STATISTICAL SCIENCE

Volume 37, Number 1 February 2022

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

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

Statistical Science Volume 37, Number 1 (1–158) February 2022

Volume 37 **Number 1** **February 2022**

Statistical Science
2022, Vol. 37, No. 1, 1–23
https://doi.org/10.1214/20-STS818

Bayesian Analysis of Rank Data with Covariates and Heterogeneous Rankers

Xinran Li, Dingdong Yi and Jun S. Liu

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

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

REFERENCES

AGRESTI, A. (1990). *Categorical Data Analysis*. *Wiley Series in Probability and Mathematical Statistics*: *Applied Probability and Statistics*. Wiley, New York. MR1044993

AILON, N. (2010). Aggregation of partial rankings, p-ratings and top-m lists. *Algorithmica* **57** 284–300. MR2587555 https://doi.org/10.1007/s00453-008-9211-1

ALBERT, J. H. and CHIB, S. (1993). Bayesian analysis of binary and polychotomous response data. *J. Amer. Statist. Assoc.* **88** 669–679. MR1224394

ALLISON, P. D. and CHRISTAKIS, N. A. (1994). Logit models for sets of ranked items. *Sociol. Method.* **24** 199–228.

ALVO, M. and YU, P. L. H. (2014). *Statistical Methods for Ranking Data*. *Frontiers in Probability and the Statistical Sciences*. Springer, New York. MR3308929 https://doi.org/10.1007/978-1-4939-1471-5

ANTONIAK, C. E. (1974). Mixtures of Dirichlet processes with applications to Bayesian nonparametric problems. *Ann. Statist.* **2** 1152–1174. MR0365969

AZARI SOUFIANI, H. (2014). *Revisiting Random Utility Models*. Ph.D. thesis, Harvard Univ.

AZARI SOUFIANI, H., CHEN, W., PARKES, D. C. and XIA, L. (2013). Generalized method-of-moments for rank aggregation. In *Advances in Neural Information Processing Systems* (C. J. C. Burges, L. Bottou, M. Welling, Z. Ghahramani and K. Q. Weinberger, eds.) **26** 2706–2714. Curran Associates, Red Hook, NY.

BADGELEY, M. A., CHIKINA, M. D. and SEALFON, S. C. (2014). Hybrid Bayesian-rank integration approach improves the predictive power of genomic dataset aggregation. *Bioinformatics* **31** 209–215.

BENTER, W. (1994). Computer-based horse race handicapping and wagering systems: A report. In *Efficiency of Racetrack Betting Markets* (W. T. Ziemba, V. S. Lo and D. B. Haush, eds.) 183–198. Academic Press, London.

BHOWMIK, A. and GHOSH, J. (2017). LETOR methods for unsupervised rank aggregation. In *Proceedings of the* 26*th International Conference on World Wide Web* 1331–1340. International World Wide Web Conferences Steering Committee.

BLACKWELL, D. and MACQUEEN, J. B. (1973). Ferguson distributions via Pólya urn schemes. *Ann. Statist.* **1** 353–355. MR0362614

BLOCK, H. D. and MARSCHAK, J. (1960). Random orderings and stochastic theories of responses. In *Contributions to Probability and Statistics* 97–132. Stanford Univ. Press, Stanford, CA.

Xinran Li is Assistant Professor, Department of Statistics, University of Illinois, Champaign, Illinois 61820, USA (e-mail: xinranli@illinois.edu). Dingdong Yi is Quantitative Researcher, Citadel Americas LLC, New York, New York 10022, USA (e-mail: yidingdong@gmail.com). Jun S. Liu is Professor, Department of Statistics, Harvard University, Cambridge, Massachusetts 02138, USA (e-mail: jliu@stat.harvard.edu).

MR0121236
BÖCKENHOLT, U. (1992). Thurstonian representation for partial ranking data. *Br. J. Math. Stat. Psychol.* **45** 31–49.
BÖCKENHOLT, U. (1993). Applications of Thurstonian models to ranking data. In *Probability Models and Statistical Analyses for Ranking Data* (*Amherst, MA*, 1990) (M. A. M. A. and J. S. J. S., eds.). *Lect. Notes Stat.* **80** 157–172. Springer, New York. MR1237206 https://doi.org/10.1007/978-1-4612-2738-0_9
BÖCKENHOLT, U. (2006). Thurstonian-based analyses: Past, present, and future utilities. *Psychometrika* **71** 615–629. MR2312235 https://doi.org/10.1007/s11336-006-1598-5
BORDA, J. C. (1781). Mémoire sur les élections au scrutin.
BRADLEY, R. A. and TERRY, M. E. (1952). Rank analysis of incomplete block designs. I. The method of paired comparisons. *Biometrika* **39** 324–345. MR0070925 https://doi.org/10.2307/2334029
CARON, F. and DOUCET, A. (2012). Efficient Bayesian inference for generalized Bradley–Terry models. *J. Comput. Graph. Statist.* **21** 174–196. MR2913362 https://doi.org/10.1080/10618600.2012.638220
CRITCHLOW, D. E., FLIGNER, M. A. and VERDUCCI, J. S. (1991). Probability models on rankings. *J. Math. Psych.* **35** 294–318. MR1128236 https://doi.org/10.1016/0022-2496(91)90050-4
D'AMBROSIO, A., IORIO, C., STAIANO, M. and SICILIANO, R. (2019). Median constrained bucket order rank aggregation. *Comput. Statist.* **34** 787–802. MR3953682 https://doi.org/10.1007/s00180-018-0858-z
DANIELS, H. E. (1950). Rank correlation and population models. *J. Roy. Statist. Soc. Ser. B* **12** 171–181. MR0040629
DAVID, H. A. (1963). *The Method of Paired Comparisons*, **12**. Oxford University Press, Oxford.
DAVIDSON, R. R. and FARQUHAR, P. H. (1976). A bibliography on the method of paired comparisons. *Biometrics* **32** 241–252. MR0408134
DECONDE, R. P., HAWLEY, S., FALCON, S., CLEGG, N., KNUDSEN, B. and ETZIONI, R. (2006). Combining results of microarray experiments: A rank aggregation approach. *Stat. Appl. Genet. Mol. Biol.* **5** Art. 15, 25. MR2240846 https://doi.org/10.2202/1544-6115.1204
DENG, K., HAN, S., LI, K. J. and LIU, J. S. (2014). Bayesian aggregation of order-based rank data. *J. Amer. Statist. Assoc.* **109** 1023–1039. MR3265678 https://doi.org/10.1080/01621459.2013.878660
DIACONIS, P. (1988). *Group Representations in Probability and Statistics. Institute of Mathematical Statistics Lecture Notes—Monograph Series* **11**. IMS, Hayward, CA. MR0964069
DWORK, C., KUMAR, R., NAOR, M. and SIVAKUMAR, D. (2001). Rank aggregation methods for the web. In *Proceedings of the* 10*th International Conference on World Wide Web* 613–622. ACM, New York.
ESCOBAR, M. D. and WEST, M. (1995). Bayesian density estimation and inference using mixtures. *J. Amer. Statist. Assoc.* **90** 577–588. MR1340510
FERGUSON, T. S. (1983). Bayesian density estimation by mixtures of normal distributions. In *Recent Advances in Statistics* 287–302. Academic Press, New York. MR0736538
FONG, D. K. H., KIM, S., CHEN, Z. and DESARBO, W. S. (2016). A Bayesian multinomial probit model for the analysis of panel choice data. *Psychometrika* **81** 161–183. MR3463498 https://doi.org/10.1007/s11336-014-9437-6
FRÜHWIRTH-SCHNATTER, S. and MALSINER-WALLI, G. (2019). From here to infinity: Sparse finite versus Dirichlet process mixtures in model-based clustering. *Adv. Data Anal. Classif.* **13** 33–64. MR3935190 https://doi.org/10.1007/s11634-018-0329-y
GIONIS, A., MANNILA, H., PUOLAMÄKI, K. and UKKONEN, A. (2006). Algorithms for discovering bucket orders from data. In *Proceedings of the* 12*th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 561–566.
GORMLEY, I. C. and MURPHY, T. B. (2006). Analysis of Irish third-level college applications data. *J. Roy. Statist. Soc. Ser. A* **169** 361–379. MR2225548 https://doi.org/10.1111/j.1467-985X.2006.00412.x
GORMLEY, I. C. and MURPHY, T. B. (2008a). Exploring voting blocs within the Irish electorate: A mixture modeling approach. *J. Amer. Statist. Assoc.* **103** 1014–1027. MR2528824 https://doi.org/10.1198/016214507000001049
GORMLEY, I. C. and MURPHY, T. B. (2008b). A mixture of experts model for rank data with applications in election studies. *Ann. Appl. Stat.* **2** 1452–1477. MR2655667 https://doi.org/10.1214/08-AOAS178
GORMLEY, I. C. and MURPHY, T. B. (2009). A grade of membership model for rank data. *Bayesian Anal.* **4** 265–295. MR2507364 https://doi.org/10.1214/09-BA410
GORMLEY, I. C. and MURPHY, T. B. (2010). Clustering ranked preference data using sociodemographic covariates. In *Choice Modelling*: *The State-of-the-Art and the State-of-Practice*: *Proceedings from the Inaugural International Choice Modelling Conference* 543–569.
GRAY-DAVIES, T., HOLMES, C. C. and CARON, F. (2016). Scalable Bayesian nonparametric regression via a Plackett–Luce model for conditional ranks. *Electron. J. Stat.* **10** 1807–1828. MR3522661 https://doi.org/10.1214/15-EJS1032
GUIVER, J. and SNELSON, E. (2009). Bayesian inference for Plackett–Luce ranking models. In *Proceedings of the* 26*th Annual International Conference on Machine Learning* 377–384. ACM, New York.
HASTIE, T. and TIBSHIRANI, R. (1998). Classification by pairwise coupling. *Ann. Statist.* **26** 451–471. MR1626055 https://doi.org/10.1214/aos/1028144844
HAUSMAN, J. A. and RUUD, P. A. (1987). Specifying and testing econometric models for rank-ordered data. *J. Econometrics* **34** 83–104. MR0887389 https://doi.org/10.1016/0304-4076(87)90068-6
HOFF, P. D. (2009). *A First Course in Bayesian Statistical Methods. Springer Texts in Statistics*. Springer, New York. MR2648134 https://doi.org/10.1007/978-0-387-92407-6
HUANG, T.-K., WENG, R. C. and LIN, C.-J. (2006). Generalized Bradley–Terry models and multi-class probability estimates. *J. Mach. Learn. Res.* **7** 85–115. MR2274363
HUNTER, D. R. (2004). MM algorithms for generalized Bradley–Terry models. *Ann. Statist.* **32** 384–406. MR2051012 https://doi.org/10.1214/aos/1079120141
JOHNSON, V. E., DEANER, R. O. and VAN SCHAIK, C. P. (2002). Bayesian analysis of rank data with application to primate intelligence experiments. *J. Amer. Statist. Assoc.* **97** 8–17. MR1947270 https://doi.org/10.1198/016214502753479185
JOHNSON, T. R. and KUHN, K. M. (2013). Bayesian thurstonian models for ranking data using jags. *Behavior Research Methods* **45** 857–872.
KENDALL, M. G. (1938). A new measure of rank correlation. *Biometrika* 81–93.
KENKRE, S., KHAN, A. and PANDIT, V. (2011). On discovering bucket orders from preference data. In *Proceedings of the* 2011 *SIAM International Conference on Data Mining* 872–883. SIAM, Philadelphia.
LEE, M. D., STEYVERS, M. and MILLER, B. (2014). A cognitive model for aggregating people's rankings. *PLoS ONE* **9** 1–9.
LI, X., WANG, X. and XIAO, G. (2017). A comparative study of rank aggregation methods for partial and top ranked lists in genomic applications. *Brief. Bioinform.* **20** 178–189.

LI, X., YI, D. and LIU, J. S (2022). Supplement to "Bayesian Analysis of Rank Data with Covariates and Heterogeneous Rankers." https://doi.org/10.1214/20-STS818SUPP

LI, X., CHOUDHARY, P. K., BISWAS, S. and WANG, X. (2018). A Bayesian latent variable approach to aggregation of partial and top-ranked lists in genomic studies. *Stat. Med.* **37** 4266–4278. MR3879427 https://doi.org/10.1002/sim.7920

LI, H., XU, M., LIU, J. S. and FAN, X. (2020). An extended mallows model for ranked data aggregation. *J. Amer. Statist. Assoc.* **115** 730–746. MR4107676 https://doi.org/10.1080/01621459.2019.1573733

LIN, S. (2010). Rank aggregation methods. *Wiley Interdiscip. Rev.: Comput. Stat.* **2** 555–570.

LIN, S. and DING, J. (2009). Integration of ranked lists via cross entropy Monte Carlo with applications to mRNA and microRNA studies. *Biometrics* **65** 9–18. MR2665841 https://doi.org/10.1111/j.1541-0420.2008.01044.x

LIU, J. S. (1994). The collapsed Gibbs sampler in Bayesian computations with applications to a gene regulation problem. *J. Amer. Statist. Assoc.* **89** 958–966. MR1294740

LIU, T.-Y. (2009). Learning to rank for information retrieval. *Foundations and Trends in Information Retrieval* **3** 225–331.

LIU, J. S. and WU, Y. N. (1999). Parameter expansion for data augmentation. *J. Amer. Statist. Assoc.* **94** 1264–1274. MR1731488 https://doi.org/10.2307/2669940

LIU, S.-Q., SHEN, G., BAI, D., ZHOU, H., LI, S., CHEN, W.-J., WANG, D.-W., LI, W.-R., GENG, Z. et al. (2012). Consistency of the subjective evaluation of malocclusion severity by the Chinese orthodontic experts. *Journal of Peking University. Health Sciences* **44** 98–102.

LIU, A., ZHAO, Z., LIAO, C., LU, P. and XIA, L. (2019). Learning Plackett–Luce mixtures from partial preferences. In *Proceedings of the Thirty-Third AAAI Conference on Artificial Intelligence* (*AAAI-19*).

LUCE, R. D. (1959). *Individual Choice Behavior: A Theoretical Analysis*. Wiley, New York. MR0108411

MACEACHERN, S. N. (1994). Estimating normal means with a conjugate style Dirichlet process prior. *Comm. Statist. Simulation Comput.* **23** 727–741. MR1293996 https://doi.org/10.1080/03610919408813196

MALLOWS, C. L. (1957). Non-null ranking models. I. *Biometrika* **44** 114–130. MR0087267 https://doi.org/10.1093/biomet/44.1-2.114

MARDEN, J. I. (1995). *Analyzing and Modeling Rank Data*. *Monographs on Statistics and Applied Probability* **64**. CRC Press, London. MR1346107

MAYDEU-OLIVARES, A. (1999). Thurstonian modeling of ranking data via mean and covariance structure analysis. *Psychometrika* **64** 325–340. MR1720087 https://doi.org/10.1007/BF02294299

MCFADDEN, D. (1980). Econometric models for probabilistic choice among products. *J. Bus.* **53** S13–S29.

MEILA, M. and CHEN, H. (2012). Dirichlet process mixtures of generalized mallows models. Preprint. Available at arXiv:1203.3496.

MOSTELLER, F. (1951). Remarks on the method of paired comparisons: I. the least squares solution assuming equal standard deviations and equal correlations. *Psychometrika* **16** 3–9.

NEAL, R. M. (1992). Bayesian mixture modeling. In *Maximum Entropy and Bayesian Methods* 197–211. Springer, Berlin.

NEAL, R. M. (2000). Markov chain sampling methods for Dirichlet process mixture models. *J. Comput. Graph. Statist.* **9** 249–265. MR1823804 https://doi.org/10.2307/1390653

PLACKETT, R. L. (1975). The analysis of permutations. *J. R. Stat. Soc. Ser. C. Appl. Stat.* **24** 193–202. MR0391338 https://doi.org/10.2307/2346567

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

RAO, P. V. and KUPPER, L. L. (1967). Ties in paired-comparison experiments: A generalization of the Bradley–Terry model. *J. Amer. Statist. Assoc.* **62** 194–204. MR0217963

RICHMOND, S., SHAW, W. C., O'BRIEN, K. D., BUCHANAN, I. B., JONES, R., STEPHENS, C. D., ROBERTS, C. T. and ANDREWS, M. (1992). The development of the par index (peer assessment rating): Reliability and validity. *The European Journal of Orthodontics* **14** 125–139.

SCHIMEK, M. G., BUDINSKÁ, E., KUGLER, K. G., ŠVENDOVÁ, V., DING, J. and LIN, S. (2015). TopKLists: A comprehensive R package for statistical inference, stochastic aggregation, and visualization of multiple omics ranked lists. *Stat. Appl. Genet. Mol. Biol.* **14** 311–316. MR3353637 https://doi.org/10.1515/sagmb-2014-0093

SONG, G.-Y., ZHAO, Z.-H., DING, Y., BAI, Y.-X., WANG, L., HE, H., SHEN, G., LI, W.-R., BAUMRIND, S. et al. (2014). Reliability assessment and correlation analysis of evaluating orthodontic treatment outcome in Chinese patients. *International Journal of Oral Science* **6** 50–55.

SONG, G.-Y., JIANG, R.-P., ZHANG, X.-Y., LIU, S.-Q., YU, X.-N., CHEN, Q., WENG, X.-R., WU, W.-Z., SU, H. et al. and RESEARCH GROUP OF ESTABLISHING CHINESE EVALUATION STANDARD OF ORTHODONTIC TREATMENT OUTCOME (2015). Validation of subjective and objective evaluation methods for orthodontic treatment outcome. *Journal of Peking University. Health Sciences* **47** 90–97.

STERN, H. (1990). Models for distributions on permutations. *J. Amer. Statist. Assoc.* **85** 558–564.

ŠVENDOVÁ, V. and SCHIMEK, M. G. (2017). A novel method for estimating the common signals for consensus across multiple ranked lists. *Comput. Statist. Data Anal.* **115** 122–135. MR3683133 https://doi.org/10.1016/j.csda.2017.05.010

TANNER, M. A. and WONG, W. H. (1987). The calculation of posterior distributions by data augmentation. *J. Amer. Statist. Assoc.* **82** 528–550. With discussion and with a reply by the authors. MR0898357

THURSTONE, L. L. (1927). A law of comparative judgment. *Psychological Review* **34** 273.

VAN ERP, M. and SCHOMAKER, L. (2000). Variants of the borda count method for combining ranked classifier hypotheses. In *Proceedings of the 7th International Workshop on Frontiers in Handwriting Recognition*.

VITELLI, V., SØRENSEN, Ø., CRISPINO, M., FRIGESSI, A. and ARJAS, E. (2017). Probabilistic preference learning with the Mallows rank model. *J. Mach. Learn. Res.* **18** Paper No. 158, 49. MR3813807

WALKER, J. and BEN-AKIVA, M. (2002). Generalized random utility model. *Math. Social Sci.* **43** 303–343. MR2072961 https://doi.org/10.1016/S0165-4896(02)00023-9

XIA, L. and CONITZER, V. (2011). A maximum likelihood approach towards aggregating partial orders. In *Twenty-Second International Joint Conference on Artificial Intelligence*.

YAO, G. and BÖCKENHOLT, U. (1999). Bayesian estimation of thurstonian ranking models based on the Gibbs sampler. *Br. J. Math. Stat. Psychol.* **52** 79–92.

YU, P. L. H. (2000). Bayesian analysis of order-statistics models for ranking data. *Psychometrika* **65** 281–299. https://doi.org/10.1007/BF02296147

Statistical Science
2022, Vol. 37, No. 1, 24–41
https://doi.org/10.1214/20-STS819

The Dependent Dirichlet Process and Related Models

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

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

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

REFERENCES

ALDOUS, D. J. (1985). Exchangeability and related topics. In *École D'été de Probabilités de Saint-Flour, XIII*—1983. *Lecture Notes in Math*. **1117** 1–198. Springer, Berlin. MR0883646 https://doi.org/10.1007/BFb0099421

ARGIENTO, R., GUGLIELMI, A. and PIEVATOLO, A. (2010). Bayesian density estimation and model selection using nonparametric hierarchical mixtures. *Comput. Statist. Data Anal.* **54** 816–832. MR2580918 https://doi.org/10.1016/j.csda.2009.11.002

ASCOLANI, F., LIJOI, A. and RUGGIERO, M. (2021). Predictive inference with Fleming–Viot-driven dependent Dirichlet processes. *Bayesian Anal.* Advance publication. https://doi.org/10.1214/20-BA1206

BARNDORFF-NIELSEN, O. (1973). On M-ancillarity. *Biometrika* **60** 447–455. MR0345255 https://doi.org/10.1093/biomet/60.3.447

BARNDORFF-NIELSEN, O. (1978). *Information and Exponential Families in Statistical Theory*. *Wiley Series in Probability and Mathematical Statistics*. Wiley, Chichester. MR0489333

BARRIENTOS, A. F., JARA, A. and QUINTANA, F. A. (2012). On the support of MacEachern's dependent Dirichlet processes and extensions. *Bayesian Anal.* **7** 277–309. MR2934952 https://doi.org/10.1214/12-BA709

BARRIOS, E., LIJOI, A., NIETO-BARAJAS, L. E. and PRÜNSTER, I. (2013). Modeling with normalized random measure mixture models. *Statist. Sci.* **28** 313–334. MR3135535 https://doi.org/10.1214/13-STS416

BERAHA, M., GUGLIELMI, A. and QUINTANA, F. A. (2020). The semi-hierarchical Dirichlet process and its application to clustering homogeneous distributions. Preprint. Available at arXiv:2005.10287.

CAMERLENGHI, F., DUNSON, D. B., LIJOI, A., PRÜNSTER, I. and RODRÍGUEZ, A. (2019a). Latent nested nonparametric priors (with discussion). *Bayesian Anal.* **14** 1303–1356. MR4044854 https://doi.org/10.1214/19-BA1169

CAMERLENGHI, F., LIJOI, A., ORBANZ, P. and PRÜNSTER, I. (2019b). Distribution theory for hierarchical processes. *Ann. Statist.* **47** 67–92. MR3909927 https://doi.org/10.1214/17-AOS1678

CAMPBELL, T., SYED, S., YANG, C.-Y., JORDAN, M. I. and BRODERICK, T. (2019). Local exchangeability. Preprint. Available at arXiv:1906.09507.

Fernando A. Quintana is Professor, Departamento de Estadística, Pontificia Universidad Católica de Chile, Santiago, Chile, and Deputy Director, ANID—Millennium Science Initiative Program—Millennium Nucleus Center for the Discovery of Structures in Complex Data, Santiago, Chile (e-mail: quintana@mat.uc.cl). Peter Müller is Professor, Department of Statistics and Data Science, University of Texas at Austin, Austin, Texas, USA (e-mail: pmueller@math.utexas.edu). Alejandro Jara is Associate Professor, Departamento de Estadística, Pontificia Universidad Católica de Chile, Santiago, Chile, and Director, ANID—Millennium Science Initiative Program—Millennium Nucleus Center for the Discovery of Structures in Complex Data, Santiago, Chile (e-mail: atjara@uc.cl). Steven N. MacEachern is Professor, Department of Statistics, Ohio State University, Columbus, Ohio, USA (e-mail: snm@stat.osu.edu).

CARON, F., DAVY, M., DOUCET, A., DUFLOS, E. and VANHEEGHE, P. (2008). Bayesian inference for linear dynamic models with Dirichlet process mixtures. *IEEE Trans. Signal Process.* **56** 71–84. MR2439814 https://doi.org/10.1109/TSP.2007.900167

CHEN, C., DING, N. and BUNTINE, W. (2012). Dependent hierarchical normalized random measures for dynamic topic modeling. In *Proceedings of the* 29*th International Conference on Machine Learning* (*ICML*-12) (J. Langford and J. Pineau, eds.). *ICML* '12 895–902. Omnipress, New York.

CHEN, C., RAO, V., BUNTINE, W. and TEH, Y. W. (2013). Dependent normalized random measures. In *Proceedings of the* 30*th International Conference on Machine Learning* (S. Dasgupta and D. McAllester, eds.). *Proceedings of Machine Learning Research* **28** 969–977. PMLR, Atlanta, GA.

CHIPMAN, H. A., GEORGE, E. I. and MCCULLOCH, R. E. (2010). BART: Bayesian additive regression trees. *Ann. Appl. Stat.* **4** 266–298. MR2758172 https://doi.org/10.1214/09-AOAS285

CHUNG, Y. and DUNSON, D. B. (2009). Nonparametric Bayes conditional distribution modeling with variable selection. *J. Amer. Statist. Assoc.* **104** 1646–1660. MR2750582 https://doi.org/10.1198/jasa.2009.tm08302

CIFARELLI, D. and REGAZZINI, E. (1978). Problemi statistici non parametrici in condizioni di scambialbilita parziale e impiego di medie associative. Technical report. Quaderni Istituto Matematica Finanziaria, Torino.

DE IORIO, M., MÜLLER, P., ROSNER, G. L. and MACEACHERN, S. N. (2004). An ANOVA model for dependent random measures. *J. Amer. Statist. Assoc.* **99** 205–215. MR2054299 https://doi.org/10.1198/016214504000000205

DE IORIO, M., JOHNSON, W. O., MÜLLER, P. and ROSNER, G. L. (2009). Bayesian nonparametric nonproportional hazards survival modeling. *Biometrics* **65** 762–771. MR2649849 https://doi.org/10.1111/j.1541-0420.2008.01166.x

DEVROYE, L. (1986). *Nonuniform Random Variate Generation*. Springer, New York. MR0836973 https://doi.org/10.1007/978-1-4613-8643-8

DE LA CRUZ-MESÍA, R., QUINTANA, F. A. and MÜLLER, P. (2007). Semiparametric Bayesian classification with longitudinal markers. *J. R. Stat. Soc. Ser. C. Appl. Stat.* **56** 119–137. MR2359237 https://doi.org/10.1111/j.1467-9876.2007.00569.x

DI LUCCA, M. A., GUGLIELMI, A., MÜLLER, P. and QUINTANA, F. A. (2013). A simple class of Bayesian nonparametric autoregression models. *Bayesian Anal.* **8** 63–87. MR3036254 https://doi.org/10.1214/13-BA803

DUAN, J. A., GUINDANI, M. and GELFAND, A. E. (2007). Generalized spatial Dirichlet process models. *Biometrika* **94** 809–825. MR2416794 https://doi.org/10.1093/biomet/asm071

DUNSON, D. B. and HERRING, A. H. (2006). Semiparametric Bayesian latent trajectory models. Technical report. ISDS Discussion Paper 16, Duke Univ., Durham, NC.

DUNSON, D. B. and PARK, J.-H. (2008). Kernel stick-breaking processes. *Biometrika* **95** 307–323. MR2521586 https://doi.org/10.1093/biomet/asn012

DUNSON, D. B., PILLAI, N. and PARK, J.-H. (2007). Bayesian density regression. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **69** 163–183. MR2325270 https://doi.org/10.1111/j.1467-9868.2007.00582.x

EPIFANI, I. and LIJOI, A. (2010). Nonparametric priors for vectors of survival functions. *Statist. Sinica* **20** 1455–1484. MR2777332

FARAWAY, J. J. (2016). *Extending the Linear Model with R*: *Generalized Linear, Mixed Effects and Nonparametric Regression Models*. *Chapman & Hall/CRC Texts in Statistical Science Series*. CRC Press, Boca Raton, FL. MR3617626

FAVARO, S. and TEH, Y. W. (2013). MCMC for normalized random measure mixture models. *Statist. Sci.* **28** 335–359. MR3135536 https://doi.org/10.1214/13-STS422

FERGUSON, T. S. (1973). A Bayesian analysis of some nonparametric problems. *Ann. Statist.* **1** 209–230. MR0350949

FERGUSON, T. S. (1974). Prior distributions on spaces of probability measures. *Ann. Statist.* **2** 615–629. MR0438568

FUENTES-GARCÍA, R., MENA, R. H. and WALKER, S. G. (2009). A nonparametric dependent process for Bayesian regression. *Statist. Probab. Lett.* **79** 1112–1119. MR2510777 https://doi.org/10.1016/j.spl.2009.01.005

GELFAND, A. E. and KOTTAS, A. (2001). Nonparametric Bayesian modeling for stochastic order. *Ann. Inst. Statist. Math.* **53** 865–876. MR1880817 https://doi.org/10.1023/A:1014629724913

GELFAND, A. E., KOTTAS, A. and MACEACHERN, S. N. (2005). Bayesian nonparametric spatial modeling with Dirichlet process mixing. *J. Amer. Statist. Assoc.* **100** 1021–1035. MR2201028 https://doi.org/10.1198/016214504000002078

GIUDICI, P., MEZZETTI, M. and MULIERE, P. (2003). Mixtures of products of Dirichlet processes for variable selection in survival analysis. *J. Statist. Plann. Inference* **111** 101–115. MR1955875 https://doi.org/10.1016/S0378-3758(02)00291-4

GREEN, P. J. and RICHARDSON, S. (2001). Modelling heterogeneity with and without the Dirichlet process. *Scand. J. Stat.* **28** 355–375. MR1842255 https://doi.org/10.1111/1467-9469.00242

GRIFFIN, J. E., KOLOSSIATIS, M. and STEEL, M. F. J. (2013). Comparing distributions by using dependent normalized random-measure mixtures. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **75** 499–529. MR3065477 https://doi.org/10.1111/rssb.12002

GRIFFIN, J. E. and LEISEN, F. (2017). Compound random measures and their use in Bayesian non-parametrics. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **79** 525–545. MR3611758 https://doi.org/10.1111/rssb.12176

GRIFFIN, J. E. and STEEL, M. F. J. (2006). Order-based dependent Dirichlet processes. *J. Amer. Statist. Assoc.* **101** 179–194. MR2268037 https://doi.org/10.1198/016214505000000727

GRIFFIN, J. E. and STEEL, M. F. J. (2010). Bayesian nonparametric modelling with the Dirichlet process regression smoother. *Statist. Sinica* **20** 1507–1527. MR2777334

GRIFFIN, J. E. and STEEL, M. F. J. (2011). Stick-breaking autoregressive processes. *J. Econometrics* **162** 383–396. MR2795625 https://doi.org/10.1016/j.jeconom.2011.03.001

GUTIÉRREZ, L., MENA, R. H. and RUGGIERO, M. (2016). A time dependent Bayesian nonparametric model for air quality analysis. *Comput. Statist. Data Anal.* **95** 161–175. MR3425946 https://doi.org/10.1016/j.csda.2015.10.002

GUTIÉRREZ, L., BARRIENTOS, A. F., GONZÁLEZ, J. and TAYLOR-RODRÍGUEZ, D. (2019). A Bayesian nonparametric multiple testing procedure for comparing several treatments against a control. *Bayesian Anal.* **14** 649–675. MR3959876 https://doi.org/10.1214/18-BA1122

GYÖRFI, L., KOHLER, M., KRZYŻAK, A. and WALK, H. (2002). *A Distribution-Free Theory of Nonparametric Regression*. *Springer Series in Statistics*. Springer, New York. MR1920390 https://doi.org/10.1007/b97848

HÄRDLE, W. (1991). *Smoothing Techniques*: *With Implementation in S. Springer Series in Statistics*. Springer, New York. MR1140190 https://doi.org/10.1007/978-1-4612-4432-5

JAMES, L. F., LIJOI, A. and PRÜNSTER, I. (2009). Posterior analysis for normalized random measures with independent increments. *Scand. J. Stat.* **36** 76–97. MR2508332 https://doi.org/10.1111/j.1467-9469.2008.00609.x

JARA, A. and HANSON, T. E. (2011). A class of mixtures of dependent tail-free processes. *Biometrika* **98** 553–566. MR2836406 https://doi.org/10.1093/biomet/asq082

JARA, A., LESAFFRE, E., DE IORIO, M. and QUINTANA, F. (2010). Bayesian semiparametric inference for multivariate doubly-interval-censored data. *Ann. Appl. Stat.* **4** 2126–2149. MR2829950 https://doi.org/10.1214/10-AOAS368

JARA, A., HANSON, T., QUINTANA, F., MÜLLER, P. and ROSNER, G. L. (2011). DPpackage: Bayesian semi- and nonparametric modeling in R. *J. Stat. Softw.* **40** 1–30.

KESSLER, D. C., HOFF, P. D. and DUNSON, D. B. (2015). Marginally specified priors for non-parametric Bayesian estimation. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **77** 35–58. MR3299398 https://doi.org/10.1111/rssb.12059

KLEMELÄ, J. (2014). *Multivariate Nonparametric Regression and Visualization*: *With R and Applications to Finance*. *Wiley Series in Computational Statistics*. Wiley, Hoboken, NJ. MR3222314

KOLOSSIATIS, M., GRIFFIN, J. E. and STEEL, M. F. J. (2013). On Bayesian nonparametric modelling of two correlated distributions. *Stat. Comput.* **23** 1–15. MR3018346 https://doi.org/10.1007/s11222-011-9283-7

LAU, J. W. and SO, M. K. P. (2008). Bayesian mixture of autoregressive models. *Comput. Statist. Data Anal.* **53** 38–60. MR2528591 https://doi.org/10.1016/j.csda.2008.06.001

LAVINE, M. (1992). Some aspects of Pólya tree distributions for statistical modelling. *Ann. Statist.* **20** 1222–1235. MR1186248 https://doi.org/10.1214/aos/1176348767

LEISEN, F. and LIJOI, A. (2011). Vectors of two-parameter Poisson–Dirichlet processes. *J. Multivariate Anal.* **102** 482–495. MR2755010 https://doi.org/10.1016/j.jmva.2010.10.008

LIJOI, A., NIPOTI, B. and PRÜNSTER, I. (2014). Bayesian inference with dependent normalized completely random measures. *Bernoulli* **20** 1260–1291. MR3217444 https://doi.org/10.3150/13-BEJ521

LIJOI, A. and PRÜNSTER, I. (2009). Distributional properties of means of random probability measures. *Stat. Surv.* **3** 47–95. MR2529667 https://doi.org/10.1214/09-SS041

LIN, D., GRIMSON, E. and FISHER, J. W. (2010). Construction of dependent Dirichlet processes based on Poisson processes. In *Advances in Neural Information Processing Systems* 23 (J. D. Lafferty, C. K. I. Williams, J. Shawe-Taylor, R. S. Zemel and A. Culotta, eds.) 1396–1404. Curran Associates, Red Hook.

LO, A. Y. (1984). On a class of Bayesian nonparametric estimates. I. Density estimates. *Ann. Statist.* **12** 351–357. MR0733519 https://doi.org/10.1214/aos/1176346412

MACEACHERN, S. N. (1999). Dependent nonparametric processes. In *ASA Proceedings of the Section on Bayesian Statistical Science* Amer. Statist. Assoc., Alexandria, VA.

MACEACHERN, S. N. (2000). Dependent Dirichlet processes. Technical report. Department of Statistics, The Ohio State Univ.

MENA, R. H. and RUGGIERO, M. (2016). Dynamic density estimation with diffusive Dirichlet mixtures. *Bernoulli* **22** 901–926. MR3449803 https://doi.org/10.3150/14-BEJ681

MENA, R. H., RUGGIERO, M. and WALKER, S. G. (2011). Geometric stick-breaking processes for continuous-time Bayesian nonparametric modeling. *J. Statist. Plann. Inference* **141** 3217–3230. MR2796026 https://doi.org/10.1016/j.jspi.2011.04.008

MIRA, A. and PETRONE, S. (1996). Bayesian hierarchical nonparametric inference for change-point problems. In *Bayesian Statistics*, 5 (*Alicante*, 1994). *Oxford Sci. Publ.* 693–703. Oxford Univ. Press, New York. MR1425440

MULIERE, P. and PETRONE, S. (1993). A Bayesian predictive approach to sequential search for an optimal dose: Parametric and nonparametric models. *J. Ital. Stat. Soc.* **2** 349–364.

MÜLLER, P., ERKANLI, A. and WEST, M. (1996). Bayesian curve fitting using multivariate normal mixtures. *Biometrika* **83** 67–79. MR1399156 https://doi.org/10.1093/biomet/83.1.67

MÜLLER, P. and QUINTANA, F. (2010). Random partition models with regression on covariates. *J. Statist. Plann. Inference* **140** 2801–2808. MR2651966 https://doi.org/10.1016/j.jspi.2010.03.002

MÜLLER, P., QUINTANA, F. and ROSNER, G. (2004). A method for combining inference across related nonparametric Bayesian models. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **66** 735–749. MR2088779 https://doi.org/10.1111/j.1467-9868.2004.05564.x

MÜLLER, P., QUINTANA, F. and ROSNER, G. L. (2011). A product partition model with regression on covariates. *J. Comput. Graph. Statist.* **20** 260–278. MR2816548 https://doi.org/10.1198/jcgs.2011.09066

MÜLLER, P., QUINTANA, F. A., JARA, A. and HANSON, T. (2015). *Bayesian Nonparametric Data Analysis*. Springer, New York.

PETRONE, S., GUINDANI, M. and GELFAND, A. E. (2009). Hybrid Dirichlet mixture models for functional data. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **71** 755–782. MR2750094 https://doi.org/10.1111/j.1467-9868.2009.00708.x

PRÜNSTER, I. and RUGGIERO, M. (2013). A Bayesian nonparametric approach to modeling market share dynamics. *Bernoulli* **19** 64–92. MR3019486 https://doi.org/10.3150/11-BEJ392

QUINTANA, F. A. and IGLESIAS, P. L. (2003). Bayesian clustering and product partition models. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **65** 557–574. MR1983764 https://doi.org/10.1111/1467-9868.00402

REGAZZINI, E., LIJOI, A. and PRÜNSTER, I. (2003). Distributional results for means of normalized random measures with independent increments. *Ann. Statist.* **31** 560–585. MR1983542 https://doi.org/10.1214/aos/1051027881

REICH, B. J. and FUENTES, M. (2007). A multivariate semiparametric Bayesian spatial modeling framework for hurricane surface wind fields. *Ann. Appl. Stat.* **1** 249–264. MR2393850 https://doi.org/10.1214/07-AOAS108

REN, L., DU, L., CARIN, L. and DUNSON, D. B. (2011). Logistic stick-breaking process. *J. Mach. Learn. Res.* **12** 203–239. MR2773552

RIGON, T. and DURANTE, D. (2021). Tractable Bayesian density regression via logit stick-breaking priors. *J. Statist. Plann. Inference* **211** 131–142. MR4117446 https://doi.org/10.1016/j.jspi.2020.05.009

RODRÍGUEZ, A. and DUNSON, D. B. (2011). Nonparametric Bayesian models through probit stick-breaking processes. *Bayesian Anal.* **6** 145–177. MR2781811 https://doi.org/10.1214/11-BA605

RODRÍGUEZ, A., DUNSON, D. B. and GELFAND, A. E. (2008). The nested Dirichlet process. *J. Amer. Statist. Assoc.* **103** 1131–1144. MR2528831 https://doi.org/10.1198/016214508000000553

RODRIGUEZ, A. and TER HORST, E. (2008). Bayesian dynamic density estimation. *Bayesian Anal.* **3** 339–365. MR2407430 https://doi.org/10.1214/08-BA313

SETHURAMAN, J. (1994). A constructive definition of Dirichlet priors. *Statist. Sinica* **4** 639–650. MR1309433

SKLAR, M. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Stat. Univ. Paris* **8** 229–231. MR0125600

TEH, Y. W., JORDAN, M. I., BEAL, M. J. and BLEI, D. M. (2006). Hierarchical Dirichlet processes. *J. Amer. Statist. Assoc.* **101** 1566–1581. MR2279480 https://doi.org/10.1198/016214506000000302

TOKDAR, S. T., ZHU, Y. M. and GHOSH, J. K. (2010). Bayesian density regression with logistic Gaussian process and subspace projection. *Bayesian Anal.* **5** 319–344. MR2719655 https://doi.org/10.1214/10-BA605

TRIPPA, L., MÜLLER, P. and JOHNSON, W. (2011). The multivariate beta process and an extension of the Polya tree model. *Biometrika* **98** 17–34. MR2804207 https://doi.org/10.1093/biomet/asq072

WADE, S., MONGELLUZZO, S. and PETRONE, S. (2011). An enriched conjugate prior for Bayesian nonparametric inference. *Bayesian Anal.* **6** 359–385. MR2843536 https://doi.org/10.1214/ba/1339616468

WANG, C. and ROSNER, G. L. (2019). A Bayesian nonparametric causal inference model for synthesizing randomized clinical trial and real-world evidence. *Stat. Med.* **38** 2573–2588. MR3962129 https://doi.org/10.1002/sim.8134

XU, Z., MACEACHERN, S. N. and XU, X. (2015). Modeling non-Gaussian time series with nonparametric Bayesian model. *IEEE Trans. Pattern Anal. Mach. Intell.* **37** 372–382.

Statistical Science
2022, Vol. 37, No. 1, 42–63
https://doi.org/10.1214/20-STS820

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

Marco Oesting and Kirstin Strokorb

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

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

REFERENCES

BÜCHER, A. and SEGERS, J. (2017). On the maximum likelihood estimator for the generalized extreme-value distribution. *Extremes* **20** 839–872. MR3737387 https://doi.org/10.1007/s10687-017-0292-6

CAPÉRAÀ, P., FOUGÈRES, A.-L. and GENEST, C. (1997). A nonparametric estimation procedure for bivariate extreme value copulas. *Biometrika* **84** 567–577. MR1603985 https://doi.org/10.1093/biomet/84.3.567

CHANDLER, R. E. and BATE, S. (2007). Inference for clustered data using the independence loglikelihood. *Biometrika* **94** 167–183. MR2367830 https://doi.org/10.1093/biomet/asm015

DAVISON, A. C. and HUSER, R. (2015). Statistics of extremes. *Annu. Rev. Stat. Appl.* **2** 203–235.

DAVISON, A. C., HUSER, R. and THIBAUD, E. (2018). Spatial extremes with application to climate and environmental exposure. In *Handbook of Environmental and Ecological Statistics* (A. E. Gelfand, M. Fuentes, J. A. Hoeting and R. L. Smith, eds.) CRC Press, Boca Raton, FL.

DAVISON, A. C., PADOAN, S. A. and RIBATET, M. (2012). Statistical modeling of spatial extremes. *Statist. Sci.* **27** 161–186. MR2963980 https://doi.org/10.1214/11-STS376

DE FONDEVILLE, R. and DAVISON, A. C. (2018). High-dimensional peaks-over-threshold inference. *Biometrika* **105** 575–592. MR3842886 https://doi.org/10.1093/biomet/asy026

DE HAAN, L. (1978). A characterization of multidimensional extreme-value distributions. *Sankhyā Ser. A* **40** 85–88. MR0545467

DE HAAN, L. (1984). A spectral representation for max-stable processes. *Ann. Probab.* **12** 1194–1204. MR0757776

DE HAAN, L. and FERREIRA, A. (2006). *Extreme Value Theory*: *An Introduction*. *Springer Series in Operations Research and Financial Engineering*. Springer, New York. MR2234156 https://doi.org/10.1007/0-387-34471-3

DE HAAN, L. and LIN, T. (2001). On convergence toward an extreme value distribution in $C[0,1]$. *Ann. Probab.* **29** 467–483. MR1825160 https://doi.org/10.1214/aop/1008956340

DIEKER, A. B. and MIKOSCH, T. (2015). Exact simulation of Brown–Resnick random fields at a finite number of locations. *Extremes* **18** 301–314. MR3351818 https://doi.org/10.1007/s10687-015-0214-4

DOMBRY, C., ENGELKE, S. and OESTING, M. (2016). Exact simulation of max-stable processes. *Biometrika* **103** 303–317. MR3509888 https://doi.org/10.1093/biomet/asw008

DOMBRY, C. and EYI-MINKO, F. (2012). Strong mixing properties of max-infinitely divisible random fields. *Stochastic Process. Appl.* **122** 3790–3811. MR2965926 https://doi.org/10.1016/j.spa.2012.06.013

DOMBRY, C. and EYI-MINKO, F. (2013). Regular conditional distributions of continuous max-infinitely divisible random fields. *Electron. J. Probab.* **18** 7. MR3024101 https://doi.org/10.1214/EJP.v18-1991

DOMBRY, C. and RIBATET, M. (2015). Functional regular variations, Pareto processes and peaks over threshold. *Stat. Interface* **8** 9–17. MR3320385 https://doi.org/10.4310/SII.2015.v8.n1.a2

EINMAHL, J. H. J., KIRILIOUK, A., KRAJINA, A. and SEGERS, J. (2016). An M-estimator of spatial tail dependence. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **78** 275–298. MR3453656 https://doi.org/10.1111/rssb.12114

ENGELKE, S., MALINOWSKI, A., KABLUCHKO, Z. and SCHLATHER, M. (2015). Estimation of Hüsler–Reiss distributions and Brown–Resnick processes. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **77** 239–265. MR3299407 https://doi.org/10.1111/rssb.12074

FISHER, R. A. and TIPPETT, L. H. C. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Math. Proc. Cambridge Philos. Soc.* **24** 180–190.

FRÉCHET, M. (1927). Sur la loi de probabilité de l'écart maximum. *Ann. Soc. Math. Polon.* **6** 93–116.

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

GILLELAND, E. and KATZ, R. W. (2016). extRemes 2.0: An extreme value analysis package in R. *J. Stat. Softw.* **72** 1–39.

GINÉ, E., HAHN, M. G. and VATAN, P. (1990). Max-infinitely divisible and max-stable sample continuous processes. *Probab. Theory Related Fields* **87** 139–165. MR1080487 https://doi.org/10.1007/BF01198427

GNEDENKO, B. (1943). Sur la distribution limite du terme maximum d'une série aléatoire. *Ann. of Math.* (2) **44** 423–453. MR0008655 https://doi.org/10.2307/1968974

HIJMANS, R. J. (2019). geosphere: Spherical Trigonometry. R package version 1.5-10.

HO, Z. W. O. and DOMBRY, C. (2019). Simple models for multivariate regular variation and the Hüsler–Reiss Pareto distribution. *J. Multivariate Anal.* **173** 525–550. MR3954072 https://doi.org/10.1016/j.jmva.2019.04.008

HUSER, R., DAVISON, A. C. and GENTON, M. G. (2016). Likelihood estimators for multivariate extremes. *Extremes* **19** 79–103. MR3454032 https://doi.org/10.1007/s10687-015-0230-4

JENKINSON, A. F. (1955). The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Q. J. R. Meteorol. Soc.* **81** 158–171.

KABLUCHKO, Z. (2011). Extremes of independent Gaussian processes. *Extremes* **14** 285–310. MR2824498 https://doi.org/10.1007/s10687-010-0110-x

KABLUCHKO, Z., SCHLATHER, M. and DE HAAN, L. (2009). Stationary max-stable fields associated to negative definite functions. *Ann. Probab.* **37** 2042–2065. MR2561440 https://doi.org/10.1214/09-AOP455

KIRILIOUK, A. (2016). tailDepFun: Minimum Distance Estimation of Tail Dependence Models. R package version 1.0.0.

LIU, Z., BLANCHET, J. H., DIEKER, A. B. and MIKOSCH, T. (2019). On logarithmically optimal exact simulation of max-stable and related random fields on a compact set. *Bernoulli* **25** 2949–2981. MR4003570 https://doi.org/10.3150/18-BEJ1076

MØLLER, J., ed. (2003) *Spatial Statistics and Computational Methods*. *Lecture Notes in Statistics* **173**. Springer, New York. MR2001383 https://doi.org/10.1007/978-0-387-21811-3

NORBERG, T. (1986). Random capacities and their distributions. *Probab. Theory Related Fields* **73** 281–297. MR0855227 https://doi.org/10.1007/BF00339941

OESTING, M., KABLUCHKO, Z. and SCHLATHER, M. (2012). Simulation of Brown–Resnick processes. *Extremes* **15** 89–107. MR2891311 https://doi.org/10.1007/s10687-011-0128-8

OESTING, M., RIBATET, M. and DOMBRY, C. (2016). Simulation of max-stable processes. In *Extreme Value Modeling and Risk Analysis* 195–214. CRC Press, Boca Raton, FL. MR3644314

OESTING, M., SCHLATHER, M. and SCHILLINGS, C. (2019). Sampling sup-normalized spectral functions for Brown–Resnick processes. *Stat* **8** e228. MR3978404 https://doi.org/10.1002/sta4.228

OESTING, M., SCHLATHER, M. and ZHOU, C. (2018). Exact and fast simulation of max-stable processes on a compact set using the normalized spectral representation. *Bernoulli* **24** 1497–1530. MR3706800 https://doi.org/10.3150/16-BEJ905

OESTING, M. and STROKORB, K. (2018). Efficient simulation of Brown–Resnick processes based on variance reduction of Gaussian processes. *Adv. in Appl. Probab.* **50** 1155–1175. MR3881113 https://doi.org/10.1017/apr.2018.54

OPITZ, T. (2013). Extremal t processes: Elliptical domain of attraction and a spectral representation. *J. Multivariate Anal.* **122** 409–413. MR3189331 https://doi.org/10.1016/j.jmva.2013.08.008

PENROSE, M. D. (1992). Semi-min-stable processes. *Ann. Probab.* **20** 1450–1463. MR1175271

R CORE TEAM (2020). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. R version 3.6.3.

RESNICK, S. I. and ROY, R. (1991). Random usc functions, max-stable processes and continuous choice. *Ann. Appl. Probab.* **1** 267–292. MR1102320

ROWLINGSON, B. (2019). geonames: Interface to the "Geonames" Spatial Query Web Service. R package version 0.999.

SCHLATHER, M. (2002). Models for stationary max-stable random fields. *Extremes* **5** 33–44. MR1947786 https://doi.org/10.1023/A:1020977924878

SCHLATHER, M., MALINOWSKI, A., OESTING, M., BOECKER, D., STROKORB, K., ENGELKE, S., MARTINI, J., BALLANI, F., MOREVA, O. et al. (2020). RandomFields: Simulation and Analysis of Random Fields. R package version 3.3.8.

SMITH, R. L. (1990). Max-stable processes and spatial extremes. Unpublished Manuscript.

SOUTH, A. (2017). rnaturalearth: World Map Data from Natural Earth. R package version 0.1.0.

STEPHENSON, A. G. (2002). evd: Extreme value distributions. *R News* **2** 0.

STUBNER, R. (2019). dqrng: Fast Pseudo Random Number Generators. R package version 0.2.1.

U.S./JAPAN ASTER SCIENCE TEAM ASTER Global Digital Elevation Model, V2.

VON MISES, R. (1936). La distribution de la plus grande de n valuers. *Rev. Math. Union Interbalcanique* **1** 141–160. Reproduced in: *Selected papers of Richard von Mises*, Amer. Math. Soc., Vol 2, 271-294 (1964).

Statistical Science
2022, Vol. 37, No. 1, 64–89
https://doi.org/10.1214/21-STS822

Analyzing Stochastic Computer Models: A Review with Opportunities

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

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

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

REFERENCES

ABT, M. and WELCH, W. J. (1998). Fisher information and maximum-likelihood estimation of covariance parameters in Gaussian stochastic processes. *Canad. J. Statist.* **26** 127–137. MR1624393 https://doi.org/10.2307/3315678

ALKEMA, L., RAFTERY, A. E. and CLARK, S. J. (2007). Probabilistic projections of HIV prevalence using Bayesian melding. *Ann. Appl. Stat.* **1** 229–248. MR2393849 https://doi.org/10.1214/07-AOAS111

ANDRIANAKIS, I., VERNON, I. R., MCCREESH, N., MCKINLEY, T. J., OAKLEY, J. E., NSUBUGA, R. N., GOLDSTEIN, M. and WHITE, R. G. (2015). Bayesian history matching of complex infectious disease models using emulation: A tutorial and a case study on HIV in Uganda. *PLoS Comput. Biol.* **11** e1003968.

ANDRIANAKIS, I., MCCREESH, N., VERNON, I., MCKINLEY, T. J., OAKLEY, J. E., NSUBUGA, R. N., GOLDSTEIN, M. and WHITE, R. G. (2017). Efficient history matching of a high dimensional individual-based HIV transmission model. *SIAM/ASA J. Uncertain. Quantificat.* **5** 694–719. MR3681376 https://doi.org/10.1137/16M1093008

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

Evan Baker is Postdoctoral Research Fellow, Living Systems Institute, University of Exeter, Stocker Road, Exeter, EX4 4QD, UK (e-mail: e.baker@exeter.ac.uk). Pierre Barbillon is Associate Professor, Université Paris-Saclay, AgroParisTech, INRAE, UMR MIA-Paris, 16 rue Claude Bernard, 75231 Paris Cedex 05, France. Arindam Fadikar is Postdoctoral Appointee, Mathematics and Computer Science Division, Argonne National Laboratory, 9700 South Cass Ave., Lemont, Illinois 60439, USA. Robert B. Gramacy is Professor, Department of Statistics, Virginia Tech, 250 Drillfield Drive Blacksburg, Virginia 24061, USA. Radu Herbei is Professor of Statistics, Department of Statistics, College of Arts and Sciences, The Ohio State University, 1958 Neil Ave., Columbus, Ohio 43210, USA. David Higdon is Professor, Department of Statistics, Virginia Tech, MC0439, Virginia Tech, Blacksburg, Virginia 24061, USA. Jiangeng Huang is Senior Statistical Scientist, Genentech, Inc., 1 DNA Way, South San Francisco, California 94080, USA. Leah R. Johnson is Associate Professor, Department of Statistics, Computational Modeling and Data Analytics (CMDA), Virginia Tech, Hutcheson Hall, RM 409-B, 250 Drillfield Drive, Blacksburg, Virginia 24061, USA. Pulong Ma is Postdoctoral Fellow, Duke University and Statistical and Applied Mathematical Sciences Institute, 19 T.W. Alexander Drive, P.O. Box 110207, Durham, North Carolina 27709, USA. Anirban Mondal is Assistant Professor, Department of Mathematics, Applied Mathematics, and Statistics, Case Western Reserve University, 10900 Euclid Avenue, Yost Hall Room 337, Cleveland, Ohio 44106-7058, USA. Bianica Pires is Lead Modeling & Simulation Engineer, The MITRE Corporation, 7515 Colshire Dr, McLean, Virginia 22102, USA. Jerome Sacks is Ph.D., NISS, 1460 N. Sandburg Ter, Apt 2902, Chicago, Illinois 60610, USA. Vadim Sokolov is Assistant Professor, Systems Engineering and Operations Research, George Mason University, Nguyen Engineering Building MS 4A6, Fairfax, Virginia 22302, USA.

tic kriging for simulation metamodeling. *Oper. Res.* **58** 371–382. MR2674803 https://doi.org/10.1287/opre.1090.0754

BA, S. (2019). SLHD: Maximin-distance (sliced) Latin hypercube designs. R package version 1.1.2.

BA, S. and JOSEPH, V. R. (2012). Composite Gaussian process models for emulating expensive functions. *Ann. Appl. Stat.* **6** 1838–1860. MR3058685 https://doi.org/10.1214/12-AOAS570

BAKER, E., CHALLENOR, P. and EAMES, M. (2020). Predicting the output from a stochastic computer model when a deterministic approximation is available. *J. Comput. Graph. Statist.* **29** 786–797. MR4191243 https://doi.org/10.1080/10618600.2020.1750416

BAKER, E., BARBILLON, P., FADIKAR, A., GRAMACY, R. B., HERBEI, R., HIGDON, D., HUANG, J., JOHNSON, L. R., MA, P., MONDAL, A., PIRES, B., SACKS, J. and SOKOLOV, V. (2022). Supplement to "Analyzing stochastic computer models: A review with opportunities." https://doi.org/10.1214/21-STS822SUPPA, https://doi.org/10.1214/21-STS822SUPPB

BARTON, R., NAKAYAMA, M. K. and SCHRUBEN, L. (2017). History of improving statistical efficiency. In 2017 *Winter Simulation Conference* (*WSC*) 158–180. IEEE, New York.

BASTOS, L. S. and O'HAGAN, A. (2009). Diagnostics for Gaussian process emulators. *Technometrics* **51** 425–438. MR2756478 https://doi.org/10.1198/TECH.2009.08019

BAYARRI, M. J., BERGER, J. O., CAFEO, J., GARCIA-DONATO, G., LIU, F., PALOMO, J., PARTHASARATHY, R. J., PAULO, R., SACKS, J. et al. (2007a). Computer model validation with functional output. *Ann. Statist.* **35** 1874–1906. MR2363956 https://doi.org/10.1214/009053607000000163

BAYARRI, M. J., BERGER, J. O., PAULO, R., SACKS, J., CAFEO, J. A., CAVENDISH, J., LIN, C.-H. and TU, J. (2007b). A framework for validation of computer models. *Technometrics* **49** 138–154. MR2380530 https://doi.org/10.1198/004017007000000092

BAYARRI, M. J., BERGER, J. O., KENNEDY, M. C., KOTTAS, A., PAULO, R., SACKS, J., CAFEO, J. A., LIN, C.-H. and TU, J. (2009). Predicting vehicle crashworthiness: Validation of computer models for functional and hierarchical data. *J. Amer. Statist. Assoc.* **104** 929–943. MR2750226 https://doi.org/10.1198/jasa.2009.ap06623

BEGON, M. et al. (1979). *Investigating Animal Abundance*: *Capture–Recapture for Biologists*. Edward Arnold Publishers Ltd., London.

BERNARDO, M. C., BUCK, R., LIU, L., NAZARET, W. A., SACKS, J. and WELCH, W. J. (1992). Integrated circuit design optimization using a sequential strategy. *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.* **11** 361–372.

BINOIS, M. and GRAMACY, R. B. (2018). hetGP: Heteroskedastic Gaussian process modeling and design under replication. R package version 1.1.1.

BINOIS, M., GRAMACY, R. B. and LUDKOVSKI, M. (2018). Practical heteroscedastic Gaussian process modeling for large simulation experiments. *J. Comput. Graph. Statist.* **27** 808–821. MR3890872 https://doi.org/10.1080/10618600.2018.1458625

BINOIS, M., HUANG, J., GRAMACY, R. B. and LUDKOVSKI, M. (2019). Replication or exploration? Sequential design for stochastic simulation experiments. *Technometrics* **61** 7–23. MR3933655 https://doi.org/10.1080/00401706.2018.1469433

BISSET, K. R., CHEN, J., FENG, X., KUMAR, V. and MARATHE, M. V. (2009). Epifast: A fast algorithm for large scale realistic epidemic simulations on distributed memory systems. In *Proceedings of the* 23*rd International Conference on Supercomputing* 430–439. ACM.

BOUKOUVALAS, A. and CORNFORD, D. (2009). Learning heteroscedastic Gaussian processes for complex datasets. Technical report, Aston University, Neural Computing Research Group.

BOUKOUVALAS, A., CORNFORD, D. and STEHLÍK, M. (2014). Optimal design for correlated processes with input-dependent noise. *Comput. Statist. Data Anal.* **71** 1088–1102. MR3132030 https://doi.org/10.1016/j.csda.2013.09.024

BOUKOUVALAS, A., SYKES, P., CORNFORD, D. and MARURI-AGUILAR, H. (2014). Bayesian precalibration of a large stochastic microsimulation model. *IEEE Trans. Intell. Transp. Syst.* **15** 1337–1347.

BREIMAN, L., FRIEDMAN, J. H., OLSHEN, R. A. and STONE, C. J. (1984). *Classification and Regression Trees*. *Wadsworth Statistics/Probability Series*. Wadsworth Advanced Books and Software, Belmont, CA. MR0726392

BRODERICK, T. and GRAMACY, R. B. (2011). Classification and categorical inputs with treed Gaussian process models. *J. Classification* **28** 244–270. MR2821001 https://doi.org/10.1007/s00357-011-9083-y

BRYNJARSDÓTTIR, J. and O'HAGAN, A. (2014). Learning about physical parameters: The importance of model discrepancy. *Inverse Probl.* **30** 114007, 24. MR3274591 https://doi.org/10.1088/0266-5611/30/11/114007

CHALABI, Y., DUTANG, C., SAVICKY, P. and WUERTZ, D. (2019). randtoolbox: Toolbox for pseudo and quasi random number generation and random generator tests. R package version 1.30.0.

CHEN, X., WANG, K. and YANG, F. (2013). Stochastic kriging with qualitative factors. In 2013 *Winter Simulations Conference* (*WSC*) 790–801. IEEE, New York.

CHEN, H., LOEPPKY, J. L., SACKS, J. and WELCH, W. J. (2016). Analysis methods for computer experiments: How to assess and what counts? *Statist. Sci.* **31** 40–60. MR3458592 https://doi.org/10.1214/15-STS531

CHIPMAN, H. A., GEORGE, E. I. and MCCULLOCH, R. E. (2010). BART: Bayesian additive regression trees. *Ann. Appl. Stat.* **4** 266–298. MR2758172 https://doi.org/10.1214/09-AOAS285

CHIPMAN, H. and MCCULLOCH, R. (2016). BayesTree: Bayesian additive regression trees. R package version 0.3-1.4.

CHUNG, M., BINOIS, M., GRAMACY, R. B., BARDSLEY, J. M., MOQUIN, D. J., SMITH, A. P. and SMITH, A. M. (2019). Parameter and uncertainty estimation for dynamical systems using surrogate stochastic processes. *SIAM J. Sci. Comput.* **41** A2212–A2238. MR3981381 https://doi.org/10.1137/18M1213403

CONTI, S. and O'HAGAN, A. (2010). Bayesian emulation of complex multi-output and dynamic computer models. *J. Statist. Plann. Inference* **140** 640–651. MR2558393 https://doi.org/10.1016/j.jspi.2009.08.006

CRAIG, P. S., GOLDSTEIN, M., SEHEULT, A. H. and SMITH, J. A. (1997). Pressure matching for hydrocarbon reservoirs: A case study in the use of Bayes linear strategies for large computer experiments. In *Case Studies in Bayesian Statistics* 37–93. Springer, Berlin.

CRESSIE, N. A. C. (1993). *Statistics for Spatial Data*. *Wiley Series in Probability and Mathematical Statistics*: *Applied Probability and Statistics*. Wiley, New York. Revised reprint of the 1991 edition, A Wiley-Interscience Publication. MR1239641 https://doi.org/10.1002/9781119115151

CURRIN, C., MITCHELL, T., MORRIS, M. and YLVISAKER, D. (1991). Bayesian prediction of deterministic functions, with applications to the design and analysis of computer experiments. *J. Amer. Statist. Assoc.* **86** 953–963. MR1146343

DUAN, W., ANKENMAN, B. E., SANCHEZ, S. M. and SANCHEZ, P. J. (2017). Sliced full factorial-based Latin hypercube designs as a framework for a batch sequential design algorithm. *Technometrics* **59** 11–22. MR3604185 https://doi.org/10.1080/00401706.2015.1108233

ERICKSON, C. B., ANKENMAN, B. E., PLUMLEE, M. and SANCHEZ, S. M. (2018). Gradient based criteria for sequential design. In 2018 *Winter Simulation Conference* (*WSC*) 467–478. IEEE, New York.

FADIKAR, A., HIGDON, D., CHEN, J., LEWIS, B., VENKATRAMANAN, S. and MARATHE, M. (2018). Calibrating a stochastic, agent-based model using quantile-based emulation. *SIAM/ASA J. Uncertain. Quantificat.* **6** 1685–1706. MR3890793 https://doi.org/10.1137/17M1161233

FARAH, M., BIRRELL, P., CONTI, S. and DE ANGELIS, D. (2014). Bayesian emulation and calibration of a dynamic epidemic model for A/H1N1 influenza. *J. Amer. Statist. Assoc.* **109** 1398–1411. MR3293599 https://doi.org/10.1080/01621459.2014.934453

FEYNMAN, R. P. (1948). Space-time approach to non-relativistic quantum mechanics. *Rev. Modern Phys.* **20** 367–387. MR0026940 https://doi.org/10.1103/revmodphys.20.367

FRAZIER, P. I. (2018). A tutorial on Bayesian optimization. arXiv preprint, arXiv:1807.02811.

FRENKLACH, M., PACKARD, A., GARCIA-DONATO, G., PAULO, R. and SACKS, J. (2016). Comparison of statistical and deterministic frameworks of uncertainty quantification. *SIAM/ASA J. Uncertain. Quantificat.* **4** 875–901. MR3531738 https://doi.org/10.1137/15M1019131

FRICKER, T. E., OAKLEY, J. E. and URBAN, N. M. (2013). Multivariate Gaussian process emulators with nonseparable covariance structures. *Technometrics* **55** 47–56. MR3038484 https://doi.org/10.1080/00401706.2012.715835

FUENTES, M. and RAFTERY, A. E. (2005). Model evaluation and spatial interpolation by Bayesian combination of observations with outputs from numerical models. *Biometrics* **61** 36–45. MR2129199 https://doi.org/10.1111/j.0006-341X.2005.030821.x

GAL, Y. and GHAHRAMANI, Z. (2016). Dropout as a Bayesian approximation: Representing model uncertainty in deep learning. In *Proceedings of the* 33*rd International Conference on International Conference on Machine Learning* 1050–1059.

GAO, F., SACKS, J. and WELCH, W. J. (1996). Predicting urban ozone levels and trends with semiparametric modeling. *J. Agric. Biol. Environ. Stat.* **1** 404–425. MR1807773 https://doi.org/10.2307/1400436

GNEITING, T. and RAFTERY, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *J. Amer. Statist. Assoc.* **102** 359–378. MR2345548 https://doi.org/10.1198/016214506000001437

GOLDBERG, P. W., WILLIAMS, C. K. and BISHOP, C. M. (1997). Regression with input-dependent noise: A Gaussian process treatment. In *Proceedings of the* 10*th International Conference on Neural Information Processing Systems* 493–499.

GRAMACY, R. (2007). tgp: An R package for Bayesian nonstationary, semiparametric nonlinear regression and design by treed Gaussian process models. *J. Stat. Softw.* **19** 6.

GRAMACY, R. B. (2020). *Surrogates*: *Gaussian Process Modeling*, *Design and Optimization for the Applied Sciences*. Chapman Hall/CRC, Boca Raton, FL.

GRAMACY, R. B. and LEE, H. K. H. (2008). Bayesian treed Gaussian process models with an application to computer modeling. *J. Amer. Statist. Assoc.* **103** 1119–1130. MR2528830 https://doi.org/10.1198/016214508000000689

GRAMACY, R. B. and LEE, H. K. H. (2009). Adaptive design and analysis of supercomputer experiments. *Technometrics* **51** 130–145. MR2668170 https://doi.org/10.1198/TECH.2009.0015

GRAMACY, R. B. and LEE, H. K. H. (2011). Optimization under unknown constraints. In *Bayesian Statistics* 9 (J. Bernardo, S. Bayarri, J. O. Berger, A. P. Dawid, D. Heckerman, A. F. M. Smith and M. West, eds.) 229–256. Oxford Univ. Press, Oxford. MR3204008 https://doi.org/10.1093/acprof:oso/9780199694587.003.0008

GRAMACY, R. and TADDY, M. (2010). Categorical inputs, sensitivity analysis, optimization and importance tempering with tgp version 2, an R package for treed Gaussian process models. *J. Stat. Softw.* **33** 1–48.

GRAMACY, R. B. and TADDY, M. A. (2016). tgp: Bayesian treed Gaussian process models. R package version 2.4-14.

GRAVES, A. (2011). Practical variational inference for neural networks. In *Proceedings of the* 24*th International Conference on Neural Information Processing Systems* 2348–2356.

GRIMM, V., BERGER, U., BASTIANSEN, F., ELIASSEN, S., GINOT, V., GISKE, J., GOSS-CUSTARD, J., GRAND, T., HEINZ, S. K. et al. (2006). A standard protocol for describing individual-based and agent-based models. *Ecol. Model.* **198** 115–126.

GU, M. and WANG, L. (2018). Scaled Gaussian stochastic process for computer model calibration and prediction. *SIAM/ASA J. Uncertain. Quantificat.* **6** 1555–1583. MR3875809 https://doi.org/10.1137/17M1159890

HART, J. L., ALEXANDERIAN, A. and GREMAUD, P. A. (2017). Efficient computation of Sobol' indices for stochastic models. *SIAM J. Sci. Comput.* **39** A1514–A1530. MR3686808 https://doi.org/10.1137/16M106193X

HARVILLE, D. A. (1997). *Matrix Algebra from a Statistician's Perspective*. Springer, New York. MR1467237 https://doi.org/10.1007/b98818

HENDERSON, D. A., BOYS, R. J., KRISHNAN, K. J., LAWLESS, C. and WILKINSON, D. J. (2009). Bayesian emulation and calibration of a stochastic computer model of mitochondrial DNA deletions in substantia nigra neurons. *J. Amer. Statist. Assoc.* **104** 76–87. MR2663034 https://doi.org/10.1198/jasa.2009.0005

HERBEI, R. and BERLINER, L. M. (2014). Estimating ocean circulation: An MCMC approach with approximated likelihoods via the Bernoulli factory. *J. Amer. Statist. Assoc.* **109** 944–954. MR3265667 https://doi.org/10.1080/01621459.2014.914439

HIGDON, D., KENNEDY, M., CAVENDISH, J. C., CAFEO, J. A. and RYNE, R. D. (2004). Combining field data and computer simulations for calibration and prediction. *SIAM J. Sci. Comput.* **26** 448–466. MR2116355 https://doi.org/10.1137/S1064827503426693

HIGDON, D., GATTIKER, J., WILLIAMS, B. and RIGHTLEY, M. (2008). Computer model calibration using high-dimensional output. *J. Amer. Statist. Assoc.* **103** 570–583. MR2523994 https://doi.org/10.1198/016214507000000888

JALALI, H., VAN NIEUWENHUYSE, I. and PICHENY, V. (2017). Comparison of Kriging-based algorithms for simulation optimization with heterogeneous noise. *European J. Oper. Res.* **261** 279–301. MR3629301 https://doi.org/10.1016/j.ejor.2017.01.035

JOHNSON, L. R. (2010). Implications of dispersal and life history strategies for the persistence of linyphiid spider populations. *Ecol. Model.* **221** 1138–1147.

JOHNSON, L. R. and BRIGGS, C. J. (2011). Parameter inference for an individual based model of chytridiomycosis in frogs. *J. Theoret. Biol.* **277** 90–98.

JOHNSON, L. R., GRAMACY, R. B., COHEN, J., MORDECAI, E., MURDOCK, C., ROHR, J., RYAN, S. J., STEWART-IBARRA, A. M. and WEIKEL, D. (2018). Phenomenological forecasting of disease incidence using heteroskedastic Gaussian processes: A Dengue case study. *Ann. Appl. Stat.* **12** 27–66. MR3773385 https://doi.org/10.1214/17-AOAS1090

JOLLIFFE, I. (2011). *Principal Component Analysis*. Springer, Berlin. MR0841268 https://doi.org/10.1007/978-1-4757-1904-8

JONES, D. R., SCHONLAU, M. and WELCH, W. J. (1998). Efficient global optimization of expensive black-box functions. *J. Global Optim.* **13** 455–492. Workshop on Global Optimization (Trier, 1997). MR1673460 https://doi.org/10.1023/A:1008306431147

KAC, M. (1949). On distributions of certain Wiener functionals. *Trans. Amer. Math. Soc.* **65** 1–13. MR0027960 https://doi.org/10.2307/1990512

KENNEDY, J. C., HENDERSON, D. A. and WILSON, K. J. (2020). Multilevel emulation for stochastic computer models with an application to large offshore windfarms. arXiv preprint, arXiv:2003.08921.

KENNEDY, M. C. and O'HAGAN, A. (2000). Predicting the output from a complex computer code when fast approximations are available. *Biometrika* **87** 1–13. MR1766824 https://doi.org/10.1093/biomet/87.1.1

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

KERSAUDY, P., SUDRET, B., VARSIER, N., PICON, O. and WIART, J. (2015). A new surrogate modeling technique combining kriging and polynomial chaos expansions—application to uncertainty analysis in computational dosimetry. *J. Comput. Phys.* **286** 103–117. MR3314189 https://doi.org/10.1016/j.jcp.2015.01.034

KERSTING, K., PLAGEMANN, C., PFAFF, P. and BURGARD, W. (2007). Most likely heteroscedastic Gaussian process regression. In *Proceedings of the 24th International Conference on Machine Learning* 393–400.

KIM, H.-M., MALLICK, B. K. and HOLMES, C. C. (2005). Analyzing nonstationary spatial data using piecewise Gaussian processes. *J. Amer. Statist. Assoc.* **100** 653–668. MR2160567 https://doi.org/10.1198/016214504000002014

KLEIJNEN, J. P. C. (2009). Kriging metamodeling in simulation: A review. *European J. Oper. Res.* **192** 707–716. MR2457613 https://doi.org/10.1016/j.ejor.2007.10.013

KLEIJNEN, J. P. C. (2017). Regression and kriging metamodels with their experimental designs in simulation: A review. *European J. Oper. Res.* **256** 1–16. MR3543078 https://doi.org/10.1016/j.ejor.2016.06.041

KOENKER, R. and BASSETT, G. JR. (1978). Regression quantiles. *Econometrica* **46** 33–50. MR0474644 https://doi.org/10.2307/1913643

KONOMI, B. A., KARAGIANNIS, G., LAI, K. and LIN, G. (2017). Bayesian treed calibration: An application to carbon capture with AX sorbent. *J. Amer. Statist. Assoc.* **112** 37–53. MR3646551 https://doi.org/10.1080/01621459.2016.1190279

KUSHNER, H. J. (1964). A new method of locating the maximum point of an arbitrary multipeak curve in the presence of noise. *J. Basic Eng.* **86** 97–106.

LAKSHMINARAYANAN, B., PRITZEL, A. and BLUNDELL, C. (2017). Simple and scalable predictive uncertainty estimation using deep ensembles. In *Advances in Neural Information Processing Systems* 6402–6413.

LEE, A. (2015). pyDOE: The experimental design package for python. Python package version 0.3.8.

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

LOEPPKY, J. L., MOORE, L. M. and WILLIAMS, B. J. (2010). Batch sequential designs for computer experiments. *J. Statist. Plann. Inference* **140** 1452–1464. MR2592224 https://doi.org/10.1016/j.jspi.2009.12.004

LOEPPKY, J. L., SACKS, J. and WELCH, W. J. (2009). Choosing the sample size of a computer experiment: A practical guide. *Technometrics* **51** 366–376. MR2756473 https://doi.org/10.1198/TECH.2009.08040

LYU, X., BINOIS, M. and LUDKOVSKI, M. (2018). Evaluating Gaussian process metamodels and sequential designs for noisy level set estimation. arXiv preprint, arXiv:1807.06712.

MA, P., MONDAL, A., KONOMI, B., HOBBS, J., SONG, J. and KANG, E. (2021). Computer model emulation with high-dimensional functional output in large-scale observing system uncertainty experiments. *Technometrics*. https://doi.org/10.1080/00401706.2021.1895890

MARREL, A., IOOSS, B., LAURENT, B. and ROUSTANT, O. (2009). Calculations of Sobol indices for the Gaussian process metamodel. *Reliab. Eng. Syst. Saf.* **94** 742–751.

MARREL, A., IOOSS, B., DA VEIGA, S. and RIBATET, M. (2012). Global sensitivity analysis of stochastic computer models with joint metamodels. *Stat. Comput.* **22** 833–847. MR2909625 https://doi.org/10.1007/s11222-011-9274-8

MCCULLOCH, R., SPARAPANI, R., GRAMACY, R., SPANBAUER, C. and PRATOLA, M. (2019). BART: Bayesian additive regression trees. R package version 2.7.

MCKAY, M. D., BECKMAN, R. J. and CONOVER, W. J. (1979). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics* **21** 239–245. MR0533252 https://doi.org/10.2307/1268522

MCKEAGUE, I. W., NICHOLLS, G., SPEER, K. and HERBEI, R. (2005). Statistical inversion of South Atlantic circulation in an abyssal neutral density layer. *J. Mar. Res.* **63** 683–704.

MCKINLEY, T. J., VERNON, I., ANDRIANAKIS, I., MCCREESH, N., OAKLEY, J. E., NSUBUGA, R. N., GOLDSTEIN, M. and WHITE, R. G. (2018). Approximate Bayesian computation and simulation-based inference for complex stochastic epidemic models. *Statist. Sci.* **33** 4–18. MR3757500 https://doi.org/10.1214/17-STS618

MOCKUS, J., TIESIS, V. and ZILINSKAS, A. (1978). The application of Bayesian methods for seeking the extremum. In *Towards Global Optimization* **2** 117–129.

MOLINA, G., BAYARRI, M. J. and BERGER, J. O. (2005). Statistical inverse analysis for a network microsimulator. *Technometrics* **47** 388–398. MR2208308 https://doi.org/10.1198/004017005000000409

MORTVEIT, H., ADIGA, A., AGASHE, A., ALAM, M., ALEXANDER, K., ARIFUZZAMAN, S., BARRETT, C., BECKMAN, R., BISSET, K. et al. (2015). Synthetic populations and interaction networks for Guinea, Liberia and Sierra Leone. Technical report, NDSSL.

MOUTOUSSAMY, V., NANTY, S. and PAUWELS, B. (2015). Emulators for stochastic simulation codes. In *CEMRACS* 2013—*Modelling and Simulation of Complex Systems*: *Stochastic and Deterministic Approaches*. *ESAIM Proc. Surveys* **48** 116–155. EDP Sci., Les Ulis. MR3415390 https://doi.org/10.1051/proc/201448005

NEAL, R. M. (1996). *Bayesian Learning for Neural Networks*. *Lecture Notes in Statistics* **118**. Springer, New York.

NGUYEN, H., CRESSIE, N. and BRAVERMAN, A. (2017). Multivariate spatial data fusion for very large remote sensing datasets. *Remote Sens.* **9** 142.

O'HAGAN, A. (2006). Bayesian analysis of computer code outputs: A tutorial. *Reliab. Eng. Syst. Saf.* **91** 1290–1300.

OAKLEY, J. E. and O'HAGAN, A. (2004). Probabilistic sensitivity analysis of complex models: A Bayesian approach. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **66** 751–769. MR2088780 https://doi.org/10.1111/j.1467-9868.2004.05304.x

OAKLEY, J. E. and YOUNGMAN, B. D. (2017). Calibration of stochastic computer simulators using likelihood emulation. *Technometrics* **59** 80–92. MR3604191 https://doi.org/10.1080/00401706.2015.1125391

OPITZ, T., HUSER, R., BAKKA, H. and RUE, H. (2018). INLA goes extreme: Bayesian tail regression for the estimation of high spatio-temporal quantiles. *Extremes* **21** 441–462. MR3855716 https://doi.org/10.1007/s10687-018-0324-x

PAPAMAKARIOS, G., NALISNICK, E., REZENDE, D. J., MOHAMED, S. and LAKSHMINARAYANAN, B. (2019). Normalizing flows for probabilistic modeling and inference. arXiv preprint, arXiv:1912.02762.

PARK, C. and APLEY, D. (2018). Patchwork Kriging for large-scale Gaussian process regression. *J. Mach. Learn. Res.* **19** Paper No. 7, 43. MR3862414

PAULO, R., GARCÍA-DONATO, G. and PALOMO, J. (2012). Calibration of computer models with multivariate output. *Comput. Statist. Data Anal.* **56** 3959–3974. MR2957846 https://doi.org/10.1016/j.csda.2012.05.023

PEDREGOSA, F., VAROQUAUX, G., GRAMFORT, A. et al. (2011). Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **12** 2825–2830. MR2854348

PELEG, N., FATICHI, S., PASCHALIS, A., MOLNAR, P. and BURLANDO, P. (2017). An advanced stochastic weather generator for simulating 2-d high-resolution climate variables. *J. Adv. Model. Earth Syst.* **9** 1595–1627.

PICHENY, V. and GINSBOURGER, D. (2014). Noisy kriging-based optimization methods: A unified implementation within the DiceOptim package. *Comput. Statist. Data Anal.* **71** 1035–1053. MR3132025 https://doi.org/10.1016/j.csda.2013.03.018

PICHENY, V., GINSBOURGER, D. and ROUSTANT, O. (2016). DiceOptim: Kriging-based optimization for computer experiments. R package version 2.0.

PICHENY, V., WAGNER, T. and GINSBOURGER, D. (2013). A benchmark of kriging-based infill criteria for noisy optimization. *Struct. Multidiscip. Optim.* **48** 607–626.

PLUMLEE, M. (2017). Bayesian calibration of inexact computer models. *J. Amer. Statist. Assoc.* **112** 1274–1285. MR3735376 https://doi.org/10.1080/01621459.2016.1211016

PLUMLEE, M. and TUO, R. (2014). Building accurate emulators for stochastic simulations via quantile kriging. *Technometrics* **56** 466–473. MR3290616 https://doi.org/10.1080/00401706.2013.860919

POOLE, D. and RAFTERY, A. E. (2000). Inference for deterministic simulation models: The Bayesian melding approach. *J. Amer. Statist. Assoc.* **95** 1244–1255. MR1804247 https://doi.org/10.2307/2669764

PRATOLA, M. T., CHIPMAN, H. A., GEORGE, E. I. and MCCULLOCH, R. E. (2020). Heteroscedastic BART via multiplicative regression trees. *J. Comput. Graph. Statist.* **29** 405–417. MR4116052 https://doi.org/10.1080/10618600.2019.1677243

PRONZATO, L. and MÜLLER, W. G. (2012). Design of computer experiments: Space filling and beyond. *Stat. Comput.* **22** 681–701. MR2909615 https://doi.org/10.1007/s11222-011-9242-3

PUKELSHEIM, F. (1994). The three sigma rule. *Amer. Statist.* **48** 88–91. MR1292524 https://doi.org/10.2307/2684253

QIAN, P. Z. G., WU, H. and WU, C. F. J. (2008). Gaussian process models for computer experiments with qualitative and quantitative factors. *Technometrics* **50** 383–396. MR2457574 https://doi.org/10.1198/004017008000000262

RADTKE, P. J., BURK, T. E. and BOLSTAD, P. V. (2002). Bayesian melding of a forest ecosystem model with correlated inputs. *For. Sci.* **48** 701–711.

RAFTERY, A. E., GIVENS, G. H. and ZEH, J. E. (1995). Inference from a deterministic population dynamics model for bowhead whales. *J. Amer. Statist. Assoc.* **90** 402–416.

RAMSEY, D. S. and EFFORD, M. G. (2010). Management of bovine tuberculosis in brushtail possums in New Zealand: Predictions from a spatially explicit, individual-based model. *J. Appl. Ecol.* **47** 911–919.

RANNOU, V., BROUAYE, F., HÉLIER, M. and TABBARA, W. (2002). Kriging the quantile: Application to a simple transmission line model. *Inverse Probl.* **18** 37.

RASMUSSEN, C. E. and WILLIAMS, C. K. I. (2006). *Gaussian Processes for Machine Learning. Adaptive Computation and Machine Learning*. MIT Press, Cambridge, MA. MR2514435

REYNOLDS, C. W. (1987). Flocks, herds and schools: A distributed behavioral model. In *Proceedings of the* 14*th Annual Conference on Computer Graphics and Interactive Techniques* 25–34.

RICHARDSON, C. W. (1981). Stochastic simulation of daily precipitation, temperature, and solar radiation. *Water Resour. Res.* **17** 182–190.

ROUSTANT, O., GINSBOURGER, D. and DEVILLE, Y. (2018). DiceKriging: Kriging methods for computer experiments. R package version 1.5.6.

RULLIÈRE, D., DURRANDE, N., BACHOC, F. and CHEVALIER, C. (2018). Nested Kriging predictions for datasets with a large number of observations. *Stat. Comput.* **28** 849–867. MR3766047 https://doi.org/10.1007/s11222-017-9766-2

RUSHDI, A. A., SWILER, L. P., PHIPPS, E. T., D'ELIA, M. and EBEIDA, M. S. (2017). VPS: Voronoi piecewise surrogate models for high-dimensional data fitting. *Int. J. Uncertain. Quantif.* **7** 1–21. MR3638631 https://doi.org/10.1615/Int.J.UncertaintyQuantification.2016018697

RUTTER, C. M., OZIK, J., DEYOREO, M. and COLLIER, N. (2019). Microsimulation model calibration using incremental mixture approximate Bayesian computation. *Ann. Appl. Stat.* **13** 2189–2212. MR4037427 https://doi.org/10.1214/19-aoas1279

SACKS, J., WELCH, W. J., MITCHELL, T. J. and WYNN, H. P. (1989). Design and analysis of computer experiments. *Statist. Sci.* **4** 409–435. With comments and a rejoinder by the authors. MR1041765

SALTER, J. M., WILLIAMSON, D. B., SCINOCCA, J. and KHARIN, V. (2019). Uncertainty quantification for computer models with spatial output using calibration-optimal bases. *J. Amer. Statist. Assoc.* **114** 1800–1814. MR4047301 https://doi.org/10.1080/01621459.2018.1514306

SANTNER, T. J., WILLIAMS, B. J. and NOTZ, W. I. (2018). *The Design and Analysis of Computer Experiments. Springer Series in Statistics*. Springer, New York. Second edition of [MR2160708]. MR3887662

SCHONLAU, M. and WELCH, W. J. (2006). Screening the input variables to a computer model via analysis of variance and visualization. In *Screening* 308–327. Springer.

SCHULTZ, L. and SOKOLOV, V. (2018). Practical Bayesian optimization for transportation simulators. arXiv preprint, arXiv:1810.03688.

ŠEVČÍKOVÁ, H., RAFTERY, A. E. and WADDELL, P. A. (2007). Assessing uncertainty in urban simulations using Bayesian melding. *Transp. Res., Part B, Methodol.* **41** 652–669.

SHAH, A., WILSON, A. and GHAHRAMANI, Z. (2014). Student-t processes as alternatives to Gaussian processes. In *Proceedings of the* 17*th International Conference on Artificial Intelligence and Statistics. PMLR* **33** 877–885.

SMIESZEK, T., BALMER, M., HATTENDORF, J., AXHAUSEN, K. W., ZINSSTAG, J. and SCHOLZ, R. W. (2011). Reconstructing the 2003/2004 H3N2 influenza epidemic in Switzerland with a spatially explicit, individual-based model. *BMC Infect. Dis.* **11** 115. https://doi.org/10.1186/1471-2334-11-115

SOBOL, I. M. (1967). On the distribution of points in a cube and the approximate evaluation of integrals. *Comput. Math. Math. Phys.* **7** 86–112. MR0219238

SOBOL, I. M. (1993). Sensitivity estimates for nonlinear mathematical models. *Math. Model. Comput. Exp.* **1** 407–414. MR1335161

SPILLER, E. T., BAYARRI, M. J., BERGER, J. O., CALDER, E. S., PATRA, A. K., PITMAN, E. B. and WOLPERT, R. L. (2014). Automating emulator construction for geophysical hazard maps.

SIAM/ASA J. Uncertain. Quantificat. **2** 126–152. MR3283903 https://doi.org/10.1137/120899285

SRINIVAS, N., KRAUSE, A., KAKADE, S. M. and SEEGER, M. W. (2012). Information-theoretic regret bounds for Gaussian process optimization in the bandit setting. *IEEE Trans. Inf. Theory* **58** 3250–3265. MR2952544 https://doi.org/10.1109/TIT.2011.2182033

STEIN, M. L. (2012). *Interpolation of Spatial Data*: *Some Theory for Kriging*. Springer, New York. MR1697409 https://doi.org/10.1007/978-1-4612-1494-6

SULLIVAN, T. J. (2015). *Introduction to Uncertainty Quantification*. *Texts in Applied Mathematics* **63**. Springer, Cham. MR3364576 https://doi.org/10.1007/978-3-319-23395-6

SUN, F., GRAMACY, R. B., HAALAND, B., LU, S. and HWANG, Y. (2019). Synthesizing simulation and field data of solar irradiance. *Stat. Anal. Data Min.* **12** 311–324. MR3994771 https://doi.org/10.1002/sam.11414

SUNG, C.-L., BARBER, B. D. and WALKER, B. J. (2019). Calibration of computer models with heteroscedastic errors and application to plant relative growth rates. arXiv preprint, arXiv:1910.11518.

TUO, R. and JEFF WU, C. F. (2016). A theoretical framework for calibration in computer models: Parametrization, estimation and convergence properties. *SIAM/ASA J. Uncertain. Quantificat.* **4** 767–795. MR3523087 https://doi.org/10.1137/151005841

TUO, R. and WU, C. F. J. (2015). Efficient calibration for imperfect computer models. *Ann. Statist.* **43** 2331–2352. MR3405596 https://doi.org/10.1214/15-AOS1314

VAZQUEZ, E., VILLEMONTEIX, J., SIDORKIEWICZ, M. and WALTER, E. (2008). Global optimization based on noisy evaluations: An empirical study of two statistical approaches. *J. Phys., Conf. Ser.* **135**. 012100.

VERNON, I., GOLDSTEIN, M. and BOWER, R. G. (2010). Galaxy formation: A Bayesian uncertainty analysis. *Bayesian Anal.* **5** 619–669. MR2740148 https://doi.org/10.1214/10-BA524

WANG, W. and HAALAND, B. (2019). Controlling sources of inaccuracy in stochastic kriging. *Technometrics* **61** 309–321. MR3987622 https://doi.org/10.1080/00401706.2018.1514328

WANG, S. and NG, S. H. (2020). Enhancing response predictions with a joint Gaussian process model for stochastic simulation models. *ACM Trans. Model. Comput. Simul.* **30** Art. 4, 25. MR4066845 https://doi.org/10.1145/3364219

WANG, Z., SHI, J. Q. and LEE, Y. (2017). Extended t-process regression models. *J. Statist. Plann. Inference* **189** 38–60. MR3659848 https://doi.org/10.1016/j.jspi.2017.05.006

WELLING, M. and TEH, Y. W. (2011). Bayesian learning via stochastic gradient Langevin dynamics. In *Proceedings of the* 28*th International Conference on Machine Learning* (*ICML*-11) 681–688.

WILENSKY, U. (1999). NetLogo. Available at http://ccl.northwestern.edu/netlogo/. Center for Connected Learning and Computer-Based Modeling, Northwestern Univ., Evanston, IL.

WILKINSON, R. D. (2013). Approximate Bayesian computation (ABC) gives exact results under the assumption of model error. *Stat. Appl. Genet. Mol. Biol.* **12** 129–141. MR3071024 https://doi.org/10.1515/sagmb-2013-0010

XIE, G. and CHEN, X. (2017). A heteroscedastic t-process simulation metamodeling approach and its application in inventory control and optimization. In *Simulation Conference* (*WSC*), 2017 *Winter* 3242–3253. IEEE, New York.

ZHANG, B., COLE, D. A. and GRAMACY, R. B. (2021). Distance-Distributed Design for Gaussian Process Surrogates. *Technometrics* **63** 40–52. MR4205690 https://doi.org/10.1080/00401706.2019.1677269

ZHANG, J., CRESSIE, N. and CRAIGMILE, P. F. (2008). Loss function approaches to predict a spatial quantile and its exceedance region. *Technometrics* **50** 216–227. MR2439879 https://doi.org/10.1198/004017008000000226

ZHANG, Q. and XIE, W. (2017). Asymmetric kriging emulator for stochastic simulation. In *Proceedings of the* 2017 *Winter Simulation Conference* 137. IEEE Press, New York.

ZHANG, Y., TAO, S., CHEN, W. and APLEY, D. W. (2020). A latent variable approach to Gaussian process modeling with qualitative and quantitative factors. *Technometrics* **62** 291–302. MR4125497 https://doi.org/10.1080/00401706.2019.1638834

ZHOU, J., CHANG, H. H. and FUENTES, M. (2012). Estimating the health impact of climate change with calibrated climate model output. *J. Agric. Biol. Environ. Stat.* **17** 377–394. MR2993272 https://doi.org/10.1007/s13253-012-0105-y

ZHOU, Q., QIAN, P. Z. G. and ZHOU, S. (2011). A simple approach to emulation for computer models with qualitative and quantitative factors. *Technometrics* **53** 266–273. MR2857704 https://doi.org/10.1198/TECH.2011.10025

Statistical Science
2022, Vol. 37, No. 1, 90–109
https://doi.org/10.1214/21-STS823

Statistical Dependence: Beyond Pearson's ρ

Dag Tjøstheim, Håkon Otneim and Bård Støve

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

Key words and phrases: Statistical dependence, Pearson's ρ, nonlinear dependence, distance covariance, HSIC, mutual information, local Gaussian correlation.

REFERENCES

AAS, K., CZADO, C., FRIGESSI, A. and BAKKEN, H. (2009). Pair-copula constructions of multiple dependence. *Insurance Math. Econom.* **44** 182–198. MR2517884 https://doi.org/10.1016/j.insmatheco.2007.02.001

ARONSZAJN, N. (1950). Theory of reproducing kernels. *Trans. Amer. Math. Soc.* **68** 337–404. MR0051437 https://doi.org/10.2307/1990404

BERENTSEN, G. D., KLEPPE, T. and TJØSTHEIM, D. (2014). Introducing localgauss, an R-package for estimating and visualizing local Gaussian corelation. *J. Stat. Softw.* **56** 1–18.

BERENTSEN, G. D. and TJØSTHEIM, D. (2014). Recognizing and visualizing departures from independence in bivariate data using local Gaussian correlation. *Stat. Comput.* **24** 785–801. MR3229697 https://doi.org/10.1007/s11222-013-9402-8

BERENTSEN, G. D., STØVE, B., TJØSTHEIM, D. and NORDBØ, T. (2014). Recognizing and visualizing copulas: An approach using local Gaussian approximation. *Insurance Math. Econom.* **57** 90–103. MR3225330 https://doi.org/10.1016/j.insmatheco.2014.04.005

BERGSMA, W. and DASSIOS, A. (2014). A consistent test of independence based on a sign covariance related to Kendall's tau. *Bernoulli* **20** 1006–1028. MR3178526 https://doi.org/10.3150/13-BEJ514

BERLINET, A. and THOMAS-AGNAN, C. (2004). *Reproducing Kernel Hilbert Spaces in Probability and Statistics*. Kluwer Academic, Boston, MA. With a preface by Persi Diaconis. MR2239907 https://doi.org/10.1007/978-1-4419-9096-9

BERRETT, T. B. and SAMWORTH, R. J. (2019). Nonparametric independence testing via mutual information. *Biometrika* **106** 547–566. MR3992389 https://doi.org/10.1093/biomet/asz024

BICKEL, P. J. and ROSENBLATT, M. (1973). On some global measures of the deviations of density function estimates. *Ann. Statist.* **1** 1071–1095. MR0348906

BILODEAU, M. and LAFAYE DE MICHEAUX, P. (2009). *A*-dependence statistics for mutual and serial independence of categorical variables. *J. Statist. Plann. Inference* **139** 2407–2419. MR2508002 https://doi.org/10.1016/j.jspi.2008.11.006

BILODEAU, M. and NANGUE, A. G. (2017). Tests of mutual or serial independence of random vectors with applications. *J. Mach. Learn. Res.* **18** Paper No. 74. MR3714237

BJERVE, S. and DOKSUM, K. (1993). Correlation curves: Measures of association as functions of covariate values. *Ann. Statist.* **21** 890–902. MR1232524 https://doi.org/10.1214/aos/1176349156

BLOMQVIST, N. (1950). On a measure of dependence between two random variables. *Ann. Math. Stat.* **21** 593–600. MR0039190 https://doi.org/10.1214/aoms/1177729754

BLUM, J. R., KIEFER, J. and ROSENBLATT, M. (1961). Distribution free tests of independence based on the sample distribution function. *Ann. Math. Stat.* **32** 485–498. MR0125690 https://doi.org/10.1214/aoms/1177705055

BÖTTCHER, B., KELLER-RESSEL, M. and SCHILLING, R. L. (2019). Distance multivariance: New dependence measures for random vectors. *Ann. Statist.* **47** 2757–2789. MR3988772 https://doi.org/10.1214/18-AOS1764

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

BROCK, W. A., SCHEINKMAN, J. A., DECHERT, W. D. and LEBARON, B. (1996). A test for independence based on the correlation dimension. *Econometric Rev.* **15** 197–235. MR1410877 https://doi.org/10.1080/07474939608800353

CSISZÁR, I. (1967). Information-type measures of difference of probability distributions and indirect observations. *Studia Sci. Math. Hungar.* **2** 299–318. MR0219345

CSÖRGŐ, S. (1985). Testing for independence by the empirical characteristic function. *J. Multivariate Anal.* **16** 290–299. MR0793494 https://doi.org/10.1016/0047-259X(85)90022-3

DATASTREAM (2018). Subscription service. Accessed June 2018.

DAVIS, R. A., MATSUI, M., MIKOSCH, T. and WAN, P. (2018). Applications of distance correlation to time series. *Bernoulli* **24** 3087–3116. MR3779711 https://doi.org/10.3150/17-BEJ955

DEHEUVELS, P. (1981a). A Kolmogorov–Smirnov type test for independence and multivariate samples. *Rev. Roumaine Math. Pures Appl.* **26** 213–226. MR0616038

DEHEUVELS, P. (1981b). An asymptotic decomposition for multivariate distribution-free tests of independence. *J. Multivariate Anal.* **11** 102–113. MR0612295 https://doi.org/10.1016/0047-259X(81)90136-6

DUECK, J., EDELMANN, D., GNEITING, T. and RICHARDS, D. (2014). The affinely invariant distance correlation. *Bernoulli* **20** 2305–2330. MR3263106 https://doi.org/10.3150/13-BEJ558

ENGLE, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* **50** 987–1007. MR0666121 https://doi.org/10.2307/1912773

ESCANCIANO, J. C. and HUALDE, J. (2019). Measuring asset market linkages: Nonlinear dependence and tail risk. *J. Bus. Econom. Statist.* 1–25.

ESCANCIANO, J. C. and VELASCO, C. (2006). Generalized spectral tests for the martingale difference hypothesis. *J. Econometrics* **134** 151–185. MR2328319 https://doi.org/10.1016/j.jeconom.2005.06.019

FERRATY, F. and VIEU, P. (2006). *Nonparametric Functional Data Analysis*: *Theory and Practice*. *Springer Series in Statistics*. Springer, New York. MR2229687

FISHER, R. A. (1915). Frequency distribution of the values of the correlation coefficient in samples of an indefinitely large population. *Biometrika* **10** 507–521.

FISHER, R. A. (1921). On the probable error of a coefficient of correlation deduced from a small sample. *Metron* **1** 3–32.

FOKIANOS, K. and PITSILLOU, M. (2017). Consistent testing for pairwise dependence in time series. *Technometrics* **59** 262–270. MR3635048 https://doi.org/10.1080/00401706.2016.1156024

FORBES, K. J. and RIGOBON, R. (2002). No contagion, only interdependence: Measuring stock market comovements. *J. Finance* **57** 2223–2261.

FRANCQ, C. and ZAKOÏAN, J.-M. (2011). *GARCH Models*: *Structure*, *Statistical Inference and Financial Applications*. Wiley, Chichester. MR3186556 https://doi.org/10.1002/9780470670057

GALTON, F. (1888). Co-relations and their measurement, chiefly from anthropometric data. *Proc. Roy. Soc. Lond.* **45** 135–145.

GALTON, F. (1890). Kinship and correlation. *N. Amer. Rev.* **150** 419–431.

GEBELEIN, H. (1941). Das statistische Problem der Korrelation als Variations- und Eigenwertproblem und sein Zusammenhang mit der Ausgleichsrechnung. *ZAMM Z. Angew. Math. Mech.* **21** 364–379. MR0007220 https://doi.org/10.1002/zamm.19410210604

GENEST, C., GHOUDI, K. and RÉMILLARD, B. (2007). Rank-based extensions of the Brock, Dechert, and Scheinkman test. *J. Amer. Statist. Assoc.* **102** 1363–1376. MR2372539 https://doi.org/10.1198/016214507000001076

GENEST, C. and NEŠLEHOVÁ, J. (2007). A primer on copulas for count data. *Astin Bull.* **37** 475–515. MR2422797 https://doi.org/10.2143/AST.37.2.2024077

GENEST, C., KOJADINOVIC, I., NEŠLEHOVÁ, J. and YAN, J. (2011). A goodness-of-fit test for bivariate extreme-value copulas. *Bernoulli* **17** 253–275. MR2797991 https://doi.org/10.3150/10-BEJ279

GHOUDI, K. and RÉMILLARD, B. (2018). Serial independence tests for innovations of conditional mean and variance models. *TEST* **27** 3–26. MR3764021 https://doi.org/10.1007/s11749-016-0521-3

GÓMEZ, E., GÓMEZ-VILLEGAS, M. A. and MARÍN, J. M. (2003). A survey on continuous elliptical vector distributions. *Rev. Mat. Complut.* **16** 345–361. MR2031887 https://doi.org/10.5209/rev_REMA.2003.v16.n1.16889

GORFINE, M., HELLER, R. and HELLER, Y. (2012). Comment on "Detecting novel associations in large data sets" by Reshef et al., Science Dec 16, 2011.

GRANGER, C. W., MAASOUMI, E. and RACINE, J. (2004). A dependence metric for possibly nonlinear processes. *J. Time Series Anal.* **25** 649–669. MR2086354 https://doi.org/10.1111/j.1467-9892.2004.01866.x

GRETTON, A. (2019). Introduction to RKHS, and some simple kernel algorithms. Unpublished manuscript, Lecture Notes Gatsby Computational Neuroscience Unit.

GRETTON, A. and GYÖRFI, L. (2010). Consistent nonparametric tests of independence. *J. Mach. Learn. Res.* **11** 1391–1423. MR2645456

GRETTON, A. and GYÖRFI, L. (2012). Strongly consistent nonparametric test of conditional independence. *J. Multivariate Anal.* **82** 1145–1150.

GRETTON, A., BOUSQUET, O., SMOLA, A. and SCHÖLKOPF, B. (2005). Measuring statistical dependence with Hilbert–Schmidt norms. In *Algorithmic Learning Theory* (S. Jain, U. Simon and E. Tomita, eds.). *Lecture Notes in Computer Science* **3734** 63–77. Springer, Berlin. MR2255909 https://doi.org/10.1007/11564089_7

HELLER, R., HELLER, Y. and GORFINE, M. (2013). A consistent multivariate test of association based on ranks of distances. *Biometrika* **100** 503–510. MR3068450 https://doi.org/10.1093/biomet/ass070

HJORT, N. L. and JONES, M. C. (1996). Locally parametric nonparametric density estimation. *Ann. Statist.* **24** 1619–1647. MR1416653 https://doi.org/10.1214/aos/1032298288

HOEFFDING, W. (1948). A non-parametric test of independence. *Ann. Math. Stat.* **19** 546–557. MR0029139 https://doi.org/10.1214/aoms/1177730150

HOEFFDING, W. (1963). Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.* **58** 13–30. MR0144363

HÖFFDING, W. (1940). Maszstabinvariante Korrelationstheorie. *Schr. Math. Inst. U. Inst. Angew. Math. Univ. Berlin* **5** 181–233. MR0004426

HOLLAND, P. W. and WANG, Y. J. (1987). Dependence function for continuous bivariate densities. *Comm. Statist. Theory Methods* **16** 863–876. MR0886560 https://doi.org/10.1080/03610928708829408

HONG, Y. (1999). Hypothesis testing in time series via the empirical characteristic function: A generalized spectral density approach. *J. Amer. Statist. Assoc.* **94** 1201–1220. MR1731483 https://doi.org/10.2307/2669935

HONG, Y. (2000). Generalized spectral tests for serial dependence. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **62** 557–574. MR1772415 https://doi.org/10.1111/1467-9868.00250

HONG, Y. and WHITE, H. (2005). Asymptotic distribution theory for nonparametric entropy measures of serial dependence. *Econometrica* **73** 837–901. MR2135144 https://doi.org/10.1111/j.1468-0262.2005.00597.x

HUANG, T.-M. (2010). Testing conditional independence using maximal nonlinear conditional correlation. *Ann. Statist.* **38** 2047–2091. MR2676883 https://doi.org/10.1214/09-AOS770

INCI, A. C., LI, H.-C. and MCCARTHY, J. (2011). Financial contagion: A local correlation analysis. *Res. Int. Bus. Finance* **25** 11–25.

JENTSCH, C., LEUCHT, A., MEYER, M. and BEERING, C. (2020). Empirical characteristic functions-based estimation and distance correlation for locally stationary processes. *J. Time Series Anal.* **41** 110–133. MR4048684 https://doi.org/10.1111/jtsa.12497

JOE, H. (2014). *Dependence Modeling with Copulas*. Chapman & Hall, London.

JONES, M. C. (1996). The local dependence function. *Biometrika* **83** 899–904. MR1440052 https://doi.org/10.1093/biomet/83.4.899

JONES, M. C. and KOCH, I. (2003). Dependence maps: Local dependence in practice. *Stat. Comput.* **13** 241–255. MR1982478 https://doi.org/10.1023/A:1024270700807

JORDANGER, L. A. (2020). LocalgaussSpec. Available at https://github.com/LAJordanger/localgaussSpec.

JORDANGER, L. A. and TJØSTHEIM, D. (2020). Nonlinear spectral analysis: A local Gaussian approach. *J. Amer. Statist. Assoc.* 1–55.

KENDALL, M. G. (1938). A new measure of rank correlation. *Biometrika* **30** 81–89.

KING, M. L. (1987). Testing for autocorrelation in linear regression models: A survey. In *Specification Analysis in the Linear Model* (M. L. King and D. E. A. Giles, eds.). *Internat. Lib. Econom.* 19–73. Routledge, London. MR0899966

KINNEY, J. B. and ATWAL, G. S. (2014). Equitability, mutual information, and the maximal information coefficient. *Proc. Natl. Acad. Sci. USA* **111** 3354–3359. MR3200177 https://doi.org/10.1073/pnas.1309933111

KLAASSEN, C. A. J. and WELLNER, J. A. (1997). Efficient estimation in the bivariate normal copula model: Normal margins are least favourable. *Bernoulli* **3** 55–77. MR1466545 https://doi.org/10.2307/3318652

KOJADINOVIC, I. and HOLMES, M. (2009). Tests of independence among continuous random vectors based on Cramér–von Mises functionals of the empirical copula process. *J. Multivariate Anal.* **100** 1137–1154. MR2508377 https://doi.org/10.1016/j.jmva.2008.10.013

KRASKOV, A., STÖGBAUER, H. and GRASSBERGER, P. (2004). Estimating mutual information. *Phys. Rev. E* (3) **69** 066138. MR2096503 https://doi.org/10.1103/PhysRevE.69.066138

LACAL, V. and TJØSTHEIM, D. (2017). Local Gaussian autocorrelation and tests for serial independence. *J. Time Series Anal.* **38** 51–71. MR3601314 https://doi.org/10.1111/jtsa.12195

LACAL, V. and TJØSTHEIM, D. (2019). Estimating and testing nonlinear local dependence between two time series. *J. Bus. Econom. Statist.* **37** 648–660. MR4016160 https://doi.org/10.1080/07350015.2017.1407777

LANCASTER, H. O. (1957). Some properties of the bivariate normal distribution considered in the form of a contingency table. *Biometrika* **44** 289–292.

LEHMANN, E. L. (1966). Some concepts of dependence. *Ann. Math. Stat.* **37** 1137–1153. MR0202228 https://doi.org/10.1214/aoms/1177699260

LOADER, C. R. (1996). Local likelihood density estimation. *Ann. Statist.* **24** 1602–1618. MR1416652 https://doi.org/10.1214/aos/1032298287

MARKOWITZ, H. M. (1952). Portfolio selection. *J. Finance* **7** 77–91.

MUSCAT, J. (2014). *Functional Analysis*: *An Introduction to Metric Spaces*, *Hilbert Spaces*, *and Banach Algebras*. Springer, Cham. MR3308576 https://doi.org/10.1007/978-3-319-06728-5

NELSEN, R. B. (1999). *An Introduction to Copulas*. *Lecture Notes in Statistics* **139**. Springer, New York. MR1653203 https://doi.org/10.1007/978-1-4757-3076-0

NGUYEN, Q. N., ABOURA, S., CHEVALLIER, J., ZHANG, L. and ZHU, B. (2020). Local Gaussian correlations in financial and commodity markets. *European J. Oper. Res.* **285** 306–323. MR4083070 https://doi.org/10.1016/j.ejor.2020.01.023

OTNEIM, H. (2019). `lg`: Locally gaussian distributions: Estimation and methods. Available at https://CRAN.R-project.org/package=lg.

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

OTNEIM, H. and TJØSTHEIM, D. (2017). The locally Gaussian density estimator for multivariate data. *Stat. Comput.* **27** 1595–1616. MR3687328 https://doi.org/10.1007/s11222-016-9706-6

OTNEIM, H. and TJØSTHEIM, D. (2018). Conditional density estimation using the local Gaussian correlation. *Stat. Comput.* **28** 303–321. MR3747565 https://doi.org/10.1007/s11222-017-9732-z

OTNEIM, H. and TJØSTHEIM, D. (2021). The locally Gaussian partial correlation. *J. Bus. Econom. Statist.* 1–33. To appear.

PEARSON, K. (1896). Mathematical contributions to the theory of evolution. III. Regression, heredity and panmixia. *Philos. Trans. R. Soc. Lond.* **187** 253–318.

PEARSON, K. (1922). *Francis Galton*: *A Centenary Appreciation*. Cambridge Univ. Press, Cambridge.

PEARSON, K. (1930). *The Life*, *Letters and Labors of Francis Galton*. Cambridge Univ. Press, Cambridge.

PFISTER, N. and PETERS, J. (2017). dHSIC: Independence testing via Hilbert Schmidt independence criterion. Available at https://CRAN.R-project.org/package=dHSIC.

PFISTER, N., BÜHLMANN, P., SCHÖLKOPF, B. and PETERS, J. (2018). Kernel-based tests for joint independence. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **80** 5–31. MR3744710 https://doi.org/10.1111/rssb.12235

PINKSE, J. (1998). A consistent nonparametric test for serial independence. *J. Econometrics* **84** 205–231. MR1630190 https://doi.org/10.1016/S0304-4076(97)00084-5

PRUDNIKOV, A. P., BRYCHKOV, Y. A. and MARICHEV, O. I. (1986). *Integrals and Series*. Gordon & Breach, New York.

RÉNYI, A. (1959). On measures of dependence. *Acta Math. Acad. Sci. Hung.* **10** 441–451. MR0115203 https://doi.org/10.1007/BF02024507

RESHEF, D. N., RESHEF, Y. A., FINUCANE, H. K., GROSSMAN, S. R., MCVEAN, G., TURNBAUGH, P. J., LANDER, E. S., MITZENMACHER, M. and SABETI, P. C. (2011). Detecting novel associations in large data sets. *Science* **334** 1518–1524. https://doi.org/10.1126/science.1205438

RESHEF, D., RESHEF, Y., MITZENMACHER, M. and SABETI, P. (2013). Equitability analysis of the maximal information coefficient, with comparisons.

RIZZO, M. L. and SZEKELY, G. J. (2018). Energy: E-statistics: Multivariate inference via the energy of data. Available at https://CRAN.R-project.org/package=energy.

ROBINSON, P. M. (1991). Consistent nonparametric entropy-based testing. *Rev. Econ. Stud.* **58** 437–453. MR1108130 https://doi.org/10.2307/2298005

ROSENBLATT, M. (1975). A quadratic measure of deviation of two-dimensional density estimates and a test of independence. *Ann. Statist.* **3** 1–14. MR0428579

SEJDINOVIC, D., SRIPERUMBUDUR, B., GRETTON, A. and FUKUMIZU, K. (2013). Equivalence of distance-based and RKHS-based statistics in hypothesis testing. *Ann. Statist.* **41** 2263–2291. MR3127866 https://doi.org/10.1214/13-AOS1140

SILVAPULLE, P. and GRANGER, C. W. J. (2001). Large returns, conditional correlation and portfolio diversification: A value-at-risk approach. *Quant. Finance* **1** 542–551. MR1863876 https://doi.org/10.1088/1469-7688/1/5/306

SKAUG, H. J. and TJØSTHEIM, D. (1993a). A nonparametric test of serial independence based on the empirical distribution function. *Biometrika* **80** 591–602. MR1248024 https://doi.org/10.1093/biomet/80.3.591

SKAUG, H. J. and TJØSTHEIM, D. (1993b). Nonparametric tests of serial independence. In *Developments in Time Series Analysis* (T. S. Rao, ed.) 207–229. CRC Press, London. MR1292268

SKAUG, H. J. and TJØSTHEIM, D. (1996). Testing for serial independence using measures of distance between densities. In *Athens Conference on Applied Probability and Time Series Analysis, Vol. II* (P. M. Robinson and M. Rosenblatt, eds.). *Lect. Notes Stat.* **115** 363–377. Springer, New York. MR1466759 https://doi.org/10.1007/978-1-4612-2412-9_27

SKLAR, M. (1959). Fonctions de Répartition à N Dimensions et Leurs Marges. Université Paris 8.

SPEARMAN, C. (1904). The proof and measurement of association between two things. *Am. J. Psychol.* **15** 72–101.

STANTON, J. M. (2001). Galton, Pearson, and the peas: A brief history of linear regression for statistics instructors. *J. Stat. Educ.* **9** 1–13.

STIGLER, S. M. (1989). Francis Galton's account of the invention of correlation. *Statist. Sci.* **4** 73–79. MR1007556

STØVE, B. and TJØSTHEIM, D. (2014). Asymmetric dependence patterns in financial returns: An empirical investigation using local Gaussian correlation. In *Essays in Nonlinear Time Series Econometrics* (M. Meitz N. Haldrup and P. Saikkonen, eds.) 307–329. Oxford Univ. Press, Oxford. MR3288225 https://doi.org/10.1093/acprof:oso/9780199679959.003.0013

STØVE, B. TJØSTHEIM, D. and HUFTHAMMER, K. (2014). Using local Gaussian correlation in a nonlinear re-examination of financial contagion. *J. Empir. Finance* **25** 785–801.

SU, L. and WHITE, H. (2007). A consistent characteristic function-based test for conditional independence. *J. Econometrics* **141** 807–834. MR2413488 https://doi.org/10.1016/j.jeconom.2006.11.006

SZÉKELY, G. J. (2002). $\mathcal{E}$-statistics: The energy of statistical samples. Technical report 02-16, Bowling Green State Univ., Bowling Green, OH.

SZÉKELY, G. J. and RIZZO, M. L. (2005). Hierarchical clustering via joint between-within distances: Extending Ward's minimum variance method. *J. Classification* **22** 151–183. MR2231170 https://doi.org/10.1007/s00357-005-0012-9

SZÉKELY, G. J. and RIZZO, M. L. (2009). Brownian distance covariance. *Ann. Appl. Stat.* **3** 1236–1265. MR2752127 https://doi.org/10.1214/09-AOAS312

SZÉKELY, G. J. and RIZZO, M. L. (2013). Energy statistics: A class of statistics based on distances. *J. Statist. Plann. Inference* **143** 1249–1272. MR3055745 https://doi.org/10.1016/j.jspi.2013.03.018

SZÉKELY, G. J. and RIZZO, M. L. (2014). Partial distance correlation with methods for dissimilarities. *Ann. Statist.* **42** 2382–2412. MR3269983 https://doi.org/10.1214/14-AOS1255

SZÉKELY, G. J., RIZZO, M. L. and BAKIROV, N. K. (2007). Measuring and testing dependence by correlation of distances. *Ann. Statist.* **35** 2769–2794. MR2382665 https://doi.org/10.1214/009053607000000505

TALEB, N. N. (2007). *The Black Swan*: *The Impact of the Highly Improbable*. Random House, New York.

TERÄSVIRTA, T., TJØSTHEIM, D. and GRANGER, C. W. J. (2010). *Modelling Nonlinear Economic Time Series*. *Advanced Texts in Econometrics*. Oxford Univ. Press, Oxford. MR3185399 https://doi.org/10.1093/acprof:oso/9780199587148.001.0001

TJØSTHEIM, D. and HUFTHAMMER, K. O. (2013). Local Gaussian correlation: A new measure of dependence. *J. Econometrics* **172** 33–48. MR2997128 https://doi.org/10.1016/j.jeconom.2012.08.001

TJØSTHEIM, D., OTNEIM, H. and STØVE, B. (2021). *Statistical Modeling Using Local Gaussian Approximation*. Elsevier, Amsterdam. To appear.

TJØSTHEIM, D., OTNEIM, H. and STØVE, B. (2022). Supplement to "Statistical Dependence: Beyond Pearson's ρ." https://doi.org/10.1214/21-STS823SUPP

VON NEUMANN, J. (1941). Distribution of the ratio of the mean square successive difference to the variance. *Ann. Math. Stat.* **12** 367–395. MR0006656 https://doi.org/10.1214/aoms/1177731677

VON NEUMANN, J. (1942). A further remark concerning the distribution of the ratio of the mean square successive difference to the variance. *Ann. Math. Stat.* **13** 86–88. MR0006657 https://doi.org/10.1214/aoms/1177731645

YAO, S., ZHANG, X. and SHAO, X. (2018). Testing mutual independence in high dimension via distance covariance. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **80** 455–480. MR3798874 https://doi.org/10.1111/rssb.12259

YENIGÜN, C. D., SZÉKELY, G. J. and RIZZO, M. L. (2011). A test of independence in two-way contingency tables based on maximal correlation. *Comm. Statist. Theory Methods* **40** 2225–2242. MR2862708 https://doi.org/10.1080/03610921003764274

ZHANG, K., PETERS, J., JANZING, D. and SCHÖLKOPF, B. (2012). Kernel-based conditional independence test and applications in causal discovery. In *Proceedings of the Uncertainty in Artificial Intelligence* 804–813. AUAI Press, Corvallis, OR.

ZHANG, Q., FILIPPI, S., GRETTON, A. and SEJDINOVIC, D. (2018). Large-scale kernel methods for independence testing. *Stat. Comput.* **28** 113–130. MR3741641 https://doi.org/10.1007/s11222-016-9721-7

ZHOU, Z. (2012). Measuring nonlinear dependence in time-series, a distance correlation approach. *J. Time Series Anal.* **33** 438–457. MR2915095 https://doi.org/10.1111/j.1467-9892.2011.00780.x

Statistical Science
2022, Vol. 37, No. 1, 110–122
https://doi.org/10.1214/21-STS824

Some Perspectives on Inference in High Dimensions

H. S. Battey and D. R. Cox

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

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

REFERENCES

ANASTASIOU, A. and REINERT, G. (2020). Bounds for the asymptotic distribution of the likelihood ratio. *Ann. Appl. Probab.* **30** 608–643. MR4108117 https://doi.org/10.1214/19-AAP1510

BARNDORFF-NIELSEN, O. (1983). On a formula for the distribution of the maximum likelihood estimator. *Biometrika* **70** 343–365. MR0712023 https://doi.org/10.1093/biomet/70.2.343

BARNDORFF-NIELSEN, O. E. (1984). On a formula for the distribution of the maximum likelihood estimator. *Scand. J. Stat.* **11** 157–170.

BARNDORFF-NIELSEN, O. E. (1988). *Parametric Statistical Models and Likelihood. Lecture Notes in Statistics* **50**. Springer, New York. MR0971982 https://doi.org/10.1007/978-1-4612-3934-5

BARNDORFF-NIELSEN, O. E. (1994). Adjusted versions of profile likelihood and directed likelihood, and extended likelihood. *J. Roy. Statist. Soc. Ser. B* **56** 125–140. MR1257801

BARNDORFF-NIELSEN, O. E. and COX, D. R. (1994). *Inference and Asymptotics. Monographs on Statistics and Applied Probability* **52**. CRC Press, London. MR1317097 https://doi.org/10.1007/978-1-4899-3210-5

BARTLETT, M. S. (1937). Properties of sufficiency and statistical tests. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **160** 268–282.

BATTEY, H. S. and COX, D. R. (2018). Large numbers of explanatory variables: A probabilistic assessment. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **474** 20170631.

BATTEY, H. S. and COX, D. R. (2020). High dimensional nuisance parameters: An example from parametric survival analysis. *Inf. Geom.* **3** 119–148. MR4182888 https://doi.org/10.1007/s41884-020-00030-6

BATTEY, H. S., COX, D. R. and JACKSON, M. V. (2019). On the linear in probability model for binary data. *R. Soc. Open Sci.* **6** 190067, 7. MR3980023 https://doi.org/10.1098/rsos.190067

BERK, R., BROWN, L., BUJA, A., ZHANG, K. and ZHAO, L. (2013). Valid post-selection inference. *Ann. Statist.* **41** 802–837. MR3099122 https://doi.org/10.1214/12-AOS1077

BLISS, C. I. (1935). The calculation of the dosage-mortality curve. *Ann. Appl. Biol.* **22** 135–167.

BOOTH, K. H. V. and COX, D. R. (1962). Some systematic supersaturated designs. *Technometrics* **4** 489–495. MR0184369 https://doi.org/10.2307/1266285

BRAZALLE, A., DAVISON, A. C. and REID, N. (2000). *Applied Asymptotics*. Cambridge Univ. Press, London.

BUJA, A., BROWN, L., BERK, R., GEORGE, E., PITKIN, E., TRASKIN, M., ZHANG, K. and ZHAO, L. (2019). Models as approximations I: Consequences illustrated with linear regression. *Statist. Sci.* **34** 523–544. MR4048582 https://doi.org/10.1214/18-STS693

CAI, T. T. and GUO, Z. (2017). Confidence intervals for high-dimensional linear regression: Minimax rates and adaptivity. *Ann. Statist.* **45** 615–646. MR3650395 https://doi.org/10.1214/16-AOS1461

CHERNOZHUKOV, V., CHETVERIKOV, D., DEMIRER, M., DUFLO, E., HANSEN, C., NEWEY, W. and ROBINS, J. (2018). Double/debiased machine learning for treatment and structural parameters. *Econom. J.* **21** C1–C68. MR3769544 https://doi.org/10.1111/ectj.12097

COX, D. R. (1972). Regression models and life-tables. *J. Roy. Statist. Soc. Ser. B* **34** 187–220. MR0341758

COX, D. R. (1975a). Partial likelihood. *Biometrika* **62** 269–276. MR0400509 https://doi.org/10.1093/biomet/62.2.269

COX, D. R. (1975b). A note on data-splitting for the evaluation of significance levels. *Biometrika* **62** 441–444. MR0378189 https://doi.org/10.1093/biomet/62.2.441

COX, D. R. (2020). Statistical significance. *Annu. Rev. Stat. Appl.* **7** 1–10. MR4104183 https://doi.org/10.1146/annurev-statistics-031219-041051

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

COX, D. R. and BATTEY, H. S. (2017). Large numbers of explanatory variables, a semi-descriptive analysis. *Proc. Natl. Acad. Sci. USA* **114** 8592–8595. https://doi.org/10.1073/pnas.1703764114

COX, D. R. and REID, N. (1987). Parameter orthogonality and approximate conditional inference. *J. Roy. Statist. Soc. Ser. B* **49** 1–39. With a discussion. MR0893334

COX, D. R. and REID, N. (2000). *The Theory of the Design of Experiments*. *Monographs on Statistics and Applied Probability* **67**. CRC Press, London. MR1456990

COX, D. R. and SNELL, E. J. (1974). The choice of variables in observational studies. *J. R. Stat. Soc. Ser. C. Appl. Stat.* **23** 51–59.

COX, D. R. and SNELL, E. J. (1989). *The Analysis of Binary Data*. *Monographs on Statistics and Applied Probability* **32**. CRC Press, London.

COX, D. R. and WERMUTH, N. (1996). *Multivariate Dependencies*: *Models*, *Analysis and Interpretation*. *Monographs on Statistics and Applied Probability* **67**. CRC Press, London. MR1456990

DAVISON, A. C., KOCH, E. and KOH, J. (2019). Comment: Models are approximations! [MR4048582; MR4048583]. *Statist. Sci.* **34** 584–590. MR4048589 https://doi.org/10.1214/19-STS746

DAVISON, A. C., REID, N. (2021). The tangent exponential model. arXiv:2106.10496.

FAN, Y., DEMIRKAYA, E. and LV, J. (2019). Nonuniformity of p-values can occur early in diverging dimensions. *J. Mach. Learn. Res.* **20** Paper No. 77, 33. MR3960931

FANG, E. X., NING, Y. and LIU, H. (2017). Testing and confidence intervals for high dimensional proportional hazards models. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **79** 1415–1437. MR3731669 https://doi.org/10.1111/rssb.12224

FISHER, R. A. (1922). On the mathematical foundations of theoretical statistics. *Philos. Trans. R. Soc. Lond. A* **222** 309–368.

FRASER, D. A. S. and REID, N. (1995). Ancillaries and third order significance. *Util. Math.* **47** 33–53. MR1330888

JAVANMARD, A. and MONTANARI, A. (2014). Confidence intervals and hypothesis testing for high-dimensional regression. *J. Mach. Learn. Res.* **15** 2869–2909. MR3277152

LEE, J. D., SUN, D. L., SUN, Y. and TAYLOR, J. E. (2016). Exact post-selection inference, with application to the lasso. *Ann. Statist.* **44** 907–927. MR3485948 https://doi.org/10.1214/15-AOS1371

LEI, L., BICKEL, P. J. and EL KAROUI, N. (2018). Asymptotics for high dimensional regression M-estimates: Fixed design results. *Probab. Theory Related Fields* **172** 983–1079. MR3877551 https://doi.org/10.1007/s00440-017-0824-7

LINDSAY, B. G. (1985). Using empirical partially Bayes inference for increased efficiency. *Ann. Statist.* **13** 914–931. MR0803748 https://doi.org/10.1214/aos/1176349646

LITTLE, R. J. A. and RUBIN, D. B. (1987). *Statistical Analysis with Missing Data*. *Wiley Series in Probability and Mathematical Statistics*: *Applied Probability and Statistics*. Wiley, New York. MR0890519

LOCKHART, R., TAYLOR, J., TIBSHIRANI, R. J. and TIBSHIRANI, R. (2014). A significance test for the lasso. *Ann. Statist.* **42** 413–468. MR3210970 https://doi.org/10.1214/13-AOS1175

MCCULLAGH, P. and POLSON, N. G. (2018). Statistical sparsity. *Biometrika* **105** 797–814. MR3877866 https://doi.org/10.1093/biomet/asy051

NING, Y. and LIU, H. (2017). A general theory of hypothesis tests and confidence regions for sparse high dimensional models. *Ann. Statist.* **45** 158–195. MR3611489 https://doi.org/10.1214/16-AOS1448

PATEL, L., GUSTAFSSON, N., LIN, Y., OBER, R., HENRIQUES, R. and COHEN, E. (2019). A hidden Markov model approach to characterizing the photo-switching behavior of fluorophores. *Ann. Appl. Stat.* **13** 1397–1429. MR4019144 https://doi.org/10.1214/19-AOAS1240

RASINES, D. G. and YOUNG, G. A. (2021). Splitting strategies for post-selection inference. arXiv:2102.02159.

SATTERTHWAITE, F. E. (1959). Random balance experimentation. *Technometrics* **1** 111–137. MR0130765 https://doi.org/10.2307/1266466

SHLOMOVICH, L., COHEN, E. A. K., ADAMS, N. and PATEL, L. (2020). A Monte Carlo EM algorithm for the parameter estimation of aggregated Hawkes processes. arXiv:2001.07160.

STEIN, C. (1956). Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, 1954–1955, *Vol. I* 197–206. Univ. California Press, Berkeley and Los Angeles. MR0084922

SUR, P. and CANDÈS, E. J. (2019). A modern maximum-likelihood theory for high-dimensional logistic regression. *Proc. Natl. Acad. Sci. USA* **116** 14516–14525. MR3984492 https://doi.org/10.1073/pnas.1810420116

TANG, Y. and REID, N. (2020). Modified likelihood root in high dimensions. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **82** 1349–1369. MR4176346

TIAN, X. and TAYLOR, J. E. (2018). Selective inference with a randomized response. *Ann. Statist.* **46** 679–710. https://doi.org/10.1214/17-AOS1564

TIBSHIRANI, R. (1996). Regression shrinkage and selection via the lasso. *J. Roy. Statist. Soc. Ser. B* **58** 267–288. MR1379242

TIBSHIRANI, R. J., TAYLOR, J., LOCKHART, R. and TIBSHIRANI, R. (2016). Exact post-selection inference for sequential regression procedures. *J. Amer. Statist. Assoc.* **111** 600–620. MR3538689 https://doi.org/10.1080/01621459.2015.1108848

VAN DE GEER, S., BÜHLMANN, P., RITOV, Y. and DEZEURE, R. (2014). On asymptotically optimal confidence regions and tests for high-dimensional models. *Ann. Statist.* **42** 1166–1202. MR3224285 https://doi.org/10.1214/14-AOS1221

VERSHYNIN, R. (2018). *High-Dimensional Probability*: *An Introduction with Applications in Data Science*. *Cambridge Series in Statistical and Probabilistic Mathematics* **47**. Cambridge Univ. Press, Cambridge. With a foreword by Sara van de Geer. MR3837109 https://doi.org/10.1017/9781108231596

WEATHERBURN, C. E. (1957). *An Introduction to Riemannian Geometry and the Tensor Calculus*. Cambridge Univ. Press, Cambridge, UK.

YATES, F. (1936). A new method of arranging variety trials involving a large number of varieties. *J. Agric. Sci.* **26** 424–455.

ZHANG, C.-H. and ZHANG, S. S. (2014). Confidence intervals for low dimensional parameters in high dimensional linear models. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **76** 217–242. MR3153940 https://doi.org/10.1111/rssb.12026

Statistical Science
2022, Vol. 37, No. 1, 123–142
https://doi.org/10.1214/21-STS825

Diffusion Smoothing for Spatial Point Patterns

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

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

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

REFERENCES

[1] ABRAMSON, I. S. (1982). On bandwidth variation in kernel estimates—a square root law. *Ann. Statist.* **10** 1217–1223. MR0673656

[2] BADDELEY, A., DAVIES, T. M, RAKSHIT, S., NAIR, G. and MCSWIGGAN, G. (2022). Supplement to "Diffusion Smoothing for Spatial Point Patterns." https://doi.org/10.1214/21-STS825SUPPA, https://doi.org/10.1214/21-STS825SUPPB

[3] BADDELEY, A., RUBAK, E. and TURNER, R. (2015). *Spatial Point Patterns*: *Methodology and Applications with R*. CRC Press/CRC, London.

[4] BADDELEY, A. and TURNER, R. (2005). Spatstat: An R package for analyzing spatial point patterns. *J. Stat. Softw.* **12** 1–42.

[5] BADDELEY, A., TURNER, R. and RUBAK, E. (2016). Adjusted composite likelihood ratio test for spatial Gibbs point processes. *J. Stat. Comput. Simul.* **86** 922–941. MR3437066 https://doi.org/10.1080/00949655.2015.1044530

[6] BADDELEY, A. J., MØLLER, J. and WAAGEPETERSEN, R. (2000). Non- and semi-parametric estimation of interaction in inhomogeneous point patterns. *Stat. Neerl.* **54** 329–350. MR1804002 https://doi.org/10.1111/1467-9574.00144

[7] BARRY, R. P. and MCINTYRE, J. (2011). Estimating animal densities and home range in regions with irregular boundaries and holes: A lattice-based alternative to the kernel density estimator. *Ecol. Model.* **222** 1666–1672.

[8] BARTLETT, M. S. (1964). The spectral analysis of two-dimensional point processes. *Biometrika* **51** 299–311. MR0175254 https://doi.org/10.1093/biomet/51.3-4.299

[9] BELL, M. and GRUNWALD, G. (2004). Mixed models for the analysis of replicated spatial point patterns. *Biostatistics* **5** 633–648.

[10] BERMAN, M. and DIGGLE, P. (1989). Estimating weighted integrals of the second-order intensity of a spatial point process. *J. Roy. Statist. Soc. Ser. B* **51** 81–92. MR0984995

[11] BITHELL, J. F. (1990). An application of density estimation to geographical epidemiology. *Stat. Med.* **9** 691–701. https://doi.org/10.1002/sim.4780090616

[12] BONEVA, L. I., KENDALL, D. and STEFANOV, I. (1971). Spline transformations: Three new diagnostic aids for the statistical

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

data-analyst (with discussion). *J. Roy. Statist. Soc. Ser. B* **33** 1–70. MR0288888

[13] BOTEV, Z. I., GROTOWSKI, J. F. and KROESE, D. P. (2010). Kernel density estimation via diffusion. *Ann. Statist.* **38** 2916–2957. MR2722460 https://doi.org/10.1214/10-AOS799

[14] BOYCE, W. E. and DIPRIMA, R. C. (1969). *Elementary Differential Equations and Boundary Value Problems*, 2nd ed. Wiley, New York.

[15] BRENNER, S. and SCOTT, R. (2007). *The Mathematical Theory of Finite Element Methods*. Springer, New York.

[16] BREZINSKI, C. and ZAGLIA, M. R. (1991). *Extrapolation Methods. Theory and Practice*. North-Holland, Amsterdam.

[17] BUTCHER, J. C. (2003). *Numerical Methods for Ordinary Differential Equations*. Wiley, Chichester. MR1993957 https://doi.org/10.1002/0470868279

[18] CAO, R., CUEVAS, A. and GONZÁLES-MANTEIGA, W. (1994). A comparative study of several smoothing methods in density estimation. *Comput. Statist. Data Anal.* **17** 153–176.

[19] CHAUDHURI, P. and MARRON, J. S. (2000). Scale space view of curve estimation. *Ann. Statist.* **28** 408–428. MR1790003 https://doi.org/10.1214/aos/1016218224

[20] CHOI, E. and HALL, P. (2001). Nonparametric analysis of earthquake point-process data. In *State of the Art in Probability and Statistics* (*Leiden*, 1999). *Institute of Mathematical Statistics Lecture Notes—Monograph Series* **36** 324–344. IMS, Beachwood, OH. MR1836568 https://doi.org/10.1214/lnms/1215090076

[21] CHUNG, M. K., QIU, A., SEO, S. and VORPERIAN, H. K. (2015). Unified heat kernel regression for diffusion, kernel smoothing and wavelets on manifolds and its application to mandible growth modeling in CT images. *Med. Image Anal.* **22** 63–76. https://doi.org/10.1016/j.media.2015.02.003

[22] CRESSIE, N. A. C. (1993). *Statistics for Spatial Data*, 2nd ed. Wiley, New York.

[23] CRONIE, O. and VAN LIESHOUT, M. N. M. (2018). A non-model-based approach to bandwidth selection for kernel estimators of spatial intensity functions. *Biometrika* **105** 455–462. MR3804414 https://doi.org/10.1093/biomet/asy001

[24] DALEY, D. J. and VERE-JONES, D. (1988). *An Introduction to the Theory of Point Processes. Springer Series in Statistics*. Springer, New York. MR0950166

[25] DALEY, D. J. and VERE-JONES, D. (2003). *An Introduction to the Theory of Point Processes. Vol. I: Elementary Theory and Methods*, 2nd ed. *Probability and Its Applications* (*New York*). Springer, New York. MR1950431

[26] DAVIES, T. M. and BADDELEY, A. (2018). Fast computation of spatially adaptive kernel estimates. *Stat. Comput.* **28** 937–956. MR3766052 https://doi.org/10.1007/s11222-017-9772-4

[27] DAVIES, T. M., FLYNN, C. R. and HAZELTON, M. L. (2018). On the utility of asymptotic bandwidth selectors for spatially adaptive kernel density estimation. *Statist. Probab. Lett.* **138** 75–81. MR3788721 https://doi.org/10.1016/j.spl.2018.02.067

[28] DAVIES, T. M. and LAWSON, A. B. (2019). An evaluation of likelihood-based bandwidth selectors for spatial and spatiotemporal kernel estimates. *J. Stat. Comput. Simul.* **89** 1131–1152. MR3928454 https://doi.org/10.1080/00949655.2019.1575066

[29] DAVIES, T. M., MARSHALL, J. C. and HAZELTON, M. L. (2018). Tutorial on kernel estimation of continuous spatial and spatiotemporal relative risk. *Stat. Med.* **37** 1191–1221. MR3777968 https://doi.org/10.1002/sim.7577

[30] DIGGLE, P. and MARRON, J. S. (1988). Equivalence of smoothing parameter selectors in density and intensity estimation. *J. Amer. Statist. Assoc.* **83** 793–800. MR0963807

[31] DIGGLE, P., ROWLINGSON, B. and SU, T. (2005). Point process methodology for on-line spatio-temporal disease surveillance. *Environmetrics* **16** 423–434. MR2147534 https://doi.org/10.1002/env.712

[32] DIGGLE, P. J. (1985). A kernel method for smoothing point process data. *J. R. Stat. Soc. Ser. C. Appl. Stat.* **34** 138–147.

[33] DIGGLE, P. J. (2014). *Statistical Analysis of Spatial and Spatio-Temporal Point Patterns*, 3rd ed. *Monographs on Statistics and Applied Probability* **128**. CRC Press, Boca Raton, FL. MR3113855

[34] DIGGLE, P. J., MATEU, J. and CLOUGH, H. E. (2000). A comparison between parametric and non-parametric approaches to the analysis of replicated spatial point patterns. *Adv. in Appl. Probab.* **32** 331–343. MR1778567 https://doi.org/10.1239/aap/1013540166

[35] DOYLE, P. G. and SNELL, J. L. (1984). *Random Walks and Electric Networks. Carus Mathematical Monographs* **22**. Mathematical Association of America, Washington, DC. MR0920811

[36] DUONG, T. and HAZELTON, M. L. (2003). Plug-in bandwidth matrices for bivariate kernel density estimation. *J. Nonparametr. Stat.* **15** 17–30. MR1958957 https://doi.org/10.1080/10485250306039

[37] DUONG, T. and HAZELTON, M. L. (2005). Convergence rates for unconstrained bandwidth matrix selectors in multivariate kernel density estimation. *J. Multivariate Anal.* **93** 417–433. MR2162646 https://doi.org/10.1016/j.jmva.2004.04.004

[38] DUONG, T. and HAZELTON, M. L. (2005). Cross-validation bandwidth matrices for multivariate kernel density estimation. *Scand. J. Stat.* **32** 485–506. MR2204631 https://doi.org/10.1111/j.1467-9469.2005.00445.x

[39] FUKUSHIMA, M. (1967). A construction of reflecting barrier Brownian motions for bounded domains. *Osaka J. Math.* **4** 183–215. MR0231444

[40] GHOSH, B. K. and HUANG, W.-M. (1992). Optimum bandwidths and kernels for estimating certain discontinuous densities. *Ann. Inst. Statist. Math.* **44** 563–577. MR1190148

[41] HALL, P. and MARRON, J. S. (1988). Variable window width kernel estimates of probability densities. *Probab. Theory Related Fields* **80** 37–49. MR0970470 https://doi.org/10.1007/BF00348751

[42] HOMINAL, P. and DEHEUVELS, P. (1979). Estimation non paramétrique de la densité compte-tenu d'informations sur le support. *Rev. Stat. Appl.* **27** 47–68. MR0552728

[43] ILLIAN, J., PENTTINEN, A., STOYAN, H. and STOYAN, D. (2008). *Statistical Analysis and Modelling of Spatial Point Patterns. Statistics in Practice*. Wiley, Chichester. MR2384630

[44] JONES, M. C. (1993). Simple boundary correction for kernel density estimation. *Stat. Comput.* **3** 135–146.

[45] JONES, M. C., MARRON, J. S. and SHEATHER, S. J. (1996). A brief survey of bandwidth selection for density estimation. *J. Amer. Statist. Assoc.* **91** 401–407. MR1394097 https://doi.org/10.2307/2291420

[46] KOENDERINK, J. J. (1984). The structure of images. *Biol. Cybernet.* **50** 363–370. MR0758126 https://doi.org/10.1007/BF00336961

[47] LINDEBERG, T. (1994). *Scale Space Theory in Computer Vision*. Kluwer, Boston, USA.

[48] LOADER, C. (1999). *Local Regression and Likelihood. Statistics and Computing*. Springer, New York. MR1704236

[49] LOADER, C. R. (1999). Bandwidth selection: Classical or plug-in? *Ann. Statist.* **27** 415–438. MR1714723 https://doi.org/10.1214/aos/1018031201

[50] MARSHALL, J. C. and HAZELTON, M. L. (2010). Boundary kernels for adaptive density estimators on regions with irregular boundaries. *J. Multivariate Anal.* **101** 949–963. MR2584911 https://doi.org/10.1016/j.jmva.2009.09.003

[51] McSwiggan, G., Baddeley, A. and Nair, G. (2017). Kernel density estimation on a linear network. *Scand. J. Stat.* **44** 324–345. MR3658517 https://doi.org/10.1111/sjos.12255

[52] McSwiggan, G., Baddeley, A. and Nair, G. (2020). Estimation of relative risk for events on a linear network. *Stat. Comput.* **30** 469–484. MR4064631 https://doi.org/10.1007/s11222-019-09889-7

[53] Microsoft Corporation and Weston, S. (2019). doParallel: Foreach Parallel Adaptor for the 'parallel' Package. R package version 1.0.15.

[54] Prince, M. I., Chetwynd, A., Diggle, P., Jarner, M., Metcalf, J. V. and James, O. F. W. (2001). The geographical distribution of primary biliary cirrhosis in a well-defined cohort. *Hepatology* **34** 1083–1088.

[55] Redmond, A. K. and Davies, T. M. (2018). spagmix: Artificial Spatial and Spatiotemporal Densities on Bounded Windows. R package version 0.3-1.

[56] Richardson, L. F. (1911). The approximate arithmetical solution by finite differences of physical problems including differential equations, with an application to the stresses in a masonry dam. *Philos. Trans. R. Soc. Lond. Ser. A* **210** 307–357.

[57] Richardson, L. F. and Gaunt, J. A. (1927). The deferred approach to the limit. Part I. Single lattice. Part II. Interpenetrating lattices. *Philos. Trans. R. Soc. Lond. Ser. A* **226** 299–361.

[58] Sain, S. R., Baggerly, K. A. and Scott, D. W. (1994). Cross-validation of multivariate densities. *J. Amer. Statist. Assoc.* **89** 807–817. MR1294726

[59] Schiesser, W. E. (1991). *The Numerical Method of Lines*. Academic Press, San Diego, CA. MR1117905

[60] Schuster, E. F. (1985). Incorporating support constraints into nonparametric estimators of densities. *Comm. Statist. Theory Methods* **14** 1123–1136. MR0797636 https://doi.org/10.1080/03610928508828965

[61] Scott, D. W. (1992). *Multivariate Density Estimation. Theory, Practice and Visualization*. Wiley, New York.

[62] Serra, J. (1982). *Image Analysis and Mathematical Morphology*. Academic Press, London.

[63] Silverman, B. W. (1982). Kernel density estimation using the fast Fourier transform. *Appl. Stat.* **31** 93–99.

[64] Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis. Monographs on Statistics and Applied Probability*. CRC Press, London. MR0848134 https://doi.org/10.1007/978-1-4899-3324-9

[65] Smith, B. (2018). Fantastic Pā and Where to Find Them: A Spatial Analysis of the pre-European Pā of Aotearoa Master's thesis Univ. Otago Dunedin, New Zealand. Available at http://hdl.handle.net/10523/8524.

[66] Smith, B. A., Thomas, T. and Davies, T. M. (2021). Statistical approaches to spatial variation in conflict and social organisation: Assessing the emergent distribution of fortification works (pā) in New Zealand/Aotearoa. Submitted for publication.

[67] Stachelek, J. and Madden, C. J. (2015). Application of inverse path distance weighting for high-density spatial mapping of water quality patterns. *Int. J. Geogr. Inf. Sci.* **29** 1240–1250.

[68] Stroock, D. W. and Varadhan, S. R. S. (1971). Diffusion processes with boundary conditions. *Comm. Pure Appl. Math.* **24** 147–225. MR0277037 https://doi.org/10.1002/cpa.3160240206

[69] Terrell, G. R. (1990). The maximal smoothing principle in density estimation. *J. Amer. Statist. Assoc.* **85** 470–477. MR1141750

[70] Terrell, G. R. and Scott, D. W. (1992). Variable kernel density estimation. *Ann. Statist.* **20** 1236–1265. MR1186249 https://doi.org/10.1214/aos/1176348768

[71] Wand, M. P. and Jones, M. C. (1995). *Kernel Smoothing. Monographs on Statistics and Applied Probability* **60**. CRC Press, London. MR1319818 https://doi.org/10.1007/978-1-4899-4493-1

[72] Weickert, J. (1997). *Anisotropic Diffusion in Image Processing*. Teubner, Stuttgart.

[73] Zhang, X., King, M. L. and Hyndman, R. J. (2006). A Bayesian approach to bandwidth selection for multivariate kernel density estimation. *Comput. Statist. Data Anal.* **50** 3009–3031. MR2239655 https://doi.org/10.1016/j.csda.2005.06.019

Statistical Science
2022, Vol. 37, No. 1, 143–158
https://doi.org/10.1214/21-STS829

A Conversation with Ross Prentice

Li Hsu and Charles Kooperberg

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

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

Dr. Prentice has received numerous awards for his work, including an honorary doctorate in mathematics from the University of Waterloo, the Mantel Award for Lifetime Contributions to Statistics in Epidemiology from the American Statistical Association, the Mortimer Spiegelman Award from the American Public Health Association, the Committee of Presidents of Statistical Societies Presidents' Award and RA Fisher Award, the Marvin Zelen Leadership Award for Outstanding Achievement in Statistical Science from Harvard University, the American Association of Cancer Research/American Cancer Society Award for Research Excellence in Cancer Epidemiology and Prevention, and the American Association for Cancer Research Team Science Award. He was elected to the Institute of Medicine/National Academy of Medicine in 1990. The Ross L. Prentice Endowed Professorship of Biostatistical Collaboration was created at the UW in 2005 and has been awarded every year since its inception. The interior space of the Public Health Sciences building at the Hutch has been named the Ross L. Prentice Atrium. In his spare time, Ross enjoys sports including water skiing, golf, running, and spending time with his wife, Didi, and with his daughters, sons-in-law, and grandchildren. He ran daily from when he was in his 20s until his knees objected about 10 years ago.

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

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

Li Hsu is Professor, Biostatistics Program, Fred Hutchinson Cancer Research Center, 1100 Fairview Avenue North, M2B500, Seattle, Washington 19024, USA (e-mail: lih@fredhutch.org). Charles Kooperberg is Professor and Co-Program Head, Biostatistics Program, Fred Hutchinson Cancer Research Center, 1100 Fairview Avenue North, M2B500, Seattle, Washington 19024, USA (e-mail: clk@fredhutch.org).

REFERENCES

BRESLOW, N. E. and DAY, N. E. (1980). *Statistical Methods in Cancer Research*. *Vol* 1. *The Analysis of Case-Control Studies*. International Agency for Research on Cancer Scientific Publications, Lyon, France.

CAI, J. and HSU, L. (2013). Preface *Lifetime Data Anal*. **19** 437–1441.

CHLEBOWSKI, R. T., ARAGAKI, A. K., ANDERSON, G. L., PAN, K., NEUHOUSER, M. L., MANSON, J. E., THOMSON, C. A., MOSSAVAR-RAHMANI, Y., LANE, D. S. et al. (2020). Dietary modification and breast cancer mortality: Long-term follow-up of the Women's Health Initiative randomized trial. *J. Clin. Oncol*. **38** 1419–1428.

COX, D. R. (1972). Regression models and life-tables (with discussion). *J. Roy. Statist. Soc. Ser. B* **34** 187–220. MR0341758

KALBFLEISCH, J. D. and PRENTICE, R. L. (1972). Discussion of the paper by D. R. Cox. *J. Roy. Statist. Soc. Ser. B* **34** 215–216.

KALBFLEISCH, J. D. and PRENTICE, R. L. (1980). *The Statistical Analysis of Failure Time Data*. *Wiley Series in Probability and Mathematical Statistics*. Wiley, New York-Chichester-Brisbane. MR0570114

KALBFLEISCH, J. D. and PRENTICE, R. L. (2002). *The Statistical Analysis of Failure Time Data*, 2nd ed. *Wiley Series in Probability and Statistics*. Wiley Interscience, Hoboken, NJ. MR1924807 https://doi.org/10.1002/9781118032985

LAMPE, J. W., HUANG, Y., NEUHOUSER, M. L., TINKER, L. F., SONG, X., SCHOELLER, D. A., KIM, S., RAFTERY, D., DI, C. et al. (2017). Dietary biomarker evaluation in a controlled feeding study in women from the Women's Health Initiative cohort. *Am. J. Clin. Nutr.* **105** 466–475.

LIN, D. (2013). An overview of Ross Prentice's contributions to statistical science. *Stat. Biosci*. **5** 224–231.

PRENTICE, R. L. (1989). Surrogate endpoints in clinical trials: Definition and operational criteria. *Stat. Med*. **8** 431–440.

PRENTICE, R. L. and PYKE, R. (1979). Logistic disease incidence models and case-control studies. *Biometrika* **66** 403–411. MR0556730 https://doi.org/10.1093/biomet/66.3.403

PRENTICE, R. L. and ZHAO, S. (2019). *The Statistical Analysis of Multivariate Failure Time Data*: *A Marginal Modeling Approach*. *Monographs on Statistics and Applied Probability* **163**. CRC Press, Boca Raton, FL. MR3966434 https://doi.org/10.1201/9780429162367

PRENTICE, R. L. and ZHAO, S. (2020). Regression models and multivariate life tables. *J. Amer. Statist. Assoc*. [Epub ahead of print]. https://doi.org/10.1080/01621459.2020.1713792

PRENTICE, R. L., PETTINGER, M., NEUHOUSER, M. L., TINKER, L. F., HUANG, Y., ZHENG, C., MANSON, J. E., MOSSAVAR-RAHMANI, Y., ANDERSON, G. L. et al. (2019a). Application of blood concentration biomarkers in nutritional epidemiology: Example of carotenoid and tocopherol intake in relation to chronic disease risk. *Am. J. Clin. Nutr.* **109** 1189–1196.

PRENTICE, R. L., ARAGAKI, A. K., HOWARD, B. V., CHLEBOWSKI, R. T., THOMSON, C. A., VAN HORN, L., TINKER, L. F., MANSON, J. E., ANDERSON, G. L. et al. (2019b). Low-fat dietary pattern among postmenopausal women influences long-term cancer, cardiovascular disease, and diabetes outcomes. *J. Nutr.* **149** 1565–1574.

PRENTICE, R. L., ARAGAKI, A. K., CHLEBOWSKI, R. T., ZHAO, S., ANDERSON, G. L., ROSSOUW, J. E., WALLACE, R., BANACK, H., SHADYAB, A. H. et al. (2020). Dual-outcome intention-to-treat analyses in the Women's Health Initiative randomized controlled hormone therapy trials. *Am. J. Epidemiol*. **189** 972–981.

PRENTICE, R. L., PETTINGER, M., NEUHOUSER, M. L., RAFTERY, D., ZHENG, C., NAGANA GORUDA, G. A., HUANG, Y., TINKER, L. F., HOWARD, B. V. et al. (2021). Biomarker-calibrated macronutrient intake and chronic disease risk among postmenopausal women. *J. Nutr.* [Epub ahead of print]. https://doi.org/10.1093/jn/nxab091

ROSSOUW, J., ANDERSON, G., PRENTICE, R., LACROIX, A., KOOPERBERG, C., STEFANICK, M., JACKSON, R., BERESFORD, S., HOWARD, B. et al. (2002). Risks and benefits of estrogen plus progestin in healthy postmenopausal women: Principal results From the Women's Health Initiative randomized controlled trial. *J. Am. Med. Assoc*. **288** 321–333.

SMITH, D. C., PRENTICE, R., THOMPSON, D. J. and HERRMANN, W. L. (1975). Association of exogenous estrogen and endometrial carcinoma. *N. Engl. J. Med*. **293** 1164–1167.

STORB, R., PRENTICE, R. L. and THOMAS, E. D. (1977). Marrow transplantation for treatment of aplastic anemia: An analysis of factors associated with graft rejection. *N. Engl. J. Med*. **296** 61–66.

THE WOMEN'S HEALTH INITIATIVE STEERING COMMITTEE (2004). Effects of conjugated equine estrogen in postmenopausal women with hysterectomy. *J. Am. Med. Assoc*. **291** 1701–1712.

WEIDEN, P. L., FLOURNOY, N., THOMAS, E. D., PRENTICE, R., FEFER, A., BUCKNER, C. D. and STORB, R. (1979). Antileukemic effect of graft-versus-host disease in human recipients of allogeneic-marrow grafts. *N. Engl. J. Med*. **300** 1068–1073.

INSTITUTE OF MATHEMATICAL STATISTICS

(Organized September 12, 1935)

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

IMS OFFICERS

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

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

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

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

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

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

IMS EDITORS

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

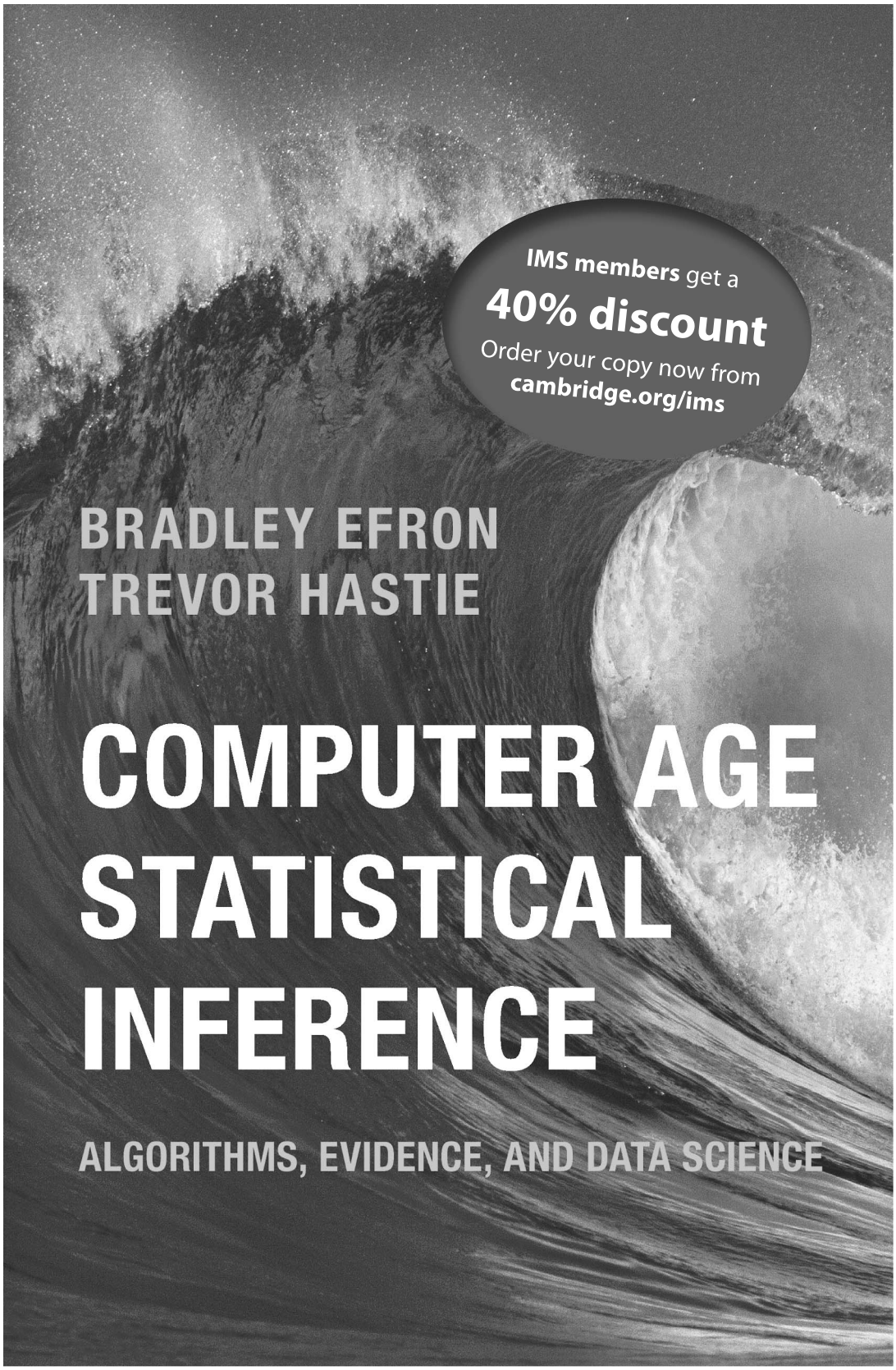
The Annals of Applied Statistics. *Editor-in-Chief*: Ji Zhu, Department of Statistics, University of Michigan, Ann Arbor, MI 48109, USA

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

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

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

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


IMS members get a
40% discount
Order your copy now from
cambridge.org/ims
BRADLEY EFRON
TREVOR HASTIE
COMPUTER AGE
STATISTICAL
INFERENCE
ALGORITHMS, EVIDENCE, AND DATA SCIENCE