study changes with predictors in ways that cannot be reasonably represented by a finite dimensional functional form. This can seriously affect the answers to the scientific questions of interest, and therefore more general approaches are indeed needed. This gives rise to the study of fully nonparametric regression models. We review some of the main Bayesian approaches that have been employed to define probability models where the complete response distribution may vary flexibly with predictors. We focus on developments based on modifications of the Dirichlet process, historically termed dependent Dirichlet processes, and some of the extensions that have been proposed to tackle this general problem using nonparametric approaches.

Key words and phrases: Related random probability distributions, Bayesian nonparametrics, nonparametric regression, quantile regression.

REFERENCES


A Comparative Tour through the Simulation Algorithms for Max-Stable Processes

Marco Oesting and Kirstin Strokorb

Abstract. Being the max-analogue of α-stable stochastic processes, max-stable processes form one of the fundamental classes of stochastic processes. With the arrival of sufficient computational capabilities, they have become a benchmark in the analysis of spatiotemporal extreme events. Simulation is often a necessary part of inference of certain characteristics, in particular for future spatial risk assessment. In this article, we give an overview over existing procedures for this task, put them into perspective of one another and use some new theoretical results to make comparisons with respect to their properties.

Key words and phrases: Spectral representation, threshold stopping, extremal functions, error assessment.

REFERENCES


Marco Oesting is Junior Professor of Computational Statistics, Stuttgart Center for Simulation Science & Institute for Stochastics and Applications, University of Stuttgart, 70569 Stuttgart, Germany (e-mail: marco.oesting@mathematik.uni-stuttgart.de). Kirstin Strokorb is Senior Lecturer, School of Mathematics, Cardiff University, Cardiff CF24 4AG, UK (e-mail: strokorbk@cardiff.ac.uk).
Analyzing Stochastic Computer Models: A Review with Opportunities

Evan Baker, Pierre Barbillon, Arindam Fadikar, Robert B. Gramacy, Radu Herbei, David Higdon, Jiangeng Huang, Leah R. Johnson, Pulong Ma, Anirban Mondal, Bianica Pires, Jerome Sacks and Vadim Sokolov

Abstract. In modern science, computer models are often used to understand complex phenomena and a thriving statistical community has grown around analyzing them. This review aims to bring a spotlight to the growing prevalence of stochastic computer models—providing a catalogue of statistical methods for practitioners, an introductory view for statisticians (whether familiar with deterministic computer models or not), and an emphasis on open questions of relevance to practitioners and statisticians. Gaussian process surrogate models take center stage in this review, and these, along with several extensions needed for stochastic settings, are explained. The basic issues of designing a stochastic computer experiment and calibrating a stochastic computer model are prominent in the discussion. Instructive examples, with data and code, are used to describe the implementation of, and results from, various methods.

Key words and phrases: Computer model, Gaussian process, uncertainty quantification, emulator, computer experiment, agent based model, surrogates, calibration.

REFERENCES


ANKENMAN, B., NELSON, B. L. and STAUM, J. (2010). Stochas-


Statistical Dependence: Beyond Pearson’s $\rho$

Dag Tjøstheim, Håkon Otneim and Bård Støve

Abstract. Pearson’s $\rho$ is the most used measure of statistical dependence. It gives a complete characterization of dependence in the Gaussian case, and it also works well in some non-Gaussian situations. It is well known; however, that it has a number of shortcomings; in particular, for heavy tailed distributions and in nonlinear situations, where it may produce misleading, and even disastrous results. In recent years, a number of alternatives have been proposed. In this paper, we will survey these developments, especially results obtained in the last couple of decades. Among measures discussed are the copula, distribution-based measures, the distance covariance, the HSIC measure popular in machine learning and finally the local Gaussian correlation, which is a local version of Pearson’s $\rho$. Throughout, we put the emphasis on conceptual developments and a comparison of these. We point out relevant references to technical details as well as comparative empirical and simulated experiments. There is a broad selection of references under each topic treated.

Key words and phrases: Statistical dependence, Pearson’s $\rho$, nonlinear dependence, distance covariance, HSIC, mutual information, local Gaussian correlation.

REFERENCES


Dag Tjøstheim is Professor Emeritus, Department of Mathematics, University of Bergen, P.b. 7803, 5020 Bergen, Norway (e-mail: dag.tjoestheim@uib.no). Håkon Otneim is Associate Professor, Department of Business and Management Science, Norwegian School of Economics, Helleveien 30, 5045 Bergen, Norway (e-mail: hakkon.otneim@nhh.no). Bård Støve is Professor, Department of Mathematics, University of Bergen, P.b. 7803, 5020 Bergen, Norway (e-mail: bard.stove@uib.no).
Some Perspectives on Inference in High Dimensions

H. S. Battey and D. R. Cox

Abstract. With very large amounts of data, important aspects of statistical analysis may appear largely descriptive in that the role of probability sometimes seems limited or totally absent. The main emphasis of the present paper lies on contexts where formulation in terms of a probabilistic model is feasible and fruitful but to be at all realistic large numbers of unknown parameters need consideration. Then many of the standard approaches to statistical analysis, for instance direct application of the method of maximum likelihood, or the use of flat priors, often encounter difficulties. After a brief discussion of broad conceptual issues, we provide some new perspectives on aspects of high-dimensional statistical theory, emphasizing a number of open problems.

Key words and phrases: Inference, likelihood, model uncertainty, nuisance parameters, parameter orthogonalization, sparsity.

REFERENCES


H. S. Battey is Assistant Professor, Department of Mathematics, Imperial College London, SW7 2AZ, UK (e-mail: h.battey@imperial.ac.uk). D. R. Cox is Honorary Fellow, Nuffield College, University of Oxford OX1 1NF, UK.


Diffusion Smoothing for Spatial Point Patterns

Adrian Baddeley, Tilman M. Davies, Suman Rakshit, Gopalan Nair and Greg McSwiggan

Abstract. Traditional kernel methods for estimating the spatially-varying density of points in a spatial point pattern may exhibit unrealistic artefacts, in addition to the familiar problems of bias and over- or under-smoothing. Performance can be improved by using diffusion smoothing, in which the smoothing kernel is the heat kernel on the spatial domain. This paper develops diffusion smoothing into a practical statistical methodology for two-dimensional spatial point pattern data. We clarify the advantages and disadvantages of diffusion smoothing over Gaussian kernel smoothing. Adaptive smoothing, where the smoothing bandwidth is spatially-varying, can be performed by adopting a spatially-varying diffusion rate: this avoids technical problems with adaptive Gaussian smoothing and has substantially better performance. We introduce a new form of adaptive smoothing using lagged arrival times, which has good performance and improved robustness. Applications in archaeology and epidemiology are demonstrated. The methods are implemented in open-source R code.

Key words and phrases: Adaptive smoothing, bandwidth, heat kernel, kernel estimation, lagged arrival method, Richardson extrapolation.

REFERENCES


Adrian Baddeley is John Curtin Distinguished Professor, School of Electrical Engineering, Computing and Mathematical Sciences, Curtin University, GPO Box U1987, Perth WA 6845, Australia (e-mail: Adrian.Baddeley@curtin.edu.au). Tilman M. Davies is Senior Lecturer, Department of Mathematics and Statistics, University of Otago, PO Box 56, Dunedin 9054, New Zealand (e-mail: tilman.davies@otago.ac.nz). Suman Rakshit is Lecturer, School of Electrical Engineering, Computing and Mathematical Sciences, and Research Fellow, SAGI-West, Curtin University, GPO Box U1987, Perth WA 6845, Australia (e-mail: Suman.Rakshit@curtin.edu.au). Gopalan Nair is Senior Lecturer, Department of Mathematics and Statistics, University of Western Australia, 35 Stirling Hwy, Nedlands WA 6009, Australia (e-mail: gopalan.nair@uwa.edu.au). Greg McSwiggan is PhD graduate, Department of Mathematics and Statistics, University of Western Australia, and professional consulting engineer, PO Box 2697 New Farm, Queensland 4005, Australia (e-mail: qfengineers@gmail.com).


A Conversation with Ross Prentice

Li Hsu and Charles Kooperberg

Abstract. Ross L. Prentice received his B.Sc. from the University of Waterloo and his Ph.D. from the University of Toronto. He joined the University of Washington (UW) and the Fred Hutchinson Cancer Research Center (the Hutch) in 1974, and is currently Professor of Biostatistics at these institutions. He was Senior Vice President at the Hutch, and Director of its Public Health Sciences Division, for more than 25 years.

Dr. Prentice’s expertise and research interests are in the fields of biostatistics, epidemiology, and disease prevention. He played a central role in the conception, design, and implementation of the Women’s Health Initiative. In statistical and medical literature he has over 500 scientific papers, including more than 40 with 500 or more citations. His substantial contributions to the theory of population and clinical research include the use of surrogate endpoints and case-cohort designs and other areas such as survival analysis, nutritional epidemiology, genetic epidemiology, biomarkers, and measurement error. Dr. Prentice is recognized for his mentoring of students and junior colleagues, and for his generous collaborations.

Dr. Prentice has received numerous awards for his work, including an honorary doctorate in mathematics from the University of Waterloo, the Mantel Award for Lifetime Contributions to Statistics in Epidemiology from the American Statistical Association, the Mortimer Spiegelman Award from the American Public Health Association, the Committee of Presidents of Statistical Societies Presidents’ Award and RA Fisher Award, the Marvin Zelen Leadership Award for Outstanding Achievement in Statistical Science from Harvard University, the American Association of Cancer Research/American Cancer Society Award for Research Excellence in Cancer Epidemiology and Prevention, and the American Association for Cancer Research Team Science Award. He was elected to the Institute of Medicine/National Academy of Medicine in 1990. The Ross L. Prentice Endowed Professorship of Biostatistical Collaboration was created at the UW in 2005 and has been awarded every year since its inception. The interior space of the Public Health Sciences building at the Hutch has been named the Ross L. Prentice Atrium.

In his spare time, Ross enjoys sports including water skiing, golf, running, and spending time with his wife, Didi, and with his daughters, sons-in-law, and grandchildren. He ran daily from when he was in his 20s until his knees objected about 10 years ago.

This interview took place with Li Hsu and Charles Kooperberg via Zoom in December 2020.

Key words and phrases: Fred Hutchinson Cancer Research Center, University of Washington, University of Waterloo, survival analysis, nutritional epidemiology, Women’s Health Initiative.
INSTITUTE OF MATHEMATICAL STATISTICS

(Organized September 12, 1935)

The purpose of the Institute is to foster the development and dissemination of the theory and applications of statistics and probability.

IMS OFFICERS

President: Krzysztof Burdzy, Department of Mathematics, University of Washington, Seattle, Washington 98195-4350, USA

President-Elect: Peter Bühlmann, Seminar für Statistik, ETH Zürich, 8092 Zürich, Switzerland

Past President: Regina Y. Liu, Department of Statistics, Rutgers University, Piscataway, New Jersey 08854-8019, USA

Executive Secretary: Edsel Peña, Department of Statistics, University of South Carolina, Columbia, South Carolina 29208-001, USA

Treasurer: Zhengjun Zhang, Department of Statistics, University of Wisconsin, Madison, Wisconsin 53706-1510, USA

Program Secretary: Annie Qu, Department of Statistics, University of California, Irvine, Irvine, CA 92697-3425, USA

IMS EDITORS

The Annals of Statistics. Editors: Enno Mammen, Institute for Applied Mathematics, Heidelberg University, 69120 Heidelberg, Germany. Lan Wang, Miami Herbert Business School, University of Miami, Coral Gables, FL 33124, USA

The Annals of Applied Statistics. Editor-in-Chief: Ji Zhu, Department of Statistics, University of Michigan, Ann Arbor, MI 48109, USA

The Annals of Probability. Editors: Alice Guionnet, Unité de Mathématiques Pures et Appliquées, ENS de Lyon, Lyon, France. Christophe Garban, Institut Camille Jordan, Université Claude Bernard Lyon 1, 69622 Villeurbanne, France

The Annals of Applied Probability. Editors: Qi-Man Shao, Department of Statistics and Data Science, Southern University of Science and Technology Shenzhen, Guangdong 518055, P.R. China. Kavita Ramanan, Division of Applied Mathematics, Brown University, Providence, RI 02912 USA

Statistical Science. Editor: Sonia Petrone, Department of Decision Sciences, Università Bocconi, 20100 Milano MI, Italy

The IMS Bulletin. Editor: Vlada Limic, UMR 7501 de l’Université de Strasbourg et du CNRS, 7 rue René Descartes, 67084 Strasbourg Cedex, France
BRADLEY EFRON
TREVOR HASTIE

COMPUTER AGE
STATISTICAL
INFERENCE

ALGORITHMS, EVIDENCE, AND DATA SCIENCE