STATISTICAL SCIENCE

Volume 38, Number 3 August 2023

Statistical Science **[ISSN 0883-4237 (print); ISSN 2168-8745 (online)]**, Volume 38, Number 3, August 2023. 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.

Statistical Science **Volume 38, Number 3 (369–524) August 2023**

Volume 38 **Number 3** **August 2023**

Statistical Science
2023, Vol. 38, No. 3, 369–385
https://doi.org/10.1214/22-STS879

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

ARJAS, E. (2012). Causal inference from observational data: A Bayesian predictive approach. In *Causality*: *Statistical Perspectives and Applications* (C. Berzuini, A. P. Dawid and L. Bernardinelli, eds.) 71–84. Wiley, NY.

ARJAS, E. and PARNER, J. (2004). Causal reasoning from longitudinal data. *Scand. J. Stat.* **31** 171–187. MR2066247 https://doi.org/10.1111/j.1467-9469.2004.02-134.x

BAKER, S. G. (2013). Causal inference, probability theory, and graphical insights. *Stat. Med.* **32** 4319–4330. MR3118357 https://doi.org/10.1002/sim.5828

BERNARDO, J.-M. and SMITH, A. F. M. (1994). *Bayesian Theory*. *Wiley Series in Probability and Mathematical Statistics*: *Probability and Mathematical Statistics*. Wiley, Chichester. MR1274699 https://doi.org/10.1002/9780470316870

BIJLSMA, M. J., TARKIAINEN, L., MYRSKYLÄ, M. and MARTIKAINEN, P. (2017). Unemployment and subsequent depression: A mediation analysis using the parametric G-formula. *Soc. Sci. Med.* **194** 142–150.

BÜHLMANN, P. (2020). Invariance, causality and robustness: 2018 Neyman Lecture. *Statist. Sci.* **35** 404–426. MR4148216 https://doi.org/10.1214/19-STS721

CHAKRABORTY, B. and MURPHY, S. A. (2014). Dynamic treatment regimes. *Annu. Rev. Stat. Appl.* **1** 447–464.

CHIB, S. (2007). Analysis of treatment response data without the joint distribution of potential outcomes. *J. Econometrics* **140** 401–412. MR2408912 https://doi.org/10.1016/j.jeconom.2006.07.009

COLE, S. R. and FRANGAKIS, C. E. (2009). The consistency statement in causal inference: A definition or an assumption? *Epidemiology* **20** 3–5.

COMMENGES, D. (2019). Causality without potential outcomes and the dynamic approach. Preprint. Available at arXiv:1905.01195.

COMMENGES, D. and GÉGOUT-PETIT, A. (2015). The stochastic system approach for estimating dynamic treatments effect. *Lifetime Data Anal.* **21** 561–578. MR3397506 https://doi.org/10.1007/s10985-015-9322-3

DAWID, A. P. (2000). Causal inference without counterfactuals. *J. Amer. Statist. Assoc.* **95** 407–448. MR1803167 https://doi.org/10.2307/2669377

DAWID, P. (2021). Decision-theoretic foundations for statistical causality. *J. Causal Inference* **9** 39–77. MR4289525 https://doi.org/10.1515/jci-2020-0008

DAWID, A. P. and DIDELEZ, V. (2010). Identifying the consequences of dynamic treatment strategies: A decision-theoretic overview. *Stat. Surv.* **4** 184–231. MR2740837 https://doi.org/10.1214/10-SS081

DAWID, A. P., MUSIO, M. and FIENBERG, S. E. (2016). From statistical evidence to evidence of causality. *Bayesian Anal.* **11** 725–752. MR3498044 https://doi.org/10.1214/15-BA968

DE FINETTI, B. (1929). Funzione caratteristica di un fenomeno aleatorio. In *Atti del Congresso Internazionale dei Matematici*: *Bologna del* 3 *al* 10 *de Settembre di* 1928 179–190.

DE FINETTI, B. (1938). Sur la condition d'équivalence partielle. *Actual. Sci. Ind.* **739**. Translated In: Studies in Inductive and Proba-

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).

bility, II. Jeffrey, R. (ed.) University of California Press: Berkeley 1980.

DIACONIS, P. (1988). Recent progress on de Finetti's notions of exchangeability. In *Bayesian Statistics*, 3 (*Valencia*, 1987). *Oxford Sci. Publ.* 111–125. Oxford Univ. Press, New York. MR1008047

FERREIRA, J. A. (2015). Some models and methods for the analysis of observational data. *Stat. Surv.* **9** 106–208. MR3396384 https://doi.org/10.1214/15-SS110

FERREIRA, J. A. (2019). Causality from the point of view of statistics. Preprint. Available at arXiv:1908.07301.

GELMAN, A., CARLIN, J. B., STERN, H. S. and RUBIN, D. B. (2004). *Bayesian Data Analysis*, 2nd ed. *Texts in Statistical Science Series*. CRC Press/CRC, Boca Raton, FL. MR2027492

GREENLAND, S. (2003). Quantifying biases in causal models: Classical confounding vs collider-stratification bias. *Epidemiology* **14** 300–306.

GREENLAND, S. (2012). Causal inference as a prediction problem: Assumptions, identification and evidence synthesis. In *Causality: Statistical Perspectives and Applications* (C. Berzuini, A. P. Dawid and L. Bernardinelli, eds.) 43–58. Wiley, NY.

GREENLAND, S. and ROBINS, J. M. (1986). Identifiability, exchangeability, and epidemiological confounding. *Int. J. Epidemiol.* **15** 413–419. https://doi.org/10.1093/ije/15.3.413

GREENLAND, S. and ROBINS, J. M. (2009). Identifiability, exchangeability, and epidemiological confounding revisited. *Epidemiol. Perspect. Innov.* **6**. https://doi.org/10.1186/1742-5573-6-4.

GREENLAND, S., ROBINS, J. M. and PEARL, J. (1999). Confounding and collapsibility in causal inference. *Statist. Sci.* **14** 29–46.

HERNÁN, M. A., BRUMBACK, B. and ROBINS, J. M. (2001). Marginal structural models to estimate the joint causal effect of nonrandomized treatments. *J. Amer. Statist. Assoc.* **96** 440–448. MR1939347 https://doi.org/10.1198/016214501753168154

HERNÁN, M. A. and ROBINS, J. M. (2006). Estimating causal effects from epidemiological data. *J. Epidemiol. Community Health* **60** 578–586.

HEWITT, E. and SAVAGE, L. J. (1955). Symmetric measures on Cartesian products. *Trans. Amer. Math. Soc.* **80** 470–501. MR0076206 https://doi.org/10.2307/1992999

HOLLAND, P. W. (1986). Statistics and causal inference. *J. Amer. Statist. Assoc.* **81** 945–970. MR0867618

JAIN, P., DANAEI, G., ROBINS, J. M., MANSON, J. E. and HERNÁN, M. A. (2016). Smoking cessation and long-term weight gain in the Framingham Heart Study: An application of the parametric g-formula for a continuous outcome. *Eur. J. Epidemiol.* **31** 1223–1229. https://doi.org/10.1007/s10654-016-0200-4

KEIDING, N. and CLAYTON, D. (2014). Standardization and control for confounding in observational studies: A historical perspective. *Statist. Sci.* **29** 529–558. MR3300358 https://doi.org/10.1214/13-STS453

KEIL, A. P., EDWARDS, J. K., RICHARDSON, D. R., NAIMI, A. I. and COLE, S. R. (2014). The parametric g-formula for time-to-event data: Towards intuition with a worked example. *Epidemiology* **25** 889.

LAURITZEN, S. L., ANDERSEN, A. H., EDWARDS, D., JÖRESKOG, K. G. and JOHANSEN, S. (1989). Mixed graphical association models [with discussion and rejoinder]. *Scand. J. Stat.* **16** 273–306.

LINDLEY, D. V. (2002). Seeing and doing: The concept of causation. *Int. Stat. Rev.* **70** 191–214.

LINDLEY, D. V. and NOVICK, M. R. (1981). The role of exchangeability in inference. *Ann. Statist.* **9** 45–58. MR0600531

NEOPHYTOU, A. M., COSTELLO, S., PICCIOTTO, S., BROWN, D. M., ATTFIELD, M. D., BLAIR, A., LUBIN, J. H., STEWART, P. A., VERMEULEN, R. et al. (2019). Diesel exhaust, respirable dust, and ischemic heart disease: An application of the parametric g-formula. *Epidemiology* **30** 177–185.

PEARL, J. (2009). *Causality: Models, Reasoning, and Inference*, 2nd ed. Cambridge Univ. Press, Cambridge. MR2548166 https://doi.org/10.1017/CBO9780511803161

PEARL, J. (2010). On the consistency rule in causal inference: Axiom, definition, assumption, or theorem? *Epidemiology* **21** 872–875. https://doi.org/10.1097/EDE.0b013e3181f5d3fd

ROBINS, J. M. and GREENLAND, S. (1992). Identifiability and exchangeability for direct and indirect effects. *Epidemiology* **3** 143–155. https://doi.org/10.1097/00001648-199203000-00013

ROBINS, J. M., HERNÁN, M. A. and BRUMBACK, B. (2000). Marginal structural models and causal inference in epidemiology. *Epidemiology* **11** 550–560.

ROBINS, J. M. and WASSERMAN, L. (1997). Estimation of effects of sequential treatments by reparameterizing directed acyclic graphs. In *Proceedings of the Thirteenth Conference on Uncertainty in Artificial Intelligence*, *Providence Rhode Island*, *August* 1–3, 1997 (D. Geiger and P. Shenoy, eds.) 409–420. Morgan Kaufmann, San Francisco, CA.

ROSENBAUM, P. R. and RUBIN, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika* **70** 41–55. MR0742974 https://doi.org/10.1093/biomet/70.1.41

RØYSLAND, K. (2011). A martingale approach to continuous-time marginal structural models. *Bernoulli* **17** 895–915. MR2817610 https://doi.org/10.3150/10-BEJ303

RUBIN, D. B. (1978). Bayesian inference for causal effects: The role of randomization. *Ann. Statist.* **6** 34–58. MR0472152

SAARELA, O., BELZILE, L. R. and STEPHENS, D. A. (2016). A Bayesian view of doubly robust causal inference. *Biometrika* **103** 667–681. MR3551791 https://doi.org/10.1093/biomet/asw025

SAARELA, O., STEPHENS, D. A and MOODIE, E. E (2023). Supplement to "The Role of Exchangeability in Causal Inference." https://doi.org/10.1214/22-STS879SUPP

SAARELA, O., STEPHENS, D. A., MOODIE, E. E. M. and KLEIN, M. B. (2015). On Bayesian estimation of marginal structural models. *Biometrics* **71** 279–288. MR3366229 https://doi.org/10.1111/biom.12269

SHAHN, Z., LI, Y., SUN, Z., MOHAN, A., SAMPAIO, C. and HU, J. (2019). G-computation and hierarchical models for estimating multiple causal effects from observational disease registries with irregular visits. *AMIA Joint Summits on Translational Science Proceedings* **2019** 789–798.

TAUBMAN, S. L., ROBINS, J. M., MITTLEMAN, M. A. and HERNÁN, M. A. (2009). Intervening on risk factors for coronary heart disease: An application of the parametric g-formula. *Int. J. Epidemiol.* **38** 1599–1611. https://doi.org/10.1093/ije/dyp192

VANDERWEELE, T. J. (2009a). Concerning the consistency assumption in causal inference. *Epidemiology* **20** 880–883. https://doi.org/10.1097/EDE.0b013e3181bd5638

VANDERWEELE, T. J. (2009b). Marginal structural models for the estimation of direct and indirect effects. *Epidemiology* **20** 18–26. https://doi.org/10.1097/EDE.0b013e31818f69ce

VANSTEELANDT, S., BEKAERT, M. and CLAESKENS, G. (2012). On model selection and model misspecification in causal inference. *Stat. Methods Med. Res.* **21** 7–30. MR2867536 https://doi.org/10.1177/0962280210387717

VAN DER VAART, A. W. (1998). *Asymptotic Statistics. Cambridge Series in Statistical and Probabilistic Mathematics* **3**. Cambridge Univ. Press, Cambridge. MR1652247 https://doi.org/10.1017/CBO9780511802256

VON PLATO, J. (1989). De Finetti's earliest works on the foundations of probability. *Erkenntnis* **31** 263–282.

WESTREICH, D., COLE, S. R., YOUNG, J. G., PALELLA, F., TIEN, P. C., KINGSLEY, L., GANGE, S. J. and HERNÁN, M. A. (2012). The parametric g-formula to estimate the effect of highly active antiretroviral therapy on incident AIDS or death. *Stat. Med.* **31** 2000–2009. MR2956032 https://doi.org/10.1002/sim.5316

Statistical Science
2023, Vol. 38, No. 3, 386–410
https://doi.org/10.1214/22-STS880

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

[1] AITCHISON, J. (1982). The statistical analysis of compositional data. *J. Roy. Statist. Soc. Ser. B* **44** 139–177. MR0676206

[2] AITCHISON, J. (1986). *The Statistical Analysis of Compositional Data. Monographs on Statistics and Applied Probability*. CRC Press, London. MR0865647 https://doi.org/10.1007/978-94-009-4109-0

[3] AITCHISON, J. (1997). The one-hour course in compositional data analysis, or compositional data analysis is simple. In *Proceedings of IAMG'97* (V. Pawlowsky-Glahn, ed.) 3–35. CIMNE, Barcelona.

[4] AITCHISON, J. (2008). The single principle of compositional data analysis, continuing fallacies, confusions and misunderstandings and some suggested remedies. In *Proceedings of CodaWork'08, Keynote Address* 3–35 URL: https://core.ac.uk/download/pdf/132548276.pdf.

[5] AITCHISON, J. and BACON-SHONE, J. (1984). Log constrast models for experiments with mixtures. *Biometrika* **71** 323–330.

[6] AITCHISON, J. and GREENACRE, M. (2002). Biplots of com-

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).

positional data. *J. R. Stat. Soc., Ser. C* **51** 375–392. MR1977249 https://doi.org/10.1111/1467-9876.00275

[7] AMARI, S. (2016). *Information Geometry and Its Applications*. *Applied Mathematical Sciences* **194**. Springer, Tokyo. MR3495836 https://doi.org/10.1007/978-4-431-55978-8

[8] BENZÉCRI, J.-P. (1980). *L'analyse des Données*. *II*: *L'analyse des correspondances*, 3rd ed. Dunod, Paris. MR0593139

[9] BOOESHAGHI, A. S., HALLGRÍMSDÓTTIR, I. B., GÁLVEZ-MERCHÁN, Á. and PACHTER, L. (2022). Depth normalization for single-cell genomics count data. bioRxiv, Cold Spring Harbor Laboratory.

[10] BÓNA, M. (2006). *A Walk Through Combinatorics*: *An Introduction to Enumeration and Graph Theory*, 2nd ed. World Scientific, Hackensack, NJ. MR2361255 https://doi.org/10.1142/6177

[11] BUCCIANTI, A. (2015). The FOREGS repository: Modelling variability in stream water on a continental scale revising classical diagrams from CoDA (compositional data analysis) perspective. *J. Geochem. Explor.* **154** 94–104.

[12] BUETTNER, M., OSTNER, J., MUELLER, C. L., THEIS, F. J. and SCHUBERT, B. (2021). scCODA is a Bayesian model for compositional single-cell data analysis. *Nat. Commun.* **12** 1–10.

[13] BUTLER, A. and GLASBEY, C. (2008). A latent Gaussian model for compositional data with zeros. *J. R. Stat. Soc., Ser. C* **57** 505–520. MR2528668 https://doi.org/10.1111/j.1467-9876.2008.00627.x

[14] COENDERS, G. and GREENACRE, M. (2022). Three approaches to supervised learning for compositional data with pairwise logratios. https://doi.org/10.1080/02664763.2022.2108007

[15] COENDERS, G. and PAWLOWSKY-GLAHN, V. (2020). On interpretations of tests and effect sizes in regression models with a compositional predictor. *SORT* **44** 201–220. MR4121257 https://doi.org/10.2436/20.8080.02.100

[16] COMBETTES, P. L. and MÜLLER, C. L. (2021). Regression models for compositional data: General log-contrast formulations, proximal optimization, and microbiome data applications. *Stat. Biosci.* **13** 217–242. https://doi.org/10.1007/s12561-020-09283-2

[17] CORTÉS, J. A. (2009). On the Harker variation diagrams; a comment on "The statistical analysis of compositional data. Where are we and where should we be heading?" by Aitchison and Egozcue (2005). *Math. Geosci.* **41** 817–828. https://doi.org/10.1007/s11004-009-9222-8

[18] DAVID, M., DAGBERT, M. and BEAUCHEMIN, Y. (1977). Statistical analysis in geology: Correspondence analysis method. *Colo. Sch. Mines Q.* **72** 11–57.

[19] EGOZCUE, J. J. and PAWLOWSKY-GLAHN, V. (2005). Groups of parts and their balances in compositional data analysis. *Math. Geol.* **37** 795–828. MR2183639 https://doi.org/10.1007/s11004-005-7381-9

[20] EGOZCUE, J. J. and PAWLOWSKY-GLAHN, V. (2019). Compositional data: The sample space and its structure. *TEST* **28** 599–638. MR3992128 https://doi.org/10.1007/s11749-019-00670-6

[21] GREENACRE, M. (2019). Discussion of "Compositional data: the sample space and its structure", by Egozcue and Pawlowsky-Glahn. *TEST* **2019** 20–24.

[22] ERB, I. and AY, N. (2021). The information-geometric perspective of compositional data analysis. In *Advances in Compositional Data Analysis* (P. Filzmoser, K. Hron, J. A. Martín-Fernández and J. Palarea-Albaladejo, eds.) 21–43. Springer, New York.

[23] ERB, I. and NOTREDAME, C. (2016). How should we measure proportionality on relative gene expression data? *Theory Biosci.* **135** 21–36.

[24] ERB, I., QUINN, T. P., LOVELL, D. and NOTREDAME, C. (2017). Differential proportionality — a normalization-free approach to differential gene expression. In *Proceedings of CoDaWork* 2017, *the 7th Compositional Data Analysis Workshop*. Available under bioRxiv, pp. 134536. https://doi.org/10.1101/134536

[25] FILZMOSER, P., HRON, K. and TEMPL, M. (2018). *Applied Compositional Data Analysis*. *Springer Series in Statistics*. Springer, Cham. MR3839314 https://doi.org/10.1007/978-3-319-96422-5

[26] FIŠEROVÁ, E. and HRON, K. (2011). On the interpretation of orthonormal coordinates for compositional data. *Math. Geosci.* **43** 455.

[27] GABRIEL, K. R. (1972). Analysis of meteorological data by means of canonical decomposition and biplots. *J. Appl. Meteorol. Climatol.* **11** 1071–1077.

[28] GORDON-RODRIGUEZ, E., QUINN, T. P. and CUNNINGHAM, J. P. (2021). Learning sparse log-ratios for high-throughput sequencing data. *Bioinformatics*. btab645. https://doi.org/10.1093/bioinformatics/btab645

[29] GOWER, J. C. and DIJKSTERHUIS, G. B. (2004). *Procrustes Problems*. *Oxford Statistical Science Series* **30**. Oxford Univ. Press, Oxford. MR2051013 https://doi.org/10.1093/acprof:oso/9780198510581.001.0001

[30] GRAEVE, M. and GREENACRE, M. (2020). The selection and analysis of fatty acid ratios: A new approach for the univariate and multivariate analysis of fatty acid trophic markers in marine organisms. *Limnol. Oceanogr., Methods* **18** 196–210.

[31] GRALINSKA, E., KOHL, C., FADAKAR, B. S. and VINGRON, M. (2022). Visualizing cluster-specific genes from single-cell transcriptomics data using association plots. *J. Mol. Biol.* **434** 167525. https://doi.org/10.1016/j.jmb.2022.167525

[32] GREENACRE, M. (2003). Singular value decomposition of matched matrices. *J. Appl. Stat.* **30** 1101–1113. MR2037841 https://doi.org/10.1080/0266476032000107132

[33] GREENACRE, M. (2009). Power transformations in correspondence analysis. *Comput. Statist. Data Anal.* **53** 3107–3116. MR2667615 https://doi.org/10.1016/j.csda.2008.09.001

[34] GREENACRE, M. (2010). Log-ratio analysis is a limiting case of correspondence analysis. *Math. Geosci.* **42** 129–34.

[35] GREENACRE, M. (2011). Measuring subcompositional incoherence. *Math. Geosci.* **43** 681–93.

[36] GREENACRE, M. (2013). Contribution biplots. *J. Comput. Graph. Statist.* **22** 107–122. MR3044325 https://doi.org/10.1080/10618600.2012.702494

[37] GREENACRE, M. (2016). Data reporting and visualization in ecology. *Polar Biol.* **39** 2189–2205.

[38] GREENACRE, M. (2016). *Correspondence Analysis in Practice*, 3rd ed. CRC Press, Boca Raton, FL.

[39] GREENACRE, M. (2017). 'Size' and 'shape' in the measurement of multivariate proximity. *Methods Ecol. Evol.* **8** 1415–1424. https://doi.org/10.1111/2041-210X.12776

[40] GREENACRE, M. (2018). *Compositional Data Analysis in Practice*. Chapman & Hall / CRC Press, Boca Raton, Florida.

[41] GREENACRE, M. (2019). Variable selection in compositional data analysis using pairwise logratios. *Math. Geosci.* **51** 649–682. MR3981436 https://doi.org/10.1007/s11004-018-9754-x

[42] GREENACRE, M. (2020). Amalgamations are valid in compositional data analysis, can be used in agglomerative clustering, and their logratios have an inverse transformation. *Appl. Comput. Geosc.* **5** 100017.

[43] GREENACRE, M. (2021). Compositional data analysis. *Annu. Rev. Stat. Appl.* **8** 271–299. MR4243548 https://doi.org/10.1146/annurev-statistics-042720-124436

[44] GREENACRE, M. (2022). Compositional data analysis – linear algebra, visualization and interpretation. In *Innovations in Multivariate Statistical Modelling*: *Navigating Theoretical and Multidisciplinary Domains* (A. Bekker and J. Ferreira, eds.) Springer, New York. https://arxiv.org/abs/2110.12439.

[45] GREENACRE, M., GRUNSKY, E. and BACON-SHONE, J. (2020). A comparison of amalgamation and isometric logratios in compositional data analysis. *Comput. Geosci.* **148** 104621.

[46] GREENACRE, M., GRUNSKY, E., BACON-SHONE, J., ERB, I. and QUINN, T. (2023). Supplement to "Aitchison's Compositional Data Analysis 40 Years on: A Reappraisal." https://doi.org/10.1214/22-STS880SUPPA, https://doi.org/10.1214/22-STS880SUPPB, https://doi.org/10.1214/22-STS880SUPPC, https://doi.org/10.1214/22-STS880SUPPD, https://doi.org/10.1214/22-STS880SUPPE

[47] GREENACRE, M. and LEWI, P. (2009). Distributional equivalence and subcompositional coherence in the analysis of compositional data, contingency tables and ratio-scale measurements. *J. Classification* **26** 29–54. MR2507824 https://doi.org/10.1007/s00357-009-9027-y

[48] GREENACRE, M., MÁRTINEZ-ÁLVARO, M. and BLASCO, A. (2021). Compositional data analysis of microbiome and any-omics datasets: A validation of the additive logratio transformation. *Front. Microbiol.* **12** 2625. https://doi.org/10.3389/fmicb.2021.727398

[49] GRUNSKY, E. C. (1985). Recognition of alteration in volcanic rocks using statistical analysis of lithogeochemical data. *J. Geochem. Explor.* **25** 157–183.

[50] HAFEMEISTER, C. and SATIJA, R. (2019). Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome biology* **20** 1–15.

[51] HAUG, T., FALK-PETERSEN, S., GREENACRE, M. et al. (2017). Trophic level and fatty acids in harp seals compared with common minke whales in the Barents Sea. *Marine Biol. Res.* **13** 919–932. https://doi.org/10.1080/17451000.2017.1313988

[52] HELLINGER, E. (1909). Neue Begründung der Theorie quadratischer Formen von unendlichvielen Veränderlichen. *J. Reine Angew. Math.* **136** 210–271. MR1580780 https://doi.org/10.1515/crll.1909.136.210

[53] HRON, K., COENDERS, G., FILZMOSER, P., PALAREA-ALBALADEJO, J., FAMĚRA, M. and GRYGAR, T. M. (2021). Analysing pairwise logratios revisited. *Math. Geosci.* **54** https://www.x-mol.com/paperRedirect/1381133593200320512.

[54] HRON, K., FILZMOSER, P., DE CARITAT, P., FIŠEROVÁ, E. and GARDLO, A. (2017). Weighted pivot coordinates for compositional data and their application to geochemical mapping. *Math. Geosci.* **49** 797–814. MR3672426 https://doi.org/10.1007/s11004-017-9684-z

[55] HSU, L. L. and CULHANE, A. C. (2023). Correspondence analysis for dimension reduction, batch integration, and visualization of single-cell RNA-seq data. *Sci. Rep.* **13** 1197.

[56] HUBERT, L. and ARABIE, P. (1985). Comparing partitions. *J. Classification* **2** 193–218.

[57] JACKSON, D. A. (1997). Compositional data in community ecology: The paradigm or peril of proportions? *Ecology* **78** 929–940.

[58] KRAFT, A., GRAEVE, M., JANSSEN, D. et al. (2017). Arctic pelagic amphipods: Lipid dynamics and life strategy. *J. Plankton Res.* **37** 790–807.

[59] KRZANOWSKI, W. (1987). Selection of variables to preserve multivariate data structure, using principal components. *J. R. Stat. Soc. Ser. C. Appl. Stat.* **36** 22–33.

[60] KYNČLOVÁ, P., HRON, K. and FILZMOSER, P. (2017). Correlation between compositional parts based on symmetric balances. *Math. Geosci.* **49** 777–796. MR3672425 https://doi.org/10.1007/s11004-016-9669-3

[61] LEWI, P. J. (1976). Spectral mapping, a technique for classifying biological activity profiles of chemical compounds. *Arz. Forsch.* **26** 1295–300.

[62] LEWI, P. J. (1986). Analysis of biological activity profiles by Spectramap. *Eur. J. Med. Chem.* **21** 155–62.

[63] LEWI, P. J. (2005). Spectral mapping, a personal and historical account of an adventure in multivariate data analysis. *Chemom. Intell. Lab. Syst.* **77** 215–23.

[64] LOVELL, D., PAWLOWSKY-GLAHN, V., EGOZCUE, J. J., MARGUERAT, S. and BÄHLER, J. (2015). Proportionality: A valid alternative to correlation for relative data. *PLoS Comput. Biol.* **11** e1004075.

[65] LUECKEN, M. D. and THEIS, F. J. (2019). Current best practices in single-cell RNA-seq analysis: a tutorial. *Mol. Syst. Biol.* **15** e8746.

[66] MARTÍN-FERNÁNDEZ, J. A., PAWLOWSKY-GLAHN, V., EGOZCUE, J. J. and TOLOSONA-DELGADO, R. (2018). Advances in principal balances for compositional data. *Math. Geosci.* **50** 273–298. MR3779069 https://doi.org/10.1007/s11004-017-9712-z

[67] MARTÍNEZ-ÁLVARO, M., AUFFRET, M. D., DUTHIE, C. A., DEWHURST, R., CLEVELAND, M., WATSON, M. and ROEHE, R. (2021). Bovine host genome acts on specific metabolism, communication and genetic processes of rumen microbes host-genomically linked to methane emissions. Submitted for Publication https://www.researchsquare.com/article/rs-290150/v1.

[68] MARTÍNEZ-ÁLVARO, M., ZUBIRI-GAITÁN, A., HERNÁNDEZ, P., GREENACRE, M., FERRER, A. and BLASCO, A. (2021). Comprehensive comparison of the cecum microbiome functional core in genetically obese and lean hosts under similar environmental conditions. Accepted by *Commun. Biol.*

[69] MCKINLEY, J. M., GRUNSKY, E. and MUELLER, U. (2018). Environmental monitoring and peat assessment using multivariate analysis of regional-scale geochemical data. *Math. Geosci.* **50** 235–246. MR3772992 https://doi.org/10.1007/s11004-017-9686-x

[70] MEIER, S., FALK-PETERSEN, S., GADE-SØRENSEN, L. A. et al. (2016). Fatty acids in common minke whale (Balaenoptera acutorostrata) blubber reflect the feeding area and food selection, but also high endogenous metabolism. *Marine Biol.*

[71] MURTAGH, F. (1984). Counting dendrograms: A survey. *Discrete Appl. Math.* **7** 191–199. MR0727923 https://doi.org/10.1016/0166-218X(84)90066-0

[72] PALAREA-ALBALADEJO, J. and MARTIN-FERNANDEZ, J. (2015). zCompositions – R package for multivariate imputation of left-censored data under a compositional approach. *Chemom. Intell. Lab. Syst.* **143** 85–96.

[73] PAWLOWSKY-GLAHN, V. and BUCCIANTI, A., eds. (2011) *Compositional Data Analysis*: *Theory and Applications* Wiley, Chichester. MR2920574 https://doi.org/10.1002/9781119976462

[74] PAWLOWSKY-GLAHN, V., EGOZCUE, J. J. and TOLOSANA-DELGADO, R. (2015). *Modeling and Analysis of Compositional Data*. *Statistics in Practice*. Wiley, Chichester. MR3328965

[75] QUINN, T. P. and ERB, I. (2020). Amalgams: Data-driven amalgamation for the dimensionality reduction of compositional data. *NAR Genomics Bioinform.* **2**. lqaa076. https://doi.org/10.1093/nargab/lqaa076

[76] QUINN, T. P., ERB, I., RICHARDSON, M. F. and CROWLEY, T. M. (2018). Understanding sequencing data as compositions: An outlook and review. *Bioinformatics* **34** 2870–2878. https://doi.org/10.1093/bioinformatics/bty175

[77] QUINN, T. P., RICHARDSON, M. F., LOVELL, D. and CROWLEY, T. M. (2017). Propr: An R-package for identifying proportionally abundant features using compositional data analysis. *Sci. Rep.* **7** 16252–16259.

[78] RAND, W. M. (1971). Objective criteria for the evaluation of clustering methods. *J. Amer. Statist. Assoc.* **66** 846–850.

[79] REN, B., BACALLADO, S., FAVARO, S., HOLMES, S. and TRIPPA, L. (2017). Bayesian nonparametric ordination for the analysis of microbial communities. *J. Amer. Statist. Assoc.* **112** 1430–1442. MR3750866 https://doi.org/10.1080/01621459.2017.1288631

[80] REY, F., GREENACRE, M., SILVA NETO, G. M., BUENO-PARDO, J., DOMINGUES, M. R. and CALADO, R. (2022). Fatty acid ratio analysis identifies changes in competent meroplanktonic larvae sampled over different supply events. *Mar. Environ. Res.* **173** 105517.

[81] SCEALY, J. L. and WELSH, A. H. (2011). Regression for compositional data by using distributions defined on the hypersphere. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **73** 351–375. MR2815780 https://doi.org/10.1111/j.1467-9868.2010.00766.x

[82] SCEALY, J. L. and WELSH, A. H. (2014). Colours and cocktails: Compositional data analysis 2013 Lancaster lecture. *Aust. N. Z. J. Stat.* **56** 145–169. MR3226434 https://doi.org/10.1111/anzs.12073

[83] SMITHSON, M. and BROOMELL, S. B. (2022). Compositional data analysis tutorial. *Psychol. Methods* **27**. https://doi.org/10.1037/met0000464

[84] SMYTH, D. (2007). Methods used in the Tellus Geochemical Mapping of Northern Ireland. British Geological Survey, Open Report, OR/07/022.

[85] STANLEY, C. R. (2019). Molar element ratio analysis of lithogeochemical data: A toolbox for use in mineral exploration and mining. *Geochem., Explor. Environ. Anal.* **20** 233–256.

[86] STEPHENS, M. A. (1982). Use of the von Mises distribution to analyse continuous proportions. *Biometrika* **69** 197–203. MR0655685 https://doi.org/10.1093/biomet/69.1.197

[87] TE BEEST, D. E., NIJHUIS, E. H., MÖHLMANN, T. W. R. and TER BRAAK, C. J. F. (2021). Log-ratio analysis of microbiome data with many zeroes is library size dependent. *Mol. Ecol. Resour.* **21** 1866–1874. https://doi.org/10.1111/1755-0998.13391

[88] R CORE TEAM (2021). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

[89] TOWNES, F. W., HICKS, S. C., ARYEE, M. J. and IRIZARRY, R. A. (2019). Feature selection and dimension reduction for single-cell RNA-Seq based on a multinomial model. *Genome Biol.* **20** 295. https://doi.org/10.1186/s13059-019-1861-6

[90] VAN DEN BOOGAART, K. G. and TOLOSANA-DELGADO, R. (2013). *Analyzing Compositional Data with R. Use R!* Springer, Heidelberg. MR3099409 https://doi.org/10.1007/978-3-642-36809-7

[91] VAN DEN WOLLENBERG, A. L. (1977). Redundancy analysis, an alternative for canonical analysis. *Psychometrika* **42** 207–219.

[92] WOOD, J. and GREENACRE, M. (2021). Making the most of expert knowledge to analyse archaeological data: A case study on Parthian and Sasanian glazed pottery. *Archael. Anthrop. Sci.* **13** 110.

[93] YOO, J., SUN, Z., GREENACRE, M., MAD, Q., CHUNG, D. and KIM, Y. M. (2022). A guideline for the statistical analysis of compositional data in immunology. *Commun. Stat. Appl. Methods* **29** 453–469.

Statistical Science
2023, Vol. 38, No. 3, 411–439
https://doi.org/10.1214/22-STS881

Statistical Embedding: Beyond Principal Components

Dag Tjøstheim, Martin Jullum and Anders Løland

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

AIZERMAN, M. A., BRAVERMAN, E. M. and ROZONOER, L. I. (1956). Theoretical foundations of the potential function method in pattern recognition learning. *Autom. Remote Control* **25** 821–137.

ARMILLOTTA, M., FOKIANOS, K. and KRIKIDIS, I. (2022). Generalized linear models network autoregression. In *Network Science* 112–125. International Conference on Network Science.

BAGLAMA, J. and REICHEL, L. (2005). Augmented implicitly restarted Lanczos bidiagonalization methods. *SIAM J. Sci. Comput.* **27** 19–42. MR2201173 https://doi.org/10.1137/04060593X

BELKIN, M. and NIYOGI, P. (2002). Laplacian eigenmaps and spectral techniques for embedding and clustering. In *Advances in Information Processing Systems* (T. K. Leen, T. G. Dietterich and V. Treps, eds.). MIT Press, Cambridge, MA.

BELKIN, M. and NIYOGI, P. (2003). Laplacian eigenmaps for dimensionality reduction and data representation. *Neural Comput.* **15** 1373–1396.

BIAN, R., KOH, Y. S., DOBBIE, G. and DIVOLI, A. (2019). Network embedding and change modeling in dynamic heterogeneous networks. In *Proceedings of the* 42*nd International ACM SIGIR Conference on Research and Development in Information Retrieval* 861–864.

BICKEL, P. and CHEN, A. (2009). A nonparametric view of network models and Newman–Girvan and other modularities. *Proc. Natl. Acad. Sci.* **106** 21068–21073.

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).

BICKEL, P. J. and SARKAR, P. (2016). Hypothesis testing for automated community detection in networks. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **78** 253–273. MR3453655 https://doi.org/10.1111/rssb.12117

BICKEL, P., CHOI, D., CHANG, X. and ZHANG, H. (2013). Asymptotic normality of maximum likelihood and its variational approximation for stochastic blockmodels. *Ann. Statist.* **41** 1922–1943. MR3127853 https://doi.org/10.1214/13-AOS1124

BLONDEL, V. D., GUILLAUME, J.-L., LAMBIOTTE, R. and LEFEBVRE, E. (2008). Fast unfolding of communities in large networks. *J. Stat. Mech. Theory Exp.* **2008** P10008.

BOSER, B. E., GUYON, I. M. and VAPNIK, V. N. (1992). A training algorithm for optimal margin classifiers. In *Fifth Annual Workshop on COLT*, ACM, Pittsburgh, PA.

BUKKURI, A., ANDOR, N. and DARCY, I. K. (2021). Applications of topological data analysis on oncology. *Front. Artif. Intell. Mach. Learn. Artif. Intell.* **4** 1–14.

CANNINGS, T. I. and SAMWORTH, R. J. (2017). Random-projection ensemble classification. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **79** 959–1035. MR3689307 https://doi.org/10.1111/rssb.12228

CARLSSON, G. (2009). Topology and data. *Bull. Amer. Math. Soc.* (*N.S.*) **46** 255–308. MR2476414 https://doi.org/10.1090/S0273-0979-09-01249-X

CARRIÈRE, M., MICHEL, B. and OUDOT, S. (2018). Statistical analysis and parameter selection for Mapper. *J. Mach. Learn. Res.* **19** Paper No. 12, 39 pp. MR3862419

CARRIÈRE, M. and RABADÁN, R. (2020). Topological data analysis of single-cell Hi-C contact maps. In *Topological Data Analysis—The Abel Symposium* 2018. *Abel Symp.* **15** 147–162. Springer, Cham. MR4338672 https://doi.org/10.1007/978-3-030-43408-3_6

CHAZAL, F. and MICHEL, B. (2017). An introduction to topological data analysis: Fundamental and practical aspects for data scientists. Preprint. Available at arXiv:1710.04019v1.

CHAZAL, F. and MICHEL, B. (2021). An introduction to topological data analysis: Fundamental and practical aspects for data scientists. *Front. Artif. Intell. Mach. Learn. Artif. Intell.* **4** 1–28.

CHEN, Y.-C., GENOVESE, C. R. and WASSERMAN, L. (2015). Asymptotic theory for density ridges. *Ann. Statist.* **43** 1896–1928. MR3375871 https://doi.org/10.1214/15-AOS1329

CHEN, Y. C., HO, S., FREEMEN, P. E., GENOVESE, C. R. and WASSERMAN, L. (2015a). Cosmic web reconstruction through density ridges: Methods and algorithm. *Mon. Not. R. Astron. Soc.* **454** 1140–1156.

CHEN, Y. C., HO, S., TENNETI, A., MANDELBAUM, R., CROFT, R., DIMATTEO, T., FREEMAN, P. E., GENOVESE, C. R. and WASSERMAN, L. (2015b). Investigating galaxy-filament alignments in hydrodynamic simulations using density ridges. *Mon. Not. R. Astron. Soc.* **454** 3341–3350.

CLAESKENS, G., CROUX, C. and VAN KERCKHOVEN, J. (2008). An information criterion for variable selection in support vector machines. *J. Mach. Learn. Res.* **9** 541–558. MR2417246 https://doi.org/10.2139/ssrn.1094652

COIFMAN, R. R. and LAFON, S. (2006). Diffusion maps. *Appl. Comput. Harmon. Anal.* **21** 5–30. MR2238665 https://doi.org/10.1016/j.acha.2006.04.006

CORMEN, T. H., LEISERSON, C. E., RIVEST, R. L. and STEIN, C. (2022). *Introduction to Algorithms*, 3rd ed. MIT Press, Cambridge, MA. MR2572804

CRANE, H. and DEMPSEY, W. (2015). A framework for statistical network modeling. Preprint. Available at arXiv:1509.08185.

CRAWFORD, L., MONOD, A., CHEN, A. X., MUKHERJEE, S. and RABADÁN, R. (2020). Predicting clinical outcomes in glioblastoma: An application of topological and functional data analysis. *J. Amer. Statist. Assoc.* **115** 1139–1150. MR4143455 https://doi.org/10.1080/01621459.2019.1671198

CUI, P., WANG, X., PEI, J. and ZHU, W. (2019). A survey on network embedding. *IEEE Trans. Knowl. Data Eng.* **31** 833–852.

DE SILVA, V. and TENENBAUM, J. (2002). Global versus local methods in nonlinear dimensionality reduction. *Adv. Neural Inf. Process. Syst.* **15**.

DECELLE, A., KRZAKALA, F., MOORE, C. and ZDEBOROVÁ, L. (2011). Asymptotic analysis of the stochastic block model for modular networks and its algorithmic applications. *Phys. Rev. E* **84** 066106.

DEVROYE, L. and WISE, G. L. (1980). Detection of abnormal behavior via nonparametric estimation of the support. *SIAM J. Appl. Math.* **38** 480–488. MR0579432 https://doi.org/10.1137/0138038

DONG, Y., CHAWLA, N. V. and SWAMI, A. (2017). Metapath2vec: Scalable representation learning for heterogeneous networks. Kid 17, 2017, Halifax, NS, Canada.

DONG, W., MOSES, C. and LI, K. (2018). Efficient *k*-nearest neighbour graph construction for generic similarity measures. In *Proceedings of the* 20*th International Conference of the World Wide Web* 577–586, New York.

DU, L., WANG, Y., SONG, G., LU, Z. and WANG, J. (2018). Dynamic network embedding: An extended approach for Skip-Gram based network embedding. In *Proceedings of the* 27*th International Joint Conference on Artificial Intelligence*, *IJ*(*AI*-18).

DUCHAMP, T. and STUETZLE, W. (1996). Extremal properties of principal curves in the plane. *Ann. Statist.* **24** 1511–1520. MR1416645 https://doi.org/10.1214/aos/1032298280

EDELSBRUNNER, H., LETCHER, D. and ZOMORODIAN, A. (2002). Topological persistence and simplification. *Discrete Comput. Geom.* **28** 511–533. MR1949898 https://doi.org/10.1007/s00454-002-2885-2

GENOVESE, C. R., PERONE-PACIFICO, M., VERDINELLI, I. and WASSERMAN, L. (2012). Manifold estimation and singular deconvolution under Hausdorff loss. *Ann. Statist.* **40** 941–963. MR2985939 https://doi.org/10.1214/12-AOS994

GENOVESE, C. R., PERONE-PACIFICO, M., VERDINELLI, I. and WASSERMAN, L. (2014). Nonparametric ridge estimation. *Ann. Statist.* **42** 1511–1545. MR3262459 https://doi.org/10.1214/14-AOS1218

GHOJOGH, B., GHODSI, A., KARRAY, F. and CROWLEY, M. (2021). Johnson–Lindenstrauss lemma, linear and nonlinear random projections, random Fourier features and random kitchen sinks: Tutorial and survey. Preprint. Available at arXiv:2108.04172v1.

GHRIST, R. (2018). Homological algebra and data. In *The Mathematics of Data. IAS/Park City Math. Ser.* **25** 273–325. Amer. Math. Soc., Providence, RI. MR3839171

GIRVAN, M. and NEWMAN, M. E. J. (2002). Community structure in social and biological networks. *Proc. Natl. Acad. Sci. USA* **99** 7821–7826. MR1908073 https://doi.org/10.1073/pnas.122653799

GREENE, D. and CUNNINGHAM, P. (2011). Tracking the evolution of communities in dynamic social networks. Report Idiro Technologies, Dublin, Ireland.

GRETTON, A. (2019). Introduction to RKHS, and some simple kernel algorithms. Lecture notes.

GROVER, A. and LESKOVEC, J. (2016). node2vec: Scalable feature learning for networks. In *Proceedings of the* 22*nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 855–864.

HAGHVERDI, L., BUETTNER, F. and THEIS, F. J. (2015). Diffusion maps for high-dimensional single-cell analysis of differentiation data. *Bioinformatics* **31** 2989–2998. https://doi.org/10.1093/bioinformatics/btv325

HASTIE, T. (1984). Principal curves and surfaces. Laboratory for Computational Statistics Technical Report 11, Stanford Univ., Dept. Statistics. MR2634007

HASTIE, T. and STUETZLE, W. (1989). Principal curves. *J. Amer. Statist. Assoc.* **84** 502–516. MR1010339

HASTIE, T., TIBSHIRANI, R. and FRIEDMAN, J. (2019). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed. *Springer Series in Statistics*. Springer, New York. MR2722294 https://doi.org/10.1007/978-0-387-84858-7

HINTON, G. E. and ROWEIS, S. T. (2002). Stochastic neighbour embedding. *Adv. Neural Inf. Process. Syst.* **15** 833–840.

HINTON, G. E. and SALAKHUTDINOV, R. R. (2006). Reducing the dimensionality of data with neural networks. *Science* **313** 504–507. MR2242509 https://doi.org/10.1126/science.1127647

HOFF, P. D., RAFTERY, A. E. and HANDCOCK, M. S. (2002). Latent space approaches to social network analysis. *J. Amer. Statist. Assoc.* **97** 1090–1098. MR1951262 https://doi.org/10.1198/016214502388618906

HOLLAND, P. W., LASKEY, K. B. and LEINHARDT, S. (1983). Stochastic blockmodels: First steps. *Soc. Netw.* **5** 109–137. MR0718088 https://doi.org/10.1016/0378-8733(83)90021-7

HOTELLING, H. (1933). Analysis of a complex of statistical variables into principal components. *J. Educ. Psychol.* **24** 417–441.

HOTELLING, H. (1936). Relations between two sets of variates. *Biometrika* **28** 321–377.

HYVÄRINEN, A. and OJA, E. (2000). Independent component analysis: Algorithms and applications. *Neural Netw.* **13** 411–430.

JOHNSON, W. B. and LINDENSTRAUSS, J. (1984). Extensions of Lipschitz mappings into a Hilbert space. In *Conference in Modern Analysis and Probability* (*New Haven, Conn.*, 1982). *Contemp. Math.* **26** 189–206. Amer. Math. Soc., Providence, RI. MR0737400 https://doi.org/10.1090/conm/026/737400

JOLLIFFE, I. T. (2002). *Principal Component Analysis*, 2nd ed. *Springer Series in Statistics*. Springer, New York. MR2036084

JOSSE, J. and HUSSON, F. (2012). Selecting the number of components in principal component analysis using cross-validation approximations. *Comput. Statist. Data Anal.* **56** 1869–1879. MR2892383 https://doi.org/10.1016/j.csda.2011.11.012

KARRER, B. and NEWMAN, M. E. J. (2011). Stochastic blockmodels and community structure in networks. *Phys. Rev. E* (3) **83** 016107, 10 pp. MR2788206 https://doi.org/10.1103/PhysRevE.83.016107

KAZEMI, S. M., GOEL, R., JAIN, K., KOBYZEV, I., SETHI, A., FORSYTH, P. and POUPART, P. (2020). Representation learning for dynamic graphs: A survey. *J. Mach. Learn. Res.* **21** Paper No. 70, 73 pp. MR4095349

KIM, J., RINALDO, A. and WASSERMAN, L. (2019). Minimax rates for estimating the dimension of a manifold. *J. Comput. Geom.* **10** 42–95. MR3918925 https://doi.org/10.20382/jocg.v10i1a3

KOBOUROV, S. (2012). Spring embedders and forced directed graph drawing algorithms. Preprint. Available at arXiv:1201.3011.

KOHONEN, T. (1982). Self-organized formation of topologically correct feature map. *Biol. Cybernet.* **43** 59–69.

KONISHI, S. and KITAGAWA, G. (2008). *Information Criteria and Statistical Modeling. Springer Series in Statistics*. Springer, New York. MR2367855 https://doi.org/10.1007/978-0-387-71887-3

KOSSINETS, G. and WATTS, D. J. (2006). Empirical analysis of an evolving social network. *Science* **311** 88–90. MR2192483 https://doi.org/10.1126/science.1116869

LEE, C. and WILKINSON, D. J. (2019). A review of stochastic block models and extensions for graph clustering. *Appl. Netw. Sci.* **4** 122.

LEI, J. and RINALDO, A. (2015). Consistency of spectral clustering in stochastic block models. *Ann. Statist.* **43** 215–237. MR3285605 https://doi.org/10.1214/14-AOS1274

LEVINA, E. and BICKEL, P. (2004). Maximum likelihood estimation of intrinsic dimension. In *Advances in Neural Information Processing Systems* (L. Saul, Y. Weiss and L. Bottou, eds.) **17**. MIT Press, Cambridge, MA.

LI, P., HASTIE, T. J. and CHURCH, K. W. (2007). Nonlinear estimators and tail bounds for dimension reduction in l_1 using Cauchy random projections. *J. Mach. Learn. Res.* **8** 2497–2532. MR2353840 https://doi.org/10.1007/978-3-540-72927-3_37

LIM, B. and ZOHREN, S. (2021). Time-series forecasting with deep learning: A survey. *Philos. Trans. R. Soc. Lond. A* **379** Paper No. 20200209, 14 pp. MR4236146 https://doi.org/10.1098/rsta.2020.0209

LITTLE, A. V., MAGGIONI, M. and ROSASCO, L. (2011). Multiscale geometric methods for estimating intrinsic dimension. In *Proc. SampTA* 4:2.

LUDKIN, M., ECKLEY, I. and NEAL, P. (2018). Dynamic stochastic block models: Parameter estimation and detection of changes in community structure. *Stat. Comput.* **28** 1201–1213. MR3850391 https://doi.org/10.1007/s11222-017-9788-9

LUNDE, B. Å. S., KLEPPE, T. S. and SKAUG, H. J. (2020). An information criterion for automatic gradient tree boosting. Preprint. Available at arXiv:2008.05926.

MARKOV, A. (1958). The insolubility of the problem of homeomorphy. *Dokl. Akad. Nauk SSSR* **121** 218–220. MR0097793

MCINNES, L., HEALY, J. and MELVILLE, J. (2018). UMAP: Uniform manifold approximation for dimension reduction. Preprint. Available at arXiv:1802.03426v2.

MIKOLOV, T., SUTSKEVER, I., CHEN, K., CORRADO, G. and DEAN, J. (2013). Distributed representation of words and phrases and their composability. In *Advances in Neural Information Processing Systems* 26: *Proceedings Annual 27th Conference on Neural Information Processing Systems*, *Lake Tahoe*, *Nevada*, *USA*.

NEWMAN, M. E. J. (2006). Modularity and community structure in networks. *Proc. Natl. Acad. Sci.* **103** 8577–8582.

NEWMAN, M. (2020). *Networks*, 2nd ed. Oxford Univ. Press, Oxford. MR3838417 https://doi.org/10.1093/oso/9780198805090.001.0001

NEWMAN, M. E. J. and GIRVAN, M. (2004). Finding and evaluating community networks. *Phys. Rev. E* **69** 026113.

NEWMAN, M. E. J. and REINERT, G. (2016). Estimating the number of communities in a network. *Phys. Rev. Lett.* **137** 078301.

NIYOGI, P., SMALE, S. and WEINBERGER, S. (2008). Finding the homology of submanifolds with high confidence from random samples. *Discrete Comput. Geom.* **39** 419–441. MR2383768 https://doi.org/10.1007/s00454-008-9053-2

OTNEIM, H., JULLUM, M. and TJØSTHEIM, D. (2020). Pairwise local Fisher and naive Bayes: Improving two standard discriminants. *J. Econometrics* **216** 284–304. MR4077395 https://doi.org/10.1016/j.jeconom.2020.01.019

OZERTEM, U. and ERDOGMUS, D. (2011). Locally defined principal curves and surfaces. *J. Mach. Learn. Res.* **12** 1249–1286. MR2804600

PEARSON, K. (1901). On lines and planes of closest fit to systems of points in space. *Philos. Mag.* **2** 559–572.

PEIXITO, T. P. (2021). Descriptive vs. inferential community detection: Pitfalls, myths and half-truths. Preprint. Available at arXiv:2112.00183v1.

PEIXOTO, T. P. (2019). Bayesian stochastic blockmodeling. In *Advances in Network Clustering and Blockmodeling* 289–332.

PEROZZI, B., AL-RFOU, R. and SKIENA, S. (2014). Deepwalk: Online learning of social representations. In *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 701–710.

QIAO, W. and POLONIK, W. (2021). Algorithms for ridge estimation with convergence guarantees. Preprint. Available at arXiv:2014.12314v1.

QIU, J., DONG, Y., MA, H., LI, J., WANG, K. and TANG, J. (2018). Network embedding as matrix factorization: Unifying DeepWalk, LINE, PTE, and node2vec. In *Proceedings WSDM*. ACM, New Tork.

QIU, J., DONG, Y., MA, H., LI, J., WANG, K. and TANG, J. (2019). NetSMF: Large-scale network embedding as sparse matrix factorization. In *Proceedings of the* 2019 *World Wide Web Conference*, *May* 13–17, *San Francisco*, *CA*, *USA*.

RAVISSHANKER, N. and CHEN, R. (2019). Topological data analysis (TDA) for time series. Preprint. Available at arXiv:1909.10604v1.

ROHE, K., CHATTERJEE, S. and YU, B. (2011). Spectral clustering and the high-dimensional stochastic blockmodel. *Ann. Statist.* **39** 1878–1915. MR2893856 https://doi.org/10.1214/11-AOS887

ROHE, K., QIN, T. and YU, B. (2016). Co-clustering directed graphs to discover asymmetries and directional communities. *Proc. Natl. Acad. Sci. USA* **113** 12679–12684. MR3576189 https://doi.org/10.1073/pnas.1525793113

ROWEIS, S. T. and SAUL, L. K. (2000). Nonlinear dimensionality reduction by locally linear embedding. *Science* **290** 2323–2326. https://doi.org/10.1126/science.290.5500.2323

SALINAS, D., FLUNKERT, V., GASTHAUS, J. and JANUSCHOWSKI, T. (2020). DeepAR: Probabilistic forecasting with autoregressive recurrent networks. *Int. J. Forecast.* **36** 1181–1191.

SAMMON, J. W. (1969). A nonlinear mapping for data structure analysis. *IEEE Trans. Comput.* **18** 403–409.

SCHÖLKOPF, B., SMOLA, A. and MÜLLER, K.-L. (2005). Kernel principal components. *Lecture Notes in Comput. Sci.* **1327** 583–588.

SHAHRIARI, B., SWERSKY, K., WANG, Z., ADAMS, R. P. and DE FREITAS, N. (2015). Taking the human out of the loop: A review of Bayesian optimization. *Proc. IEEE* **104** 148–175.

SINGH, G., MEMOLI, F. and CARLSSON, G. (2007). Topological methods for the analysis of high dimensional data sets and 3D object recognition. In *Eurographics Symposium on Point Based Graphics* (M. Botsch and R. Pajarola, eds.). The Eurographics Association.

SUN, Y., NORICK, B., HAN, J., YAN, X., YU, P. and YU, X. (2012). Integrating meta-path selection with user-guided object clustering in heterogeneous information networks. In *KDD* '12: *Proceedings of the* 18*th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 1348–1356.

TANG, J., QU, M. and MEI, Q. (2015). PTE: Predictive text embedding through large-scale heterogeneous text networks. Preprint. Available at arXiv:1508.00200v1.

TANG, J., QU, M., WANG, M., ZHANG, M., YAN, J. and MEI, Q. (2015). LINE: Large-scale information network embedding. In *Proceedings of the* 24*th International Conference on World Wide Web* 1067–1077.

TANG, J., LIU, J., ZHANG, M. and MEI, Q. (2016). Visualizing large-scale and high-dimensional data. In *Proceedings of the* 25*th International Conference on World Wide Web* 287–297.

TENENBAUM, J. B., DE SILVA, V. and LANGFORD, J. C. (2000). A global geometric framework for nonlinear dimensionality reduction. *Science* **290** 2319–2323. https://doi.org/10.1126/science.290.5500.2319

TJØSTHEIM, D., JULLUM, M. and LØLAND, A. (2023). Some recent trends in embedding of time series and dynamic networks. *J. Time Ser. Anal.* To appear. https://doi.org/10.1111/jtsa.12677

TJØSTHEIM, D., JULLUM, M. and LØLAND, A. (2023). Supplement to "Statistical embedding: Beyond principal components". https://doi.org/10.1214/22-STS881SUPP

TJØSTHEIM, D., OTNEIM, H. and STØVE, B. (2022a). Statistical dependence: Beyond Pearson's ρ. *Statist. Sci.* **37** 90–109. MR4371097 https://doi.org/10.1214/21-sts823

TJØSTHEIM, D., OTNEIM, H. and STØVE, B. (2022b). *Statistical Modeling Using Local Gaussian Approximation*. Elsevier/Academic Press, London. MR4382419

TORGERSON, W. S. (1952). Multidimensional scaling: I. Theory and method. *Psychometrika* **17** 401–419. MR0054219 https://doi.org/10.1007/BF02288916

TUTTE, W. T. (1963). How to draw a graph. *Proc. Lond. Math. Soc.* (3) **13** 743–767. MR0158387 https://doi.org/10.1112/plms/s3-13.1.743

VAN DER MAATEN, L. (2014). Accelerating t-SNE using tree-based algorithms. *J. Mach. Learn. Res.* **15** 3221–3245. MR3277169

VAN DER MAATEN, L. and HINTON, G. (2008). Visualizing data using t-SNE. *J. Mach. Learn. Res.* **9** 2579–2605.

VAN DER MAATEN, L., POSTMA, E. and VAN DER HERIK, J. (2009). Dimensionality reduction: A comparative review. Tilburg Centre for Creative Computing, TiCC TR 2009.005.

VON LUXBURG, U. (2007). A tutorial on spectral clustering. *Stat. Comput.* **17** 395–416. MR2409803 https://doi.org/10.1007/s11222-007-9033-z

WANG, Y. X. R. and BICKEL, P. J. (2017). Likelihood-based model selection for stochastic block models. *Ann. Statist.* **45** 500–528. MR3650391 https://doi.org/10.1214/16-AOS1457

WASSERMAN, L. (2018). Topological data analysis. *Annu. Rev. Stat. Appl.* **5** 501–535. MR3774757 https://doi.org/10.1146/annurev-statistics-031017-100045

WEI, Y.-C. and CHENG, C.-K. (1989). Towards efficient hierarchical designs by ratio cut partitioning. In 1989 *IEEE International Conference on Computer-Aided Design. Digest of Technical Papers* 298–301. IEEE.

XIE, H., LI, J. and XUE, H. (2018). A survey of dimensionality reduction techniques based on random projection. Preprint. Available at arXiv:1706.04371v4.

YOUNG, G. and HOUSEHOLDER, A. S. (1938). Discussion of a set of points in terms of their mutual distances. *Psychometrika* **3** 19–22.

YOUNG, T., HAZARIKA, D., PORIA, S. and CAMBRIA, E. (2018). Recent trends in deep learning based natural language processing. *IEEE Comput. Intell. Mag.* **13** 55–75.

ZHANG, J. and CHEN, Y. (2020). Modularity based community detection in heterogeneous networks. *Statist. Sinica* **30** 601–629. MR4213981

ZHENG, Q. (2016). Spectral techniques for heterogeneous social networks. Ph.D. thesis, Queen's Univ., Ontario, Canada.

ZHOU, C., LIU, Y., LIU, X. and GAO, J. (2017). Scalable graph embedding for asymmetric proximity. In *Proceedings of the* 31*st AAAI Conference on Artificial Intelligence*.

ZHU, X. and PAN, R. (2020). Grouped network vector autoregression. *Statist. Sinica* **30** 1437–1462. MR4257540 https://doi.org/10.5705/ss.202017.0533

ZHU, X., PAN, R., LI, G., LIU, Y. and WANG, H. (2017). Network vector autoregression. *Ann. Statist.* **45** 1096–1123. MR3662449 https://doi.org/10.1214/16-AOS1476

ZOMORODIAN, A. and CARLSSON, G. (2005). Computing persistent homology. *Discrete Comput. Geom.* **33** 249–274. MR2121296 https://doi.org/10.1007/s00454-004-1146-y

Statistical Science
2023, Vol. 38, No. 3, 440–457
https://doi.org/10.1214/23-STS882

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, $\widehat{\Gamma}$, 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

[1] AGENCY, U. E. P. (2021). How people are exposed to mercury. Available at www.epa.gov/mercury.

[2] ALBERS, W., BICKEL, P. J. and VAN ZWET, W. R. (1976). Asymptotic expansions for the power of distribution free tests in the one-sample problem. *Ann. Statist.* **4** 108–156. MR0391373

[3] BERGER, R. L. and BOOS, D. D. (1994). *P* values maximized over a confidence set for the nuisance parameter. *J. Amer. Statist. Assoc.* **89** 1012–1016. MR1294746

[4] BERK, R. H. and JONES, D. H. (1978). Relatively optimal combinations of test statistics. *Scand. J. Stat.* **5** 158–162. MR0509452

[5] BIRCH, M. W. (1964). The detection of partial association. I. The 2×2 case. *J. Roy. Statist. Soc. Ser. B* **26** 313–324. MR0176562

[6] BONVINI, M. and KENNEDY, E. H. (2022). Sensitivity analysis via the proportion of unmeasured confounding. *J. Amer. Statist. Assoc.* **117** 1540–1550. MR4480730 https://doi.org/10.1080/01621459.2020.1864382

[7] BROWN, B. M. (1981). Symmetric quantile averages and related estimators. *Biometrika* **68** 235–242. MR0614960 https://doi.org/10.1093/biomet/68.1.235

[8] CAMPBELL, D. T. (1969). Prospective: Artifact and control. In *Artifacts in Behavioral Research* (R. Rosenthal and R. Rosnow, eds.) Academic Press, New York.

[9] FISHER, R. A. (1935). *The Design of Experiments*. Oliver & Boyd, Edinburgh.

[10] FOGARTY, C. B. and SMALL, D. S. (2016). Sensitivity analysis for multiple comparisons in matched observational studies through quadratically constrained linear programming. *J. Amer. Statist. Assoc.* **111** 1820–1830. MR3601738 https://doi.org/10.1080/01621459.2015.1120675

[11] GASTWIRTH, J. L. (1966). On robust procedures. *J. Amer. Statist. Assoc.* **61** 929–948. MR0205397

[12] GASTWIRTH, J. L., KRIEGER, A. M. and ROSENBAUM, P. R. (2000). Asymptotic separability in sensitivity analysis. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **62** 545–555. MR1772414 https://doi.org/10.1111/1467-9868.00249

[13] GOODMAN, S. N., SCHNEEWEISS, S. and BAIOCCHI, M. (2017). Using design thinking to differentiate useful from misleading evidence in observational research. *J. Amer. Med. Assoc.* **317** 705–707.

[14] GROENEVELD, R. A. (1972). Asymptotically optimal group rank tests for location. *J. Amer. Statist. Assoc.* **67** 847–849.

Paul R. Rosenbaum is Robert G. Putzel Professor Emeritus, Department of Statistics and Data Science, Wharton School, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA (e-mail: rosenbaum@wharton.upenn.edu).

MR0395030
[15] HASEGAWA, R. B., WEBSTER, D. W. and SMALL, D. S. (2019). Evaluating Missouri's handgun purchaser law: A bracketing method for addressing concerns about history interacting with group. *Epidemiology* **30** 371–379.
[16] HOGAN, W. W. (1973). Point-to-set maps in mathematical programming. *SIAM Rev.* **15** 591–603. MR0345641 https://doi.org/10.1137/1015073
[17] KARMAKAR, B., FRENCH, B. and SMALL, D. S. (2019). Integrating the evidence from evidence factors in observational studies. *Biometrika* **106** 353–367. MR3949308 https://doi.org/10.1093/biomet/asz003
[18] KIM, H. J., JUNG, S., ELIASSEN, A. H., CHEN, W. Y., WILLETT, W. C. and CHO, E. (2017). Alcohol consumption and breast cancer risk in younger women according to family history of breast cancer and folate intake. *Amer. J. Epidemiol.* **186** 524–531.
[19] KUROKI, M. and PEARL, J. (2014). Measurement bias and effect restoration in causal inference. *Biometrika* **101** 423–437. MR3215357 https://doi.org/10.1093/biomet/ast066
[20] LEHMANN, E. L. (1975). *Nonparametrics*. Holden-Day, Oakland, CA.
[21] LEHMANN, E. L. and ROMANO, J. P. (2006). *Testing Statistical Hypotheses*. Springer, Berlin.
[22] LOCONTE, N. K., BREWSTER, A. M., KAUR, J. S., MERRILL, J. K. and ALBERG, A. J. (2018). Alcohol and cancer: A statement of the American society of clinical oncology. *J. Clin. Oncol.* **36** 83–93.
[23] LU, B., CAI, D. and TONG, X. (2018). Testing causal effects in observational survival data using propensity score matching design. *Stat. Med.* **37** 1846–1858. MR3799844 https://doi.org/10.1002/sim.7599
[24] MARITZ, J. S. (1979). A note on exact robust confidence intervals for location. *Biometrika* **66** 163–166. MR0529161 https://doi.org/10.1093/biomet/66.1.163
[25] MARKOWSKI, E. P. and HETTMANSPERGER, T. P. (1982). Inference based on simple rank step score statistics for the location model. *J. Amer. Statist. Assoc.* **77** 901–907. MR0686416
[26] MCKILLIP, J. (1992). Research without control groups: A control construct design. In *Methodological Issues in Applied Social Psychology* 159–175. Springer, Berlin.
[27] MUNAFÒ, M. R., HIGGINS, J. P. T. and SMITH, G. D. (2021). Triangulating evidence through the inclusion of genetically informed designs. *Cold Spring Harbor Perspect. Med.* **11** a040659.
[28] NEYMAN, J. (1990). On the application of probability theory to agricultural experiments. Essay on principles. Section 9. *Statist. Sci.* **5** 465–472. MR1092986 https://doi.org/10.1214/ss/1177012031
[29] NOETHER, G. E. (1973). Some simple distribution-free confidence intervals for the center of a symmetric distribution. *J. Amer. Statist. Assoc.* **68** 716–719.
[30] PEDERSEN, G. A., MORTENSEN, G. K. and LARSEN, E. H. (1994). Beverages as a source of toxic trace element intake. *Food Add. Contam.* **11** 351–363.
[31] PIMENTEL, S. D., SMALL, D. S. and ROSENBAUM, P. R. (2016). Constructed second control groups and attenuation of unmeasured biases. *J. Amer. Statist. Assoc.* **111** 1157–1167. MR3561939 https://doi.org/10.1080/01621459.2015.1076342
[32] QUADE, D. (1979). Using weighted rankings in the analysis of complete blocks with additive block effects. *J. Amer. Statist. Assoc.* **74** 680–683. MR0548265
[33] REYNOLDS, K. D. and WEST, S. G. (1987). A multiplist strategy for strengthening nonequivalent control group designs. *Eval. Rev.* **11** 691–714.
[34] ROSENBAUM, P. R. (1987). The role of a second control group in an observational study. *Statist. Sci.* **2** 292–306.
[35] ROSENBAUM, P. R. (1989). The role of known effects in observational studies. *Biometrics* **45** 557–569. MR1010518 https://doi.org/10.2307/2531497
[36] ROSENBAUM, P. R. (1989). On permutation tests for hidden biases in observational studies: An application of Holley's inequality to the Savage lattice. *Ann. Statist.* **17** 643–653. MR0994256 https://doi.org/10.1214/aos/1176347131
[37] ROSENBAUM, P. R. (2002). *Observational Studies*, 2nd ed. *Springer Series in Statistics*. Springer, New York. MR1899138 https://doi.org/10.1007/978-1-4757-3692-2
[38] ROSENBAUM, P. R. (2007). Sensitivity analysis for m-estimates, tests, and confidence intervals in matched observational studies. *Biometrics* **63** 456–464. MR2370804 https://doi.org/10.1111/j.1541-0420.2006.00717.x
[39] ROSENBAUM, P. R. (2010). Design sensitivity and efficiency in observational studies. *J. Amer. Statist. Assoc.* **105** 692–702. MR2724853 https://doi.org/10.1198/jasa.2010.tm09570
[40] ROSENBAUM, P. R. (2012). An exact adaptive test with superior design sensitivity in an observational study of treatments for ovarian cancer. *Ann. Appl. Stat.* **6** 83–105. MR2951530 https://doi.org/10.1214/11-AOAS508
[41] ROSENBAUM, P. R. (2012). Testing one hypothesis twice in observational studies. *Biometrika* **99** 763–774. MR2999159 https://doi.org/10.1093/biomet/ass032
[42] ROSENBAUM, P. R. (2014). Weighted M-statistics with superior design sensitivity in matched observational studies with multiple controls. *J. Amer. Statist. Assoc.* **109** 1145–1158. MR3265687 https://doi.org/10.1080/01621459.2013.879261
[43] ROSENBAUM, P. R. (2015). Bahadur efficiency of sensitivity analyses in observational studies. *J. Amer. Statist. Assoc.* **110** 205–217. MR3338497 https://doi.org/10.1080/01621459.2014.960968
[44] ROSENBAUM, P. R. (2017). *Observation and Experiment*. Harvard Univ. Press, Cambridge, MA. MR3702029
[45] ROSENBAUM, P. R. (2018). Sensitivity analysis for stratified comparisons in an observational study of the effect of smoking on homocysteine levels. *Ann. Appl. Stat.* **12** 2312–2334. MR3875702 https://doi.org/10.1214/18-AOAS1153
[46] ROSENBAUM, P. R. (2020). Modern algorithms for matching in observational studies. *Annu. Rev. Stat. Appl.* **7** 143–176. MR4104189 https://doi.org/10.1146/annurev-statistics-031219-041058
[47] ROSENBAUM, P. R. (2020). *Design of Observational Studies*. Springer, Berlin.
[48] ROSENBAUM, P. R. (2021). *Replication and Evidence Factors in Observational Studies*. CRC Press/CRC, Boca Raton, FL.
[49] ROSENBAUM, P. R. (2023). Sensitivity analyses informed by tests for bias in observational studies. *Biometrics* **79**. https://doi.org/10.1111/biom.13558
[50] ROSENBAUM, P. R. and SILBER, J. H. (2009). Amplification of sensitivity analysis in matched observational studies. *J. Amer. Statist. Assoc.* **104** 1398–1405. MR2750570 https://doi.org/10.1198/jasa.2009.tm08470
[51] RUBIN, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *J. Educ. Psychol.* **66** 688–701.
[52] SCHUEMIE, M., HRIPCSAK, G., RYAN, P., MADIGAN, D. and SUCHARD, M. (2018). Empirical confidence interval calibration for population-level effect estimation studies in observational healthcare data. *Proc. Natl. Acad. Sci. USA* **115** 2571–2577.
[53] SCHWARTZ, S., LI, F. and REITER, J. P. (2012). Sensitivity analysis for unmeasured confounding in principal stratification settings with binary variables. *Stat. Med.* **31** 949–962. MR2913871 https://doi.org/10.1002/sim.4472

[54] SHI, X., MIAO, W. and TCHETGEN TCHETGEN, E. (2020). A selective review of negative control methods in epidemiology. *Curr. Epidemiol. Rep.* **7** 190–202.

[55] SUH, I., SHATEN, B. J., CUTLER, J. A. and KULLER, L. H. (1992). Alcohol use and mortality from coronary heart disease: The role of high-density lipoprotein cholesterol. The multiple risk factor intervention trial research group. *Ann. Intern. Med.* **116** 881–887. https://doi.org/10.7326/0003-4819-116-11-881

[56] SUNDARAM, R. K. (1996). *A First Course in Optimization Theory*. Cambridge Univ. Press, Cambridge. MR1402910 https://doi.org/10.1017/CBO9780511804526

[57] TARDIF, S. (1987). Efficiency and optimality results for tests based on weighted rankings. *J. Amer. Statist. Assoc.* **82** 637–644. MR0898370

[58] TCHETGEN TCHETGEN, E. J. (2014). The control outcome calibration approach for causal inference with unobserved confounding. *Amer. J. Epidemiol.* **179** 633–640.

[59] TCHETGEN TCHETGEN, E. J., YING, A., CUI, Y., SHI, X. and MIAO, W. (2020). An introduction to proximal causal learning. ArXiv Preprint. Available at arXiv:2009.10982.

[60] YU, B. and GASTWIRTH, J. L. (2005). Sensitivity analysis for trend tests: Application to the risk of radiation exposure. *Biostatistics* **6** 201–209.

[61] ZHAO, Q. (2019). On sensitivity value of pair-matched observational studies. *J. Amer. Statist. Assoc.* **114** 713–722. MR3963174 https://doi.org/10.1080/01621459.2018.1429277

Statistical Science
2023, Vol. 38, No. 3, 458–476
https://doi.org/10.1214/23-STS883

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

ARONOW, P. M. (2012). A general method for detecting interference between units in randomized experiments. *Sociol. Methods Res.* **41** 3–16. MR3190698 https://doi.org/10.1177/0049124112437535

ARONOW, P. M. and SAMII, C. (2017). Estimating average causal effects under general interference, with application to a social network experiment. *Ann. Appl. Stat.* **11** 1912–1947. MR3743283 https://doi.org/10.1214/16-AOAS1005

ATHEY, S., ECKLES, D. and IMBENS, G. W. (2018). Exact p-values for network interference. *J. Amer. Statist. Assoc.* **113** 230–240. MR3803460 https://doi.org/10.1080/01621459.2016.1241178

ATHEY, S. and IMBENS, G. W. (2022). Design-based analysis in difference-in-differences settings with staggered adoption. *J. Econometrics* **226** 62–79. MR4348786 https://doi.org/10.1016/j.jeconom.2020.10.012

BACKSTROM, L. and KLEINBERG, J. (2011). Network bucket testing. In *Proceedings of the* 20*th International Conference on World Wide Web* 615–624.

BAJARI, P., BURDICK, B., IMBENS, G. W., MASOERO, L., MCQUEEN, J., RICHARDSON, T. and ROSEN, I. M. (2021). Multiple randomization designs. arXiv preprint. Available at arXiv:2112.13495.

BASSE, G. W., FELLER, A. and TOULIS, P. (2019). Randomization tests of causal effects under interference. *Biometrika* **106** 487–494. MR3949317 https://doi.org/10.1093/biomet/asy072

BOJINOV, I., SIMCHI-LEVI, D. and ZHAO, J. (2020). Design and analysis of switchback experiments. Available at SSRN 3684168.

BOND, R. M., FARISS, C. J., JONES, J. J., KRAMER, A. D., MARLOW, C., SETTLE, J. E. and FOWLER, J. H. (2012). A 61-million-person experiment in social influence and political mobilization. *Nature* **489** 295–298.

BRANDT, A. (1938). Tests of significance in reversal or switchback trials. *Iowa Agric. Home Econ. Exp. Stat. Res. Bull.* **21** 1.

BROWN, B. W. JR. (1980). The crossover experiment for clinical trials. *Biometrics* 69–79.

COCHRAN, W. (1939). Long-term agricultural experiments. *Suppl. J. R. Stat. Soc.* **6** 104–148.

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).

COCHRAN, W. G. (1977). *Sampling Techniques*, 3rd ed. *Wiley Series in Probability and Mathematical Statistics*. Wiley, New York. MR0474575

COCHRAN, W. G. and COX, G. M. (1948). *Experimental Designs*. Wiley, New York, NY.

COOK, T. D. and DEMETS, D. L. (2007). *Introduction to Statistical Methods for Clinical Trials*. CRC Press/CRC, Boca Raton, FL.

CRÉPON, B., DUFLO, E., GURGAND, M., RATHELOT, R. and ZAMORA, P. (2013). Do labor market policies have displacement effects? Evidence from a clustered randomized experiment. *Q. J. Econ*. **128** 531–580.

FISHER, R. A. (1937). *The Design of Experiments*. Oliver & Boyd, Edinburgh.

GART, J. J. (1963). A median test with sequential application. *Biometrika* **50** 55–62. MR0156424 https://doi.org/10.1093/biomet/50.1-2.55

GASTWIRTH, J. L. (1968). The first-median test: A two-sided version of the control median test. *J. Amer. Statist. Assoc*. **63** 692–706. MR0240933

GUPTA, S., KOHAVI, R., TANG, D., XU, Y., ANDERSEN, R., BAKSHY, E., CARDIN, N., CHANDRAN, S., CHEN, N. et al. (2019). Top challenges from the first practical online controlled experiments summit. *ACM SIGKDD Explor. Newsl*. **21** 20–35.

HALLORAN, M. E., and STRUCHINER, C. J. (1991). Study Designs for Dependent Happenings. *Epidemiology*. **2** 331–338.

HECKMAN, J. J., LOCHNER, L. and TABER, C. (1998). General-equilibrium treatment effects: A study of tuition policy. *Amer. Econ. Rev*. **88** 381–386.

HEMMING, K., HAINES, T. P., CHILTON, P. J., GIRLING, A. J. and LILFORD, R. J. (2015). The stepped wedge cluster randomised trial: Rationale, design, analysis, and reporting. *BMJ* **350**.

HOLLAND, P. W. (1986). Statistics and causal inference. *J. Amer. Statist. Assoc*. **81** 945–970. MR0867618

HOLTZ, D., LOBEL, R., LISKOVICH, I. and ARAL, S. (2020). Reducing interference bias in online marketplace pricing experiments. arXiv preprint. Available at arXiv:2004.12489.

HORVITZ, D. G. and THOMPSON, D. J. (1952). A generalization of sampling without replacement from a finite universe. *J. Amer. Statist. Assoc*. **47** 663–685. MR0053460

HUDGENS, M. G. and HALLORAN, M. E. (2008). Toward causal inference with interference. *J. Amer. Statist. Assoc*. **103** 832–842. MR2435472 https://doi.org/10.1198/016214508000000292

IMAI, K., JIANG, Z. and MALANI, A. (2021). Causal inference with interference and noncompliance in two-stage randomized experiments. *J. Amer. Statist. Assoc*. **116** 632–644. MR4270009 https://doi.org/10.1080/01621459.2020.1775612

IMBENS, G. W. and RUBIN, D. B. (2015). *Causal Inference in Statistics, Social, and Biomedical Sciences*. Cambridge Univ. Press, Cambridge.

JOHARI, R., LI, H. and WEINTRAUB, G. (2020). Experimental design in two-sided platforms: An analysis of bias. arXiv preprint. Available at arXiv:2002.05670.

JONES, B. and NACHTSHEIM, C. J. (2009). Split-plot designs: What, why, and how. *J. Qual. Technol*. **41** 340–361.

KOHAVI, R., CROOK, T., LONGBOTHAM, R., FRASCA, B., HENNE, R., FERRES, J. L. and MELAMED, T. (2009). Online experimentation at Microsoft. *Data Mining Case Stud*. 11.

MANSKI, C. F. (1993). Identification of endogenous social effects: The reflection problem. *Rev. Econ. Stud*. **60** 531–542. MR1236836 https://doi.org/10.2307/2298123

MATHISEN, H. C. (1943). A method of testing the hypothesis that two samples are from the same population. *Ann. Math. Stat*. **14** 188–194. MR0009285 https://doi.org/10.1214/aoms/1177731460

MUNRO, E., WAGER, S. and XU, K. (2021). Treatment effects in market equilibrium. arXiv preprint. Available at arXiv:2109.11647.

NEYMAN, J. (1923/1990). On the application of probability theory to agricultural experiments. Essay on principles. Section 9. *Statist. Sci*. **5** 465–472.

OGBURN, E. L. and VANDERWEELE, T. J. (2014). Causal diagrams for interference. *Statist. Sci*. **29** 559–578. MR3300359 https://doi.org/10.1214/14-STS501

PAPADOGEORGOU, G., MEALLI, F. and ZIGLER, C. M. (2019). Causal inference with interfering units for cluster and population level treatment allocation programs. *Biometrics* **75** 778–787. MR4012083 https://doi.org/10.1111/biom.13049.

POLLMANN, M. (2020). Causal inference for spatial treatments. arXiv preprint, arXiv:2011.00373.

POUGET-ABADIE, J., AYDIN, K., SCHUDY, W., BRODERSEN, K. and MIRROKNI, V. (2019). Variance reduction in bipartite experiments through correlation clustering. *Adv. Neural Inf. Process. Syst*. **32**.

ROSENBAUM, P. R. (2007). Interference between units in randomized experiments. *J. Amer. Statist. Assoc*. **102** 191–200. MR2345537 https://doi.org/10.1198/016214506000001112.

RUBIN, D. B. (1978). Bayesian inference for causal effects: The role of randomization. *Ann. Statist*. **6** 34–58. MR0472152.

UGANDER, J., KARRER, B., BACKSTROM, L. and KLEINBERG, J. (2013). Graph cluster randomization: Network exposure to multiple universes. In *Proceedings of the* 19*th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, *KDD* '13 329–337. Association for Computing Machinery, New York, NY, USA.

VANDERWEELE, T. J., TCHETGEN, E. J. T. and HALLORAN, M. E. (2014). Interference and sensitivity analysis. *Statist. Sci*. **29** 687–706. MR3300366 https://doi.org/10.1214/14-STS479.

WAGER, S. and XU, K. (2021). Experimenting in equilibrium. *Manage. Sci*. https://doi.org/10.1287/mnsc.2020.3844.

WU, C. J. and HAMADA, M. S. (2011). *Experiments*: *Planning*, *Analysis*, *and Optimization* **552**. Wiley, New York.

XIONG, R., ATHEY, S., BAYATI, M. and IMBENS, G. W. (2019). Optimal experimental design for staggered rollouts. http://dx.doi.org/10.2139/ssrn.3483934.

YATES, F. (1935). Complex experiments. *Suppl. J. R. Stat. Soc*. **2** 181–247.

ZIGLER, C. M. and PAPADOGEORGOU, G. (2021). Bipartite causal inference with interference. *Statist. Sci*. **36** 109–123. MR4194206 https://doi.org/10.1214/19-STS749.

Statistical Science
2023, Vol. 38, No. 3, 477–489
https://doi.org/10.1214/23-STS885

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

BERGER, J. O. (1985). *Statistical Decision Theory and Bayesian Analysis*, 2nd ed. *Springer Series in Statistics*. Springer, New York. MR0804611 https://doi.org/10.1007/978-1-4757-4286-2

CASELLA, G. and BERGER, R. L. (2002). *Statistical Inference*, 2nd ed. Duxbury, N. Scituate.

CHEN, C. F. (1985). On asymptotic normality of limiting density functions with Bayesian implications. *J. Roy. Statist. Soc. Ser. B* **47** 540–546. MR0844485

DANIELS, M. J. and HOGAN, J. W. (2008). *Missing Data in Longitudinal Studies*: *Strategies for Bayesian Modeling and Sensitivity Analysis*. Chapman & Hall, CRC Press, New York. https://doi.org/10.1201/9781420011180

FOX, M. P., MACLEHOSE, R. F. and LASH, T. L. (2022). *Applying Quantitative Bias Analysis to Epidemiologic Data*, 2nd ed. Springer, Berlin. https://doi.org/10.1007/978-3-030-82673-4

FRANKS, A. M., D'AMOUR, A. and FELLER, A. (2020). Flexible sensitivity analysis for observational studies without observable implications. *J. Amer. Statist. Assoc.* **115** 1730–1746. MR4189753 https://doi.org/10.1080/01621459.2019.1604369

GREENLAND, S. (2003). The impact of prior distributions for uncontrolled confounding and response bias: A case study of the relation of wire codes and magnetic fields to childhood leukemia. *J. Amer. Statist. Assoc.* **98** 47–54. MR1977199 https://doi.org/10.1198/016214503388619052

GREENLAND, S. (2005). Multiple-bias modelling for analysis of observational data. *J. Roy. Statist. Soc. Ser. A* **168** 267–306. MR2119402 https://doi.org/10.1111/j.1467-985X.2004.00349.x

GUSTAFSON, P. (2007). Measurement error modelling with an approximate instrumental variable. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **69** 797–815. MR2368571 https://doi.org/10.1111/j.1467-9868.2007.00611.x

GUSTAFSON, P. (2015). *Bayesian Inference for Partially Identified Models*: *Exploring the Limits of Limited Data*. *Monographs on Statistics and Applied Probability* **141**. CRC Press, Boca Raton, FL. MR3642458

GUSTAFSON, P. (2023). Supplement to "Parameter restrictions for the sake of identification: Is there utility in asserting that perhaps a restriction holds?." https://doi.org/10.1214/23-STS885SUPP

GUSTAFSON, P. and GREENLAND, S. (2006). The performance of random coefficient regression in accounting for residual confounding. *Biometrics* **62** 760–768. MR2247204 https://doi.org/10.1111/j.1541-0420.2005.00510.x

GUSTAFSON, P. and GREENLAND, S. (2009). Interval estimation for messy observational data. *Statist. Sci.* **24** 328–342. MR2757434 https://doi.org/10.1214/09-STS305

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

GUSTAFSON, P., LE, N. D. and SASKIN, R. (2001). Case-control analysis with partial knowledge of exposure misclassification probabilities. *Biometrics* **57** 598–609. MR1855698 https://doi.org/10.1111/j.0006-341X.2001.00598.x

HOETING, J. A., MADIGAN, D., RAFTERY, A. E. and VOLINSKY, C. T. (1999). Bayesian model averaging: A tutorial (with comments by M. Clyde, David Draper and E. I. George, and a rejoinder by the authors). *Statist. Sci.* **14** 382–417. MR1765176 https://doi.org/10.1214/ss/1009212519

KASS, R. E. and RAFTERY, A. E. (1995). Bayes factors. *J. Amer. Statist. Assoc.* **90** 773–795. MR3363402 https://doi.org/10.1080/01621459.1995.10476572

KEELE, L. and QUINN, K. M. (2017). Bayesian sensitivity analysis for causal effects from 2×2 tables in the presence of unmeasured confounding with application to presidential campaign visits. *Ann. Appl. Stat.* **11** 1974–1997. MR3743285 https://doi.org/10.1214/17-AOAS1048

LASH, T. L., FOX, M. P., MACLEHOSE, R. F., MALDONADO, G., MCCANDLESS, L. C. and GREENLAND, S. (2014). Good practices for quantitative bias analysis. *Int. J. Epidemiol.* **43** 1969–1985. https://doi.org/10.1007/978-3-030-82673-4_13

LITTLE, R. J. A. and RUBIN, D. B. (2019). *Statistical Analysis with Missing Data*, *Vol.* 793, 3rd ed. Wiley, New York.

MANSKI, C. F. (2003). *Partial Identification of Probability Distributions*. *Springer Series in Statistics*. Springer, New York. MR2151380

SCHARFSTEIN, D. O., DANIELS, M. J. and ROBINS, J. M. (2003). Incorporating prior beliefs about selection bias into the analysis of randomized trials with missing outcomes. *Biostatistics* **4** 495–512. https://doi.org/10.1093/biostatistics/4.4.495

VANSTEELANDT, S., GOETGHEBEUR, E., KENWARD, M. G. and MOLENBERGHS, G. (2006). Ignorance and uncertainty regions as inferential tools in a sensitivity analysis. *Statist. Sinica* **16** 953–979. MR2281311

VERSTRAETEN, T., FARAH, B., DUCHATEAU, L. and MATU, R. (1998). Pooling sera to reduce the cost of HIV surveillance: A feasibility study in a rural Kenyan district. *Trop. Med. Int. Health* **3** 747–750. https://doi.org/10.1046/j.1365-3156.1998.00293.x

WANG, F. and GELFAND, A. E. (2002). A simulation-based approach to Bayesian sample size determination for performance under a given model and for separating models. *Statist. Sci.* **17** 193–208. MR1925941 https://doi.org/10.1214/ss/1030550861

WASSERMAN, L. (2000). Bayesian model selection and model averaging. *J. Math. Psych.* **44** 92–107. MR1770003 https://doi.org/10.1006/jmps.1999.1278

WHITE, H. (1982). Maximum likelihood estimation of misspecified models. *Econometrica* **50** 1–25. MR0640163 https://doi.org/10.2307/1912526

XIA, M. and GUSTAFSON, P. (2016). Bayesian regression models adjusting for unidirectional covariate misclassification. *Canad. J. Statist.* **44** 198–218. MR3507780 https://doi.org/10.1002/cjs.11284

XIA, M. and GUSTAFSON, P. (2018). Bayesian inference for unidirectional misclassification of a binary response trait. *Stat. Med.* **37** 933–947. MR3760458 https://doi.org/10.1002/sim.7555

Statistical Science
2023, Vol. 38, No. 3, 490–509
https://doi.org/10.1214/23-STS886

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

[1] ALQUIER, P., RIDGWAY, J. and CHOPIN, N. (2016). On the properties of variational approximations of Gibbs posteriors. *J. Mach. Learn. Res.* **17** 239. MR3595173

[2] BENNETT, J. and WAKEFIELD, J. (2001). Errors-in-variables in joint population pharmacokinetic/pharmacodynamic modeling. *Biometrics* **57** 803–812. MR1863449 https://doi.org/10.1111/j.0006-341X.2001.00803.x

[3] BISSIRI, P. G., HOLMES, C. C. and WALKER, S. G. (2016). A general framework for updating belief distributions. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **78** 1103–1130. MR3557191 https://doi.org/10.1111/rssb.12158

[4] BLANGIARDO, M., HANSELL, A. and RICHARDSON, S. (2011). A Bayesian model of time activity data to investigate health effect of air pollution in time series studies. *Atmos. Environ.* **45** 379–386.

[5] BLEI, D. M., KUCUKELBIR, A. and MCAULIFFE, J. D. (2017). Variational inference: A review for statisticians. *J. Amer. Statist. Assoc.* **112** 859–877. MR3671776 https://doi.org/10.1080/01621459.2017.1285773

[6] CARMONA, C. and NICHOLLS, G. (2020). Semi-modular inference: Enhanced learning in multi-modular models by tempering the influence of components. In *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics* (S. Chiappa and R. Calandra, eds.). *Proceedings of Machine Learning Research* **108** 4226–4235.

[7] CARMONA, C. and NICHOLLS, G. (2022). Scalable Semi-Modular Inference with Variational Meta-Posteriors. Available at arXiv:2204.00296.

[8] CARPENTER, B., GELMAN, A., HOFFMAN, M. D., LEE, D., GOODRICH, B., BETANCOURT, M., BRUBAKER, M., GUO, J., LI, P. et al. (2017). Stan: A probabilistic programming language. *J. Stat. Softw.* **76** 1–32.

[9] DEL MORAL, P., DOUCET, A. and JASRA, A. (2006). Sequential Monte Carlo samplers. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **68** 411–436. MR2278333 https://doi.org/10.1111/j.1467-9868.2006.00553.x

[10] EVANS, M. (2015). *Measuring Statistical Evidence Using Relative Belief. Monographs on Statistics and Applied Probability* **144**. CRC Press, Boca Raton, FL. MR3616661

[11] FRAZIER, D. T., LOAIZA-MAYA, R., MARTIN, G. M. and KOO, B. (2021). Loss-Based Variational Bayes Prediction. Available at arXiv:2104.14054.

[12] FRAZIER, D. T. and NOTT, D. J. (2022). Cutting feedback and modularized analyses in generalized Bayesian inference. Available at arXiv:2202.09968.

[13] GELMAN, A., MENG, X.-L. and STERN, H. (1996). Posterior predictive assessment of model fitness via realized discrepancies. *Statist. Sinica* **6** 733–807. MR1422404

[14] HAN, S., LIAO, X., DUNSON, D. and CARIN, L. (2016). Variational Gaussian copula inference. In *Proceedings of the* 19*th International Conference on Artificial Intelligence and Statistics*

Xuejun Yu is Research Fellow, Department of Paediatrics, NUS Yong Loo Lin School of Medicine, National University of Singapore, 119228, Singapore. David J. Nott is Associate Professor, Department of Statistics and Data Science, National University of Singapore, 117546, Singapore (e-mail: standj@nus.edu.sg). Michael Stanley Smith is Chair of Management (Econometrics), Melbourne Business School, University of Melbourne, 200 Leicester Street, Carlton, Victoria 3053, Australia.

(A. Gretton and C. C. Robert, eds.). *Proceedings of Machine Learning Research* **51** 829–838.

[15] JACOB, P. E., MURRAY, L. M., HOLMES, C. C. and ROBERT, C. P. (2017). Better together? Statistical learning in models made of modules. Available at arXiv:1708.08719.

[16] JACOB, P. E., O'LEARY, J. and ATCHADÉ, Y. F. (2020). Unbiased Markov chain Monte Carlo methods with couplings. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **82** 543–600. MR4112777 https://doi.org/10.1111/rssb.12336

[17] KENNEDY, M. C. and O'HAGAN, A. (2001). Bayesian calibration of computer models. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **63** 425–464. MR1858398 https://doi.org/10.1111/1467-9868.00294

[18] KNOBLAUCH, J., JEWSON, J. and DAMOULAS, T. (2022). An optimization-centric view on Bayes' rule: Reviewing and generalizing variational inference. *J. Mach. Learn. Res.* **23** 132. MR4577084

[19] KNOWLES, D. A. and MINKA, T. (2011). Non-conjugate variational message passing for multinomial and binary regression. In *Advances in Neural Information Processing Systems* 24 (J. Shawe-Taylor, R. S. Zemel, P. L. Bartlett, F. Pereira and K. Q. Weinberger, eds.) 1701–1709. Curran Associates, Red Hook.

[20] KUCUKELBIR, A., TRAN, D., RANGANATH, R., GELMAN, A. and BLEI, D. M. (2017). Automatic differentiation variational inference. *J. Mach. Learn. Res.* **18** 14. MR3634881

[21] LIN, W., KHAN, M. E. and SCHMIDT, M. (2019). Fast and simple natural-gradient variational inference with mixture of exponential-family approximations. In *Proceedings of the* 36*th International Conference on Machine Learning* (K. Chaudhuri and R. Salakhutdinov, eds.). *Proceedings of Machine Learning Research* **97** 3992–4002.

[22] LITTLE, R. J. A. (1992). Regression with missing X's: A review. *J. Amer. Statist. Assoc.* **87** 1227–1237.

[23] LIU, F., BAYARRI, M. J. and BERGER, J. O. (2009). Modularization in Bayesian analysis, with emphasis on analysis of computer models. *Bayesian Anal.* **4** 119–150. MR2486241 https://doi.org/10.1214/09-BA404

[24] LIU, Y. and GOUDIE, R. J. B. (2022). A General Framework for Cutting Feedback within Modularized Bayesian Inference. Available at arXiv:2211.03274.

[25] LIU, Y. and GOUDIE, R. J. B. (2022). Stochastic approximation cut algorithm for inference in modularized Bayesian models. *Stat. Comput.* **32** 7. MR4350200 https://doi.org/10.1007/s11222-021-10070-2

[26] LIU, Y. and GOUDIE, R. J. B. (2023). Generalized geographically weighted regression model within a modularized Bayesian framework. *Bayesian Anal.* 1–36. https://doi.org/10.1214/22-BA1357

[27] LOAIZA-MAYA, R., SMITH, M. S., NOTT, D. J. and DANAHER, P. J. (2022). Fast and accurate variational inference for models with many latent variables. *J. Econometrics* **230** 339–362. MR4466728 https://doi.org/10.1016/j.jeconom.2021.05.002

[28] LUNN, D., BEST, N., SPIEGELHALTER, D., GRAHAM, G. and NEUENSCHWANDER, B. (2009). Combining MCMC with 'sequential' PKPD modelling. *J. Pharmacokinet. Pharmacodyn.* **36** 19–38.

[29] MAUCORT-BOULCH, D., FRANCESCHI, S. and PLUMMER, M. (2008). International correlation between human papillomavirus prevalence and cervical cancer incidence. *Cancer Epidemiol. Biomark. Prev.* **17** 717–720.

[30] MCCANDLESS, L. C., DOUGLAS, I. J., EVANS, S. J. and SMEETH, L. (2010). Cutting feedback in Bayesian regression adjustment for the propensity score. *Int. J. Biostat.* **6** 16. MR2602559 https://doi.org/10.2202/1557-4679.1205

[31] MCCANDLESS, L. C., RICHARDSON, S. and BEST, N. (2012). Adjustment for missing confounders using external validation data and propensity scores. *J. Amer. Statist. Assoc.* **107** 40–51. MR2949340 https://doi.org/10.1080/01621459.2011.643739

[32] MINKA, T. (2005). Divergence measures and message passing Technical Report No. MSR-TR-2005-173 Microsoft Research.

[33] MOSS, D. and ROUSSEAU, J. (2022). Efficient Bayesian estimation and use of cut posterior in semiparametric hidden Markov models. Available at arXiv:2203.06081.

[34] MURPHY, K. M. and TOPEL, R. H. (2002). Estimation and inference in two-step econometric models. *J. Bus. Econom. Statist.* **20** 88–97. MR1940632 https://doi.org/10.1198/073500102753410417

[35] NICHOLLS, G. K., LEE, J. E., WU, C.-H. and CARMONA, C. U. (2022). Valid belief updates for prequentially additive loss functions arising in Semi-Modular inference. Available at arXiv:2201.09706.

[36] NICHOLSON, G., BLANGIARDO, M., BRIERS, M., DIGGLE, P. J., FJELDE, T. E., GE, H., GOUDIE, R. J. B., JERSAKOVA, R., KING, R. E. et al. (2022). Interoperability of statistical models in pandemic preparedness: Principles and reality. *Statist. Sci.* **37** 183–206.

[37] NOTT, D. J., WANG, X., EVANS, M. and ENGLERT, B.-G. (2020). Checking for prior-data conflict using prior-to-posterior divergences. *Statist. Sci.* **35** 234–253. MR4106603 https://doi.org/10.1214/19-STS731

[38] OGLE, K., BARBER, J. and SARTOR, K. (2013). Feedback and modularization in a Bayesian meta-analysis of tree traits affecting forest dynamics. *Bayesian Anal.* **8** 133–168. MR3036257 https://doi.org/10.1214/13-BA806

[39] ORMEROD, J. T. and WAND, M. P. (2010). Explaining variational approximations. *Amer. Statist.* **64** 140–153. MR2757005 https://doi.org/10.1198/tast.2010.09058

[40] PAPAMAKARIOS, G., NALISNICK, E., REZENDE, D. J., MOHAMED, S. and LAKSHMINARAYANAN, B. (2021). Normalizing flows for probabilistic modeling and inference. *J. Mach. Learn. Res.* **22** 57. MR4253750

[41] PETRIN, A. and TRAIN, K. (2010). A control function approach to endogeneity in consumer choice models. *J. Mark. Res.* **47** 3–13.

[42] PLUMMER, M. (2015). Cuts in Bayesian graphical models. *Stat. Comput.* **25** 37–43. MR3304902 https://doi.org/10.1007/s11222-014-9503-z

[43] POMPE, E. and JACOB, P. E. (2021). Asymptotics of cut distributions and robust modular inference using posterior bootstrap. Available at arXiv:2110.11149.

[44] PRESANIS, A. M., OHLSSEN, D., SPIEGELHALTER, D. J. and DE ANGELIS, D. (2013). Conflict diagnostics in directed acyclic graphs, with applications in Bayesian evidence synthesis. *Statist. Sci.* **28** 376–397. MR3135538 https://doi.org/10.1214/13-STS426

[45] SALIMANS, T. and KNOWLES, D. A. (2013). Fixed-form variational posterior approximation through stochastic linear regression. *Bayesian Anal.* **8** 837–881. MR3150471 https://doi.org/10.1214/13-BA858

[46] SMITH, M. S., LOAIZA-MAYA, R. and NOTT, D. J. (2020). High-dimensional copula variational approximation through transformation. *J. Comput. Graph. Statist.* **29** 729–743. MR4191239 https://doi.org/10.1080/10618600.2020.1740097

[47] STYRING, A., CHARLES, M., FANTONE, F., HALD, M., MCMAHON, A., MEADOW, R., NICHOLLS, G., PATEL, A., PITRE, M. et al. (2017). Isotope evidence for agricultural extensification reveals how the world's first cities were fed. *Nature Plants* **3** 17076.

[48] TITSIAS, M. and LÁZARO-GREDILLA, M. (2014). Doubly stochastic variational Bayes for non-conjugate inference. In *Proceedings of the* 31*st International Conference on Machine Learning* (E. P. Xing and T. Jebara, eds.). *Proceedings of Machine Learning Research* **32** 1971–1979.

[49] WAND, M. P. (2017). Fast approximate inference for arbitrarily large semiparametric regression models via message passing. *J. Amer. Statist. Assoc.* **112** 137–156. MR3646558 https://doi.org/10.1080/01621459.2016.1197833

[50] WANG, Y. and BLEI, D. M. (2019). Variational Bayes under model misspecification. In *Advances in Neural Information Processing Systems* 32: *Annual Conference on Neural Information Processing Systems* 2019, *NeurIPS* 2019, *December* 8-14, 2019, *Vancouver, BC, Canada* (H. M. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. B. Fox and R. Garnett, eds.) 13357–13367.

[51] WINN, J. and BISHOP, C. M. (2005). Variational message passing. *J. Mach. Learn. Res.* **6** 661–694. MR2249835

[52] WOODARD, D. B., CRAINICEANU, C. and RUPPERT, D. (2013). Hierarchical adaptive regression kernels for regression with functional predictors. *J. Comput. Graph. Statist.* **22** 777–800. MR3173742 https://doi.org/10.1080/10618600.2012.694765

[53] YAO, Y., VEHTARI, A., SIMPSON, D. and GELMAN, A. (2018). Yes, but did it work?: Evaluating variational inference. In *Proceedings of the* 35*th International Conference on Machine Learning* (J. Dy and A. Krause, eds.). *Proceedings of Machine Learning Research* **80** 5581–5590.

[54] YE, L., BESKOS, A., DE IORIO, M. and HAO, J. (2020). Monte Carlo co-ordinate ascent variational inference. *Stat. Comput.* **30** 887–905. MR4108683 https://doi.org/10.1007/s11222-020-09924-y

[55] YU, X., NOTT, D. J., TRAN, M.-N. and KLEIN, N. (2021). Assessment and adjustment of approximate inference algorithms using the law of total variance. *J. Comput. Graph. Statist.* **30** 977–990. MR4356599 https://doi.org/10.1080/10618600.2021.1880921

[56] ZEILER, M. D. (2012). ADADELTA: An Adaptive Learning Rate Method. Available at arXiv:1212.5701.

[57] ZHANG, L., BEAL, S. L. and SHEINER, L. B. (2003). Simultaneous vs. sequential analysis for population PK/PD data I: Best-case performance. *J. Pharmacokinet. Pharmacodyn.* **30** 387–404. https://doi.org/10.1023/b:jopa.0000012998.04442.1f

[58] ZHANG, L., BEAL, S. L. and SHEINER, L. B. (2003). Simultaneous vs. sequential analysis for population PK/PD data II: Robustness of models. *J. Pharmacokinet. Pharmacodyn.* **30** 305–416.

[59] ZIGLER, C. M., WATTS, K., YEH, R. W., WANG, Y., COULL, B. A. and DOMINICI, F. (2013). Model feedback in Bayesian propensity score estimation. *Biometrics* **69** 263–273. MR3058073 https://doi.org/10.1111/j.1541-0420.2012.01830.x

Statistical Science
2023, Vol. 38, No. 3, 510–513
https://doi.org/10.1214/23-STS887

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 spheroïdique tirés de la méthode des plus grandes et plus petits. *Mémoires de Berlin* 1753 **IX** 258–293.

HALD, A. (1998). *A History of Mathematical Statistics from* 1750 *to* 1930. *Wiley Series in Probability and Statistics*: *Texts and References Section*. Wiley, New York. MR1619032

LAPLACE, P. S. (1799). *Traité de Mécanique Céleste* **2**. Duprat, Paris.

LEGENDRE, A. M. (1805). *Nouvelles méthodes pour la détermination des orbites des comètes*. Firmin Didot, Paris.

STIGLER, S. M. (1981). Gauss and the invention of least squares. *Ann. Statist.* **9** 465–474. MR0615423

STIGLER, S. M. (1986). *The History of Statistics*: *The Measurement of Uncertainty Before* 1900. The Belknap Press of Harvard Univ. Press, Cambridge, MA. MR0852410

TORGE, W. (2001). *Geodesy*. de Gruyter, Berlin.

Jukka Nyblom is Professor Emeritus of Statistics, Department of Mathematics and Statistics, University of Jyväskylä, P.O. Box 35, FI-40014, Jyväskylä, Finland (e-mail: jukka.nyblom@jyu.fi).

Statistical Science
2023, Vol. 38, No. 3, 514–524
https://doi.org/10.1214/22-STS877

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

[1] BINDER, D. A. (1983). On the variances of asymptotically normal estimators from complex surveys. *Int. Stat. Rev.* **51** 279–292. MR0731144 https://doi.org/10.2307/1402588

[2] BIRNBAUM, A. (1962). On the foundations of statistical inference. *J. Amer. Statist. Assoc.* **57** 269–326. MR0138176

[3] FISHER, R. A. (1956). *Statistical Methods and Scientific Inference*. Oliver and Boyd, Edinburgh, UK.

[4] GODAMBE, V. P. (1955). A unified theory of sampling from finite populations. *J. Roy. Statist. Soc. Ser. B* **17** 269–278. MR0077037

[5] GODAMBE, V. P. (1966). A new approach to sampling from finite populations. I. Sufficiency and linear estimation. *J. Roy. Statist. Soc. Ser. B* **28** 310–319. MR0216720

[6] GODAMBE, V. P. and THOMPSON, M. E. (1971). Bayes, fiducial and frequency aspects of statistical inference in regression analysis in survey-sampling. *J. Roy. Statist. Soc. Ser. B* **33** 361–390. MR0362603

[7] GODAMBE, V. P. and THOMPSON, M. E. (1986). Parameters of superpopulation and survey population: Their relationships and estimation. *Int. Stat. Rev.* **54** 127–138. MR0962931 https://doi.org/10.2307/1403139

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).

[8] GODAMBE, V. P. and THOMPSON, M. E. (1989). An extension of quasi-likelihood estimation. *J. Statist. Plann. Inference* **22** 137–172. MR1004344 https://doi.org/10.1016/0378-3758(89)90106-7

[9] JIANG, C., WALLACE, M. P. and THOMPSON, M. E. (2022). Dynamic treatment regimes with interference. *Canad. J. Statist.*. To appear. https://doi.org/10.1002/cjs.11702

[10] RAMÍREZ-RAMÍREZ, L. L. and THOMPSON, M. E. (2014). Applications of the variance of final outbreak size for disease spreading in networks. *Methodol. Comput. Appl. Probab.* **16** 839–862. MR3270598 https://doi.org/10.1007/s11009-013-9325-z

[11] THOMPSON, M. E. (1984). Model and design correspondence in finite population sampling. *J. Statist. Plann. Inference* **10** 323–334. MR0766648 https://doi.org/10.1016/0378-3758(84)90057-0

[12] THOMPSON, M. E. (1997). *Theory of Sample Surveys*. *Monographs on Statistics and Applied Probability* **74**. CRC Press, London. MR1462619 https://doi.org/10.1002/1097-0258(20000715)19:13<1825::AID-SIM466>3.0.CO;2-E

[13] THOMPSON, M. E. (2001). Likelihood principle and randomization in survey sampling. In *Data Analysis from Statistical Foundations* 9–25. Nova Sci. Publ., Huntington, NY. MR2034504

[14] THOMPSON, M. E., RAMIREZ RAMIREZ, L. L., LYUBCHICH, V. and GEL, Y. R. (2016). Using the bootstrap for statistical inference on random graphs. *Canad. J. Statist.* **44** 3–24. MR3474218 https://doi.org/10.1002/cjs.11271

[15] THOMPSON, M. E., SEDRANSK, J., FANG, J. and YI, G. Y. (2022). Bayesian inference for a variance component model using pairwise composite likelihood with survey data. *Surv. Methodol.* **48** 73–93.

[16] THOMPSON, M. E. B. (1969). *Some Aspects of Optimal Stopping Theory*. ProQuest LLC, Ann Arbor, MI. Thesis (Ph.D.)—University of Illinois at Urbana-Champaign. MR2619132

[17] WU, C. and THOMPSON, M. E. (2020). *Sampling Theory and Practice*. *ICSA Book Series in Statistics*. Springer, Cham. MR4180686 https://doi.org/10.1007/978-3-030-44246-0

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: Michael Kosorok, Department of Biostatistics and Department of Statistics and Operations Research, University of North Carolina, Chapel Hill, Chapel Hill, NC 27599, USA

President-Elect: Tony Cai, Department of Statistics and Data Science, University of Pennsylvania, Philadelphia, PA 19104-6304, USA

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

Executive Secretary: Peter Hoff, Department of Statistical Science, Duke University, Durham, NC 27708-0251, USA

Treasurer: Jiashun Jin, Department of Statistics, Carnegie Mellon University, Pittsburgh, PA 15213-3890, 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 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*: Kavita Ramanan, Division of Applied Mathematics, Brown University, Providence, RI 02912, USA. Qi-Man Shao, Department of Statistics and Data Science, Southern University of Science and Technology, Shenzhen, Guangdong 518055, P.R. China

Statistical Science. *Editor*: Moulinath Banerjee, Department of Statistics, University of Michigan, Ann Arbor, MI 48109, USA

The IMS Bulletin. *Editor*: Vlada Limic, UMR 7501 de l'Université de Strasbourg et du CNRS, 7 rue René Descartes, 67084 Strasbourg Cedex, France