The Role of Exchangeability in Causal Inference
................................. Olli Saarela, David A. Stephens and Erica E. M. Moodie 369
Aitchison’s Compositional Data Analysis 40 Years on: A Reappraisal
................. Michael Greenacre, Éric Grunsky, John Bacon-Shone, Jonas Erb and Thomas Quinn 386
Statistical Embedding: Beyond Principal Components
................................. Dag Tjøstheim, Martin Jullum and Anders Løland 411
Can We Reliably Detect Biases that Matter in Observational Studies? ... Paul R. Rosenbaum 440
Experimental Design in Marketplaces .... Patrick Bajari, Brian Burdick, Guido W. Imbens,
Lorenzo Masoero, James McQueen, Thomas S. Richardson and Ido M. Rosen 458
Parameter Restrictions for the Sake of Identification: Is There Utility in Asserting That
Perhaps a Restriction Holds? ......................... Paul Gustafson 477
Variational Inference for Cutting Feedback in Misspecified Models
................................. Xuejun Yu, David J. Nott and Michael Stanley Smith 490
Note on Legendre's Method of Least Squares ......................... Jukka Nyblom 510
A Conversation with Mary E. Thompson ......................... Rhonda J. Rosychuk 514
The Role of Exchangeability in Causal Inference
Olli Saarela, David A. Stephens and Erica E. M. Moodie

Aitchison’s Compositional Data Analysis 40 Years on: A Reappraisal
Michael Greenacre, Eric Grunsky, John Bacon-Shone, Ionas Erb and Thomas Quinn

Statistical Embedding: Beyond Principal Components
Dag Tjøstheim, Martin Jullum and Anders Løland

Can We Reliably Detect Biases that Matter in Observational Studies?
Paul R. Rosenbaum

Experimental Design in Marketplaces
Patrick Bajari, Brian Burdick, Guido W. Imbens, Lorenzo Masoero, James McQueen, Thomas S. Richardson and Ido M. Rosen

Parameter Restrictions for the Sake of Identification: Is There Utility in Asserting That Perhaps a Restriction Holds?
Paul Gustafson

Variational Inference for Cutting Feedback in Misspecified Models
Xuejun Yu, David J. Nott and Michael Stanley Smith

Note on Legendre’s Method of Least Squares
Jukka Nyblom

A Conversation with Mary E. Thompson
Rhonda J. Rosychuk
EDITOR
Moulinath Banerjee
University of Michigan

ASSOCIATE EDITORS
Shankar Bhamidi
University of North Carolina
Jay Breidt
University of Chicago
Alicia Carriquiry
Iowa State University
Matias D. Cattaneo
Princeton University
Nilanjan Chatterjee
Johns Hopkins University
Yang Chen
University of Michigan
Bertrand Clarke
University of Nebraska-Lincoln
Michael J. Daniels
University of Florida
Philip Dawid
University of Cambridge
Holger Dette
Ruhr-Universitat Bochum
Robin Evans
University of Oxford
Stefano Favaro
Università di Torino
Subhashis Ghoshal
North Carolina State University
Peter Green
University of Bristol and University of Technology Sydney
Chris Holmes
University of Oxford
Tailen Hsing
University of Michigan
Po-Ling Loh
University of Cambridge
Ian McKeage
Columbia University
George Michailidis
University of California, Los Angeles
Peter Müller
University of Texas
Axel Munk
Georg-August-University of Goettingen
Jean Opsomer
Westat

MANAGING EDITOR
Dan Nordman
Iowa State University

PRODUCTION EDITOR
Patrick Kelly

EDITORIAL COORDINATOR
Kristina Mattson

PAST EXECUTIVE EDITORS
George Casella, 2002–2004
David Madigan, 2008–2010
Jon A. Wellner, 2011–2013
Leon J. Gleser, 1998–2000
Peter Green, 2014–2016
Richard Tweedie, 2001
Cun-Hui Zhang, 2017–2019
Morris Eaton, 2001
Sonia Petrone, 2020–2022
The Role of Exchangeability in Causal Inference

Olli Saarela, David A. Stephens and Erica E. M. Moodie

Abstract. Though the notion of exchangeability has been discussed in the causal inference literature under various guises, it has rarely taken its original meaning as a symmetry property of probability distributions. As this property is a standard component of Bayesian inference, we argue that in Bayesian causal inference it is natural to link the causal model, including the notion of confounding and definition of causal contrasts of interest, to the concept of exchangeability. Here, we propose a probabilistic between-group exchangeability property as an identifying condition for causal effects, relate it to alternative conditions for unconfounded inferences (commonly stated using potential outcomes) and define causal contrasts in the presence of exchangeability in terms of posterior predictive expectations for further exchangeable units. While our main focus is on a point treatment setting, we also investigate how this reasoning carries over to longitudinal settings.

Key words and phrases: Bayesian inference, causal inference, confounding, exchangeability, posterior predictive inference.

REFERENCES


Olli Saarela is Associate Professor, Dalla Lana School of Public Health, University of Toronto, 155 College Street, Toronto, Ontario M5T 3M7, Canada (e-mail: olli.saarela@utoronto.ca). David A. Stephens is Professor, Department of Mathematics and Statistics, McGill University, Burnside Hall, 805 Sherbrooke Street West, Montreal, Quebec H3A 0B9, Canada (e-mail: d.stephens@math.mcgill.ca). Erica E. M. Moodie is Professor, Department of Epidemiology and Biostatistics, McGill University, 2001 McGill College Ave, Montreal, Quebec H3A 1G1, Canada (e-mail: erica.moodie@mcgill.ca).


Aitchison’s Compositional Data Analysis 40 Years on: A Reappraisal

Michael Greenacre®, Eric Grunsky®, John Bacon-Shone®, Ionas Erb® and Thomas Quinn®

Abstract. The development of John Aitchison’s approach to compositional data analysis is followed since his paper read to the Royal Statistical Society in 1982. Aitchison’s logratio approach, which was proposed to solve the problematic aspects of working with data with a fixed-sum constraint, is summarized and reappraised. It is maintained that the properties on which this approach was originally built, the main one being subcompositional coherence, are not required to be satisfied exactly—quasi-coherence is sufficient, that is near enough to being coherent for all practical purposes. This opens up the field to using simpler data transformations, such as power transformations, that permit zero values in the data. The additional property of exact isometry, which was subsequently introduced and not in Aitchison’s original conception, imposed the use of isometric logratio transformations, but these are complicated and problematic to interpret, involving ratios of geometric means. If this property is regarded as important in certain analytical contexts, for example, unsupervised learning, it can be relaxed by showing that regular pairwise logratios, as well as the alternative quasi-coherent transformations, can also be quasi-isometric, meaning they are close enough to exact isometry for all practical purposes. It is concluded that the isometric and related logratio transformations such as pivot logratios are not a prerequisite for good practice, although many authors insist on their obligatory use. This conclusion is fully supported here by case studies in geochemistry and in genomics, where the good performance is demonstrated of pairwise logratios, as originally proposed by Aitchison, or Box–Cox power transforms of the original compositions where no zero replacements are necessary.

Key words and phrases: Box–Cox transformation, compositional modeling, correspondence analysis, isometry, logratio transformations, log-contrast, principal component analysis, Procrustes analysis, subcompositional coherence.

REFERENCES


Michael Greenacre is Professor, Department of Economics & Business, Universitat Pompeu Fabra, and Barcelona School of Management, Barcelona, Spain (e-mail: michael.greenacre@upf.edu). Eric Grunsky is Adjunct Professor, Department of Earth & Environmental Sciences, University of Waterloo, Canada (e-mail: egrunsky@gmail.com). John Bacon-Shone is Honorary Professor, Faculty of Social Sciences, Hong Kong University, Hong Kong (e-mail: johnbs@hku.hk). Ionas Erb is Researcher, Centre for Genomic Regulation (CRG), The Barcelona Institute of Science and Technology, Barcelona, Spain (e-mail: ionas.erb@crg.eu). Thomas Quinn is Researcher, Applied Artificial Intelligence Institution (A2I2), Deakin University, Australia (e-mail: contacttomquinn@gmail.com).


Abstract. There has been an intense recent activity in embedding of very high-dimensional and nonlinear data structures, much of it in the data science and machine learning literature. We survey this activity in four parts. In the first part, we cover nonlinear methods such as principal curves, multidimensional scaling, local linear methods, ISOMAP, graph-based methods and diffusion mapping, kernel based methods and random projections. The second part is concerned with topological embedding methods, in particular mapping topological properties into persistence diagrams and the Mapper algorithm. Another type of data sets with a tremendous growth is very high-dimensional network data. The task considered in part three is how to embed such data in a vector space of moderate dimension to make the data amenable to traditional techniques such as cluster and classification techniques. Arguably, this is the part where the contrast between algorithmic machine learning methods and statistical modeling, represented by the so-called stochastic block model, is at its greatest. In the paper, we discuss the pros and cons for the two approaches. The final part of the survey deals with embedding in $\mathbb{R}^2$, that is, visualization. Three methods are presented: $t$-SNE, UMAP and LargeVis based on methods in parts one, two and three, respectively. The methods are illustrated and compared on two simulated data sets; one consisting of a triplet of noisy Ranunculoid curves, and one consisting of networks of increasing complexity generated with stochastic block models and with two types of nodes.

Key words and phrases: Statistical embedding, principal component, nonlinear principal component, multidimensional scaling, local linear method, ISOMAP, graph spectral theory, diffusion mapping, reproducing kernel Hilbert space, random projection, topological data analysis and embedding, persistent homology, persistence diagram, the Mapper, network embedding, spectral embedding, stochastic block modeling, Skip-Gram, neighborhood sampling strategies, visualization, $t$-SNE, LargeVis, UMAP.

REFERENCES


Dag Tjøstheim is Professor Emeritus at the Department of Mathematics, University of Bergen, Bergen, Norway and Professor II at the Norwegian Computing Center, Oslo, Norway (e-mail: Dag.Tjostheim@uib.no). Martin Jullum is Senior Research Scientist at the Norwegian Computing Center, Oslo, Norway (e-mail: Martin.Jullum@nr.no). Anders Løland is Research Director at the Norwegian Computing Center, Oslo, Norway (e-mail: Anders.Loland@nr.no).


Can We Reliably Detect Biases that Matter in Observational Studies?

Paul R. Rosenbaum

Abstract. In an observational study of the effects caused by a treatment, biases from unmeasured covariates remain a concern even after successful adjustments for measured covariates. This concern is partly addressed by demonstrating that the qualitative conclusions of the primary analysis would not be altered by small or moderate biases—that these conclusions are insensitive to small or moderate bias. Additionally, the concern is partly addressed by collecting additional information, such as outcomes known to be unaffected by the treatment, and using this information as a test of various biases. Is there a gap between these two activities? Perhaps the study is insensitive to small biases, and we can detect large biases, but the study is sensitive to moderate biases that cannot be detected—that is an informal description of a gap. The concept of ”no gap” is defined formally in Definition 3.1, and the probability of ”no gap” is determined under various sampling situations. When there is no gap, ask: Are causal conclusions measurably strengthened? If so, by how much? The answer depends upon the covering design sensitivity, \( \Gamma \), defined to be the smallest bias that can explain both the ostensible effect of the treatment on the primary outcome and the evidence of bias provided by the unaffected outcome. The covering design sensitivity is calculated in various contexts. A small observational study of the effects of light alcohol consumption on HDL cholesterol is used to illustrate ideas and methods.

Key words and phrases: Causal inference, detecting bias, observational study, sensitivity analysis.

REFERENCES


Experimental Design in Marketplaces

Patrick Bajari, Brian Burdick, Guido W. Imbens, Lorenzo Masoero, James McQueen, Thomas S. Richardson and Ido M. Rosen

Abstract. Classical Randomized Controlled Trials (RCTs), or A/B tests, are designed to draw causal inferences about a population of units, for example, individuals, plots of land or visits to a website. A key assumption underlying a standard RCT is the absence of interactions between units, or the stable unit treatment value assumption (Ann. Statist. 6 (1978) 34–58). Modern experimentation, however, is often conducted in settings characterized by complex interactions between units. Such interactions can invalidate the standard estimators and make classical experimental designs ineffective. Although the presence of interference forces us to make untestable assumptions on the nature of the interactions even under randomization, sophisticated experimental designs can ameliorate the dependence on such assumptions. In this manuscript, we review the recent and rapidly growing literature on novel experimental designs for these settings. One key feature common to many of these designs is the presence of multiple layers of randomization within the same experiment. We discuss a novel experimental design, called Multiple Randomization Designs or MRDs, that provides a general framework for such experiments. Through these complex designs, we can study questions about causal effects in the presence of interference that cannot be answered by classical RCTs.

Key words and phrases: Experimental design, causal inference, online experimentation, multiple randomization designs, two-sided marketplaces.

REFERENCES


Patrick Bajari is Vice President, Amazon, Seattle, WA 98109, USA (e-mail: patrickbajari@gmail.com). Brian Burdick was Director of Research at Core-AI at Amazon while doing this work. Guido W. Imbens is Professor of Economics, Graduate School of Business and Department of Economics, Stanford University, SIEPR, NBER, Stanford, California 94305, USA (e-mail: imbens@stanford.edu). Lorenzo Masoero is Research Scientist, Amazon, Seattle, WA 98109, USA (e-mail: masoerl@amazon.com). James McQueen is Principal Scientist, Amazon, Seattle, WA 98109, USA (e-mail: jmcq@amazon.com). Thomas S. Richardson is Professor of Statistics, University of Washington, Seattle, WA 98195, USA (e-mail: thomasr@u.washington.edu). Ido M. Rosen is Sr Principal Scientist, Core AI, Amazon, Seattle, WA 98109, USA (e-mail: ido@uchicago.edu).
Parameter Restrictions for the Sake of Identification: Is There Utility in Asserting That Perhaps a Restriction Holds?

Paul Gustafson

Abstract. Statistical modeling can involve a tension between assumptions and statistical identification. The law of the observable data may not uniquely determine the value of a target parameter without invoking a key assumption, and, while plausible, this assumption may not be obviously true in the scientific context at hand. Moreover, there are many instances of key assumptions which are untestable, hence we cannot rely on the data to resolve the question of whether the target is legitimately identified. Working in the Bayesian paradigm, we consider the grey zone of situations where a key assumption, in the form of a parameter space restriction, is scientifically reasonable but not incontrovertible for the problem being tackled. Specifically, we investigate statistical properties that ensue if we structure a prior distribution to assert that maybe or perhaps the assumption holds. Technically this simply devolves to using a mixture prior distribution putting just some prior weight on the assumption, or one of several assumptions, holding. However, while the construct is straightforward, there is very little literature discussing situations where Bayesian model averaging is employed across a mix of fully identified and partially identified models.

Key words and phrases: Bayesian model averaging, Bayes risk, large-sample theory, partial identification.

REFERENCES


Paul Gustafson is Professor, Department of Statistics, University of British Columbia, Vancouver, Canada (e-mail: gustaf@stat.ubc.ca).


Variational Inference for Cutting Feedback in Misspecified Models

Xuejun Yu, David J. Nott and Michael Stanley Smith

Abstract. Bayesian analyses combine information represented by different terms in a joint Bayesian model. When one or more of the terms is misspecified, it can be helpful to restrict the use of information from suspect model components to modify posterior inference. This is called “cutting feedback”, and both the specification and computation of the posterior for such “cut models” is challenging. In this paper, we define cut posterior distributions as solutions to constrained optimization problems, and propose variational methods for their computation. These methods are faster than existing Markov chain Monte Carlo (MCMC) approaches by an order of magnitude. It is also shown that variational methods allow for the evaluation of computationally intensive conflict checks that can be used to decide whether or not feedback should be cut. Our methods are illustrated in a number of simulated and real examples, including an application where recent methodological advances that combine variational inference and MCMC within the variational optimization are used.

Key words and phrases: Bayesian model criticism, cutting feedback, model misspecification, modular inference.

REFERENCES


Note on Legendre’s Method of Least Squares

Jukka Nyblom

Abstract. In the first published treatment of the least squares, Legendre applied his new method to the French meridian data measured for the determination of the length of the meter (Nouvelles méthodes pour la détermination des orbites des comètes (1805) 76 Firmin Didot). Legendre treated one error term as a constant. It is shown here that it turns out to be equivalent to the generalized least squares solution of his model (Nouvelles méthodes pour la détermination des orbites des comètes (1805) 77 Firmin Didot).

Key words and phrases: French meridian data, meridian arc length, generalized least squares, flattening of the Earth, determination of meter.

REFERENCES

Euler, L. P. (1755). Élémens de la Trigonométrie sphéroïdique tirés de la méthode des plus grandes et plus petits. Mémoires de Berlin 1753 IX 258–293.


A Conversation with Mary E. Thompson

Rhonda J. Rosychuk

Abstract. Mary E. Thompson (née Beattie) was born September 9, 1944, in Winnipeg, Manitoba, Canada. She obtained a B.Sc. in Mathematics from the University of Toronto in 1965, and earned M.Sc. (1966) and Ph.D. (1969) degrees in Mathematics from the University of Illinois at Urbana-Champaign. She joined the Department of Statistics at the University of Waterloo as a Lecturer in 1969 and became an Assistant Professor in 1971. In 2004, she was awarded the honour of University Professor and in 2011 became Distinguished Professor Emerita at the University of Waterloo. She has served in many leadership roles including Chair of the Department of Statistics and Actuarial Science, Acting Dean of the Faculty of Mathematics, President of the Statistical Society of Canada (SSC) and Chair of the COPSS Presidents’ Award Committee. She chaired the Development Committee for the Canadian Statistical Sciences Institute (CANSSI) and was its founding Scientific Director.

Thompson has received numerous honours and awards including the SSC’s Gold Medal, the Elizabeth L. Scott Award, the Waksberg Award of Survey Methodology and the Governor General’s Innovation Award. She is an elected member of the International Statistical Institute, an Honorary Member of the SSC and is a Fellow of the American Statistical Association, the Institute of Mathematical Statistics, the Royal Society of Canada and the Fields Institute.

Thompson has made fundamental contributions to several areas in statistics including sampling theory and the analysis of surveys. She is the author of two books in these areas: Theory of Sample Surveys (1997) and Sampling Theory and Practices (2020 with C. Wu). She has also made key contributions in estimation theory and stochastic processes. As the author of over 150 published, refereed papers, Thompson has influenced the theory and practice of statistics.

The following conversation took place virtually in September 2022 with interviewer Rhonda J. Rosychuk of the University of Alberta.

Key words and phrases: University of Waterloo, Statistical Society of Canada, Canadian Statistical Sciences Institute, sampling theory, survey methodology, estimation theory, stochastic processes.

REFERENCES


Rhonda J. Rosychuk is Professor, Department of Pediatrics, University of Alberta, 11405 87 Avenue NW, Edmonton, Alberta, Canada T6G 1C9 (e-mail: rhonda.rosychuk@ualberta.ca).


The purpose of the Institute is to foster the development and dissemination of the theory and applications of statistics and probability.
Exponential Families in Theory and Practice

Bradley Efron, Stanford University

During the past half-century, exponential families have attained a position at the center of parametric statistical inference. Theoretical advances have been matched, and more than matched, in the world of applications, where logistic regression by itself has become the go-to methodology in medical statistics, computer-based prediction algorithms, and the social sciences. This book is based on a one-semester graduate course for first year Ph.D. and advanced master’s students. After presenting the basic structure of univariate and multivariate exponential families, their application to generalized linear models including logistic and Poisson regression is described in detail, emphasizing geometrical ideas, computational practice, and the analogy with ordinary linear regression. Connections are made with a variety of current statistical methodologies: missing data, survival analysis and proportional hazards, false discovery rates, bootstrapping, and empirical Bayes analysis. The book connects exponential family theory with its applications in a way that doesn’t require advanced mathematical preparation.

Hardback $105.00
Paperback $39.99
IMS members are entitled to a 40% discount: email  imstat.org
to request your code

www.imstat.org/cup/

Cambridge University Press, with the Institute of Mathematical Statistics, established the IMS Monographs and IMS Textbooks series of high-quality books. The series editors are Mark Handcock, Ramon van Handel, Arnaud Doucet, and John Aston.