STATISTICAL SCIENCE

Volume 37, Number 4 November 2022

***Statistical Science* [ISSN 0883-4237 (print); ISSN 2168-8745 (online)]**, Volume 37, Number 4, November 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 4 (455–629) November 2022

Volume 37 **Number 4** **November 2022**

Statistical Science
2022, Vol. 37, No. 4, 455–472
https://doi.org/10.1214/21-STS833

Interpreting p-Values and Confidence Intervals Using Well-Calibrated Null Preference Priors

Michael P. Fay, Michael A. Proschan, Erica H. Brittain and Ram Tiwari

Abstract. We propose well-calibrated null preference priors for use with one-sided hypothesis tests, such that resulting Bayesian and frequentist inferences agree. Null preference priors mean that they have nearly 100% of their prior belief in the null hypothesis, and well-calibrated priors mean that the resulting posterior beliefs in the alternative hypothesis are not overconfident. This formulation expands the class of problems giving Bayes-frequentist agreement to include problems involving discrete distributions such as binomial and negative binomial one- and two-sample exact (i.e., valid) tests. When applicable, these priors give posterior belief in the null hypothesis that is a valid p-value, and the null preference prior emphasizes that large p-values may simply represent insufficient data to overturn prior belief. This formulation gives a Bayesian interpretation of some common frequentist tests, as well as more intuitively explaining lesser known and less straightforward confidence intervals for two-sample tests.

Key words and phrases: Bayesian calibration, Bayesian hypothesis testing, frequentist matching priors, objective priors, probability matching priors.

REFERENCES

ALTHAM, P. M. E. (1969). Exact Bayesian analysis of a 2×2 contingency table, and Fisher's "exact" significance test. *J. Roy. Statist. Soc. Ser. B* **31** 261–269. MR0269016

BASU, D. (1975). Statistical information and likelihood. *Sankhyā Ser. A* **37** 1–71. MR0440747

BAYARRI, M. J. and BERGER, J. O. (2004). The interplay of Bayesian and frequentist analysis. *Statist. Sci.* **19** 58–80. MR2082147 https://doi.org/10.1214/088342304000000116

BENJAMIN, D. J., BERGER, J. O., JOHANNESSON, M., NOSEK, B. A., WAGENMAKERS, E.-J., BERK, R., BOLLEN, K. A., BREMBS, B., BROWN, L. et al. (2018). Redefine statistical significance. *Nat. Hum. Behav.* **2** 6–10.

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

BERGER, J. O. (2003). Could Fisher, Jeffreys and Neyman have agreed on testing? *Statist. Sci.* **18** 1–32. MR1997064 https://doi.org/10.1214/ss/1056397485

BERGER, J. O., BERNARDO, J. M. and SUN, D. (2009). The formal definition of reference priors. *Ann. Statist.* **37** 905–938. MR2502655 https://doi.org/10.1214/07-AOS587

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

BERGER, J. O. and DELAMPADY, M. (1987). Testing precise hypotheses. *Statist. Sci.* **2** 317–335. MR0920141

BERGER, J. O. and SELLKE, T. (1987). Testing a point null hypothesis: Irreconcilability of P values and evidence. *J. Amer. Statist. Assoc.* **82** 112–139. MR0883340

BERNARDO, J. M. (2011). Integrated objective Bayesian estimation and hypothesis testing. In *Bayesian Statistics* **9** 1–68. Oxford Univ. Press, Oxford. MR3204001 https://doi.org/10.1093/acprof:oso/9780199694587.003.0001

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

CARLIN, B. P. and LOUIS, T. A. (2009). *Bayesian Methods for Data Analysis*, 3rd ed. *Texts in Statistical Science Series*. CRC Press, Boca Raton, FL. MR2442364

Michael P. Fay is Mathematical Statisticians at the Biostatistics Research Branch, National Institute of Allergy and Infectious Diseases, 5601 Fishers Lane, Rockville, Maryland 20852 USA (e-mail: mfay@niaid.nih.gov). Michael A. Proschan is Mathematical Statisticians at the Biostatistics Research Branch, National Institute of Allergy and Infectious Diseases, 5601 Fishers Lane, Rockville, Maryland 20852 USA (e-mail: proscham@niaid.nih.gov). Erica H. Brittain is Mathematical Statisticians at the Biostatistics Research Branch, National Institute of Allergy and Infectious Diseases, 5601 Fishers Lane, Rockville, Maryland 20852 USA (e-mail: ebrittain@niaid.nih.gov). Ram Tiwari is Head of Statistical Methodology, Bristol Myers Squibb, Global Biometrics and Data Sciences, Bristol Myers Squibb, 3401 Princeton Pike, Lawrence Township, New Jersey, 08648 USA (e-mail: Ram.Tiwari@bms.com).

CASELLA, G. and BERGER, R. L. (1987a). Reconciling Bayesian and frequentist evidence in the one-sided testing problem. *J. Amer. Statist. Assoc.* **82** 106–111, 123–139. MR0883339

CASELLA, G. and BERGER, R. L. (1987b). Comment on 'Testing precise hypotheses' by Berger and Delampady. *Statist. Sci.* **2** 344–347.

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

CLOPPER, C. and PEARSON, E. S. (1934). The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika* **26** 404–413.

COX, D. R. (1958). Some problems connected with statistical inference. *Ann. Math. Stat.* **29** 357–372. MR0094890 https://doi.org/10.1214/aoms/1177706618

DATTA, G. S. and MUKERJEE, R. (2004). *Probability Matching Priors: Higher Order Asymptotics. Lecture Notes in Statistics* **178**. Springer, New York. MR2053794 https://doi.org/10.1007/978-1-4612-2036-7

DATTA, G. S. and SWEETING, T. J. (2005). Probability matching priors. In *Bayesian Thinking: Modeling and Computation. Handbook of Statist.* **25** 91–114. Elsevier/North-Holland, Amsterdam. MR2490523 https://doi.org/10.1016/S0169-7161(05)25003-4

DAWID, A. P. (1982). The well-calibrated Bayesian. *J. Amer. Statist. Assoc.* **77** 605–613. MR0675887

FAY, M. P. and HUNSBERGER, S. A. (2021). Practical valid inferences for the two-sample binomial problem. *Stat. Surv.* **15** 72–110. MR4255287 https://doi.org/10.1214/21-ss131

FAY, M. P., PROSCHAN, M. A. and BRITTAIN, E. (2015). Combining one-sample confidence procedures for inference in the two-sample case. *Biometrics* **71** 146–156. MR3335359 https://doi.org/10.1111/biom.12231

FAY, M. P., PROSCHAN, M. A., BRITTAIN, E. H. and TIWARI, R. (2022). Supplement to "Interpreting p-Values and Confidence Intervals Using Well-Calibrated Null Preference Priors." https://doi.org/10.1214/21-STS833SUPP

FREEDMAN, B. (1987). Equipoise and the ethics of clinical research. *N. Engl. J. Med.* **317** 141–145.

FREEDMAN, L. S. (2008). An analysis of the controversy over classical one-sided tests. *Clin. Trials* **5** 635–640.

GELMAN, A., SIMPSON, D. and BETANCOURT, M. (2017). The prior can often only be understood in the context of the likelihood. *Entropy* **19** 555.

GELMAN, A., CARLIN, J. B., STERN, H. S., DUNSON, D. B., VEHTARI, A. and RUBIN, D. B. (2014). *Bayesian Data Analysis*, 3rd ed. *Texts in Statistical Science Series*. CRC Press, Boca Raton, FL. MR3235677

GHOSH, M. (2011). Objective priors: An introduction for frequentists. *Statist. Sci.* **26** 187–202. MR2858380 https://doi.org/10.1214/10-STS338

GREENLAND, S. (2016). The ASA guidelines and null bias in current teaching and practice. *Amer. Statist.* **70** 10.

GREENLAND, S., SENN, S. J., ROTHMAN, K. J., CARLIN, J. B., POOLE, C., GOODMAN, S. N. and ALTMAN, D. G. (2016). Statistical tests, p values, confidence intervals, and power: A guide to misinterpretations. *Eur. J. Epidemiol.* **31** 337–350.

HANNIG, J., IYER, H., LAI, R. C. S. and LEE, T. C. M. (2016). Generalized fiducial inference: A review and new results. *J. Amer. Statist. Assoc.* **111** 1346–1361. MR3561954 https://doi.org/10.1080/01621459.2016.1165102

IOANNIDIS, J. P. A. (2005). Why most published research findings are false. *PLoS Med.* **2** e124.

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

KLEIJN, B. J. K. and VAN DER VAART, A. W. (2012). The Bernstein-Von-Mises theorem under misspecification. *Electron. J. Stat.* **6** 354–381. MR2988412 https://doi.org/10.1214/12-EJS675

LINDLEY, D. V. (1957). Binomial sampling schemes and the concept of information. *Biometrika* **44** 179–186. MR0087273 https://doi.org/10.1093/biomet/44.1-2.179

LINDLEY, D. V. (1958). Fiducial distributions and Bayes' theorem. *J. Roy. Statist. Soc. Ser. B* **20** 102–107. MR0095550

LITTLE, R. J. (2006). Calibrated Bayes: A Bayes/frequentist roadmap. *Amer. Statist.* **60** 213–223. MR2246754 https://doi.org/10.1198/000313006X117837

MAYO, D. G. (2014). On the Birnbaum argument for the strong likelihood principle. *Statist. Sci.* **29** 227–239. MR3264534 https://doi.org/10.1214/13-STS457

MENG, X.-L. (1994). Posterior predictive p-values. *Ann. Statist.* **22** 1142–1160. MR1311969 https://doi.org/10.1214/aos/1176325622

MORRIS, C. N. (1982). Natural exponential families with quadratic variance functions. *Ann. Statist.* **10** 65–80. MR0642719

MORRIS, C. (1987). Testing a point null hypothesis: The irreconcilability of p values and evidence: Comment. *J. Amer. Statist. Assoc.* **82** 131–133.

PEÑA, V. and BERGER, J. O. (2017). 'A note on recent criticisms to Birnbaum's theorem'. Available at arXiv:1711.08093v1.

RIPAMONTI, E., LLOYD, C. and QUATTO, P. (2017). Contemporary frequentist views of the 2×2 binomial trial. *Statist. Sci.* **32** 600–615. MR3730524 https://doi.org/10.1214/17-STS627

ROBERT, C. P. (2007). *The Bayesian Choice: From Decision-Theoretic Foundations to Computational Implementation*, 2nd ed. *Springer Texts in Statistics*. Springer, New York. MR2723361

RÖHMEL, J. (2005). Problems with existing procedures to calculate exact unconditional p-values for non-inferiority/superiority and confidence intervals for two binomials and how to resolve them. *Biom. J.* **47** 37–47. MR2135888 https://doi.org/10.1002/bimj.200410086

ROUSSEAU, J. (2016). On the frequentist properties of Bayesian nonparametric methods. *Annu. Rev. Stat. Appl.* **3** 211–231.

SCHERVISH, M. J. (1996). *P* values: What they are and what they are not. *Amer. Statist.* **50** 203–206. MR1422069 https://doi.org/10.2307/2684655

SPIEGELHALTER, D. J., FREEDMAN, L. S. and PARMAR, M. K. B. (1994). Bayesian approaches to randomized trials. *J. Roy. Statist. Soc. Ser. A* **157** 357–416. MR1321308 https://doi.org/10.2307/2983527

STEVENS, W. L. (1950). Fiducial limits of the parameter of a discontinuous distribution. *Biometrika* **37** 117–129. MR0035955 https://doi.org/10.1093/biomet/37.1-2.117

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

VERONESE, P. and MELILLI, E. (2015). Fiducial and confidence distributions for real exponential families. *Scand. J. Stat.* **42** 471–484. MR3345116 https://doi.org/10.1111/sjos.12117

VERONESE, P. and MELILLI, E. (2018). Fiducial, confidence and objective Bayesian posterior distributions for a multidimensional parameter. *J. Statist. Plann. Inference* **195** 153–173. MR3760845 https://doi.org/10.1016/j.jspi.2017.09.015

WAGENMAKERS, E.-J., LODEWYCKX, T., KURIYAL, H. and GRASMAN, R. (2010). Bayesian hypothesis testing for psychologists: A tutorial on the Savage–Dickey method. *Cogn. Psychol.* **60** 158–189.

WANG, W. (2010). On construction of the smallest one-sided confidence interval for the difference of two proportions. *Ann. Statist.* **38** 1227–1243. MR2604711 https://doi.org/10.1214/09-AOS744

WASSERSTEIN, R. L. and LAZAR, N. A. (2016). The ASA's statement on p-values: Context, process, and purpose. *Amer. Statist.* **70** 129–133. MR3511040 https://doi.org/10.1080/00031305.2016.1154108

WASSERSTEIN, R. L. SCHIRM, A. L. and LAZAR, N. A. (2019). Moving to a world beyond "$p < 0.05$". *Amer. Statist.* **73** 1–19.

WELCH, B. L. and PEERS, H. W. (1963). On formulae for confidence points based on integrals of weighted likelihoods. *J. Roy. Statist. Soc. Ser. B* **25** 318–329. MR0173309

XIE, M. and SINGH, K. (2013). Confidence distribution, the frequentist distribution estimator of a parameter: A review. *Int. Stat. Rev.* **81** 3–39. MR3047496 https://doi.org/10.1111/insr.12000

ZABELL, S. L. (1992). R. A. Fisher and the fiducial argument. *Statist. Sci.* **7** 369–387. MR1181418

ZELLNER, A. (1984). Posterior odds ratios for regression hypotheses: General considerations and some specific results. In *Basic Issues in Econometrics* (A. Zellner, ed.) 275–305. Univ. Chicago Press, Chicago, IL.

Statistical Science
2022, Vol. 37, No. 4, 473–493
https://doi.org/10.1214/21-STS834

Measurement Error Models: From Nonparametric Methods to Deep Neural Networks

Zhirui Hu, Zheng Tracy Ke and Jun S. Liu

Abstract. The success of deep learning has inspired a lot of recent interests in exploiting neural network structures for statistical inference and learning. In this paper, we review some popular deep neural network structures and techniques under the framework of nonparametric regression with measurement errors. In particular, we demonstrate how to use a fully connected feed-forward neural network to approximate the regression function $f(x)$, explain how to use a normalizing flow to approximate the prior distribution of X, and detail how to construct an inference network to generate approximate posterior samples of X. After reviewing recent advances in variational inference for deep neural networks, such as the importance weighted autoencoder, doubly reparametrized gradient estimator, and nonlinear independent components estimation, we describe an inference procedure built upon these advances. An extensive numerical study is presented to compare the neural network approach with classical nonparametric methods, which suggests that the neural network approach is more flexible in accommodating different classes of regression functions and performs superior or comparable to the best available method in many settings.

Key words and phrases: Doubly reparametrized gradient estimator, error-in-variables, fully connected feed-forward neural network, importance weighted autoencoder, normalizing flow, SIMEX, variational autoencoder.

REFERENCES

BARTLETT, P. L. (1998). The sample complexity of pattern classification with neural networks: The size of the weights is more important than the size of the network. *IEEE Trans. Inf. Theory* **44** 525–536. MR1607706 https://doi.org/10.1109/18.661502

BAUER, B. and KOHLER, M. (2019). On deep learning as a remedy for the curse of dimensionality in nonparametric regression. *Ann. Statist.* **47** 2261–2285. MR3953451 https://doi.org/10.1214/18-AOS1747

BERRY, S. M., CARROLL, R. J. and RUPPERT, D. (2002). Bayesian smoothing and regression splines for measurement error problems. *J. Amer. Statist. Assoc.* **97** 160–169. MR1947277 https://doi.org/10.1198/016214502753479301

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

BURDA, Y., GROSSE, R. and SALAKHUTDINOV, R. (2016). Importance weighted autoencoders. In *International Conference on Learning Representations*.

BUTUCEA, C. and TAUPIN, M.-L. (2008). New M-estimators in semi-parametric regression with errors in variables. *Ann. Inst. Henri Poincaré Probab. Stat.* **44** 393–421. MR2451051 https://doi.org/10.1214/07-AIHP107

BUZAS, J. S. (1997). Instrumental variable estimation in nonlinear measurement error models. *Comm. Statist. Theory Methods* **26** 2861–2877. MR1491231 https://doi.org/10.1080/03610929708832083

CAHILL, N., KEMP, A. C., HORTON, B. P. and PARNELL, A. C. (2015). Modeling sea-level change using errors-in-variables integrated Gaussian processes. *Ann. Appl. Stat.* **9** 547–571. MR3371325 https://doi.org/10.1214/15-AOAS824

CARROLL, R. J. and HALL, P. (1988). Optimal rates of convergence for deconvolving a density. *J. Amer. Statist. Assoc.* **83** 1184–1186. MR0997599

CARROLL, R. J., MACA, J. D. and RUPPERT, D. (1999). Nonparametric regression in the presence of measurement error. *Biometrika* **86** 541–554. MR1723777 https://doi.org/10.1093/biomet/86.3.541

CARROLL, R. J., ROEDER, K. and WASSERMAN, L. (1999). Flexible parametric measurement error models. *Biometrics* **55** 44–54.

Zhirui Hu is Assistant Professor, Center for Statistical Science, Tsinghua University, Beijing, China (e-mail: zhiruihu33@gmail.com). Zheng Tracy Ke is Assistant Professor, Department of Statistics, Harvard University, Cambridge, Massachusetts 02138, USA (e-mail: zke@fas.harvard.edu). Jun S. Liu is Professor, Department of Statistics, Harvard University, Cambridge, Massachusetts 02138, USA (e-mail: jliu@stat.harvard.edu).

https://doi.org/10.1111/j.0006-341x.1999.00044.x

CARROLL, R. J. and STEFANSKI, L. A. (1994). Measurement error, instrumental variables and corrections for attenuation with applications to meta-analyses. *Stat. Med.* **13** 1265–1282.

CARROLL, R. J., RUPPERT, D., CRAINICEANU, C. M., TOSTESON, T. D. and KARAGAS, M. R. (2004). Nonlinear and nonparametric regression and instrumental variables. *J. Amer. Statist. Assoc.* **99** 736–750. MR2090907 https://doi.org/10.1198/016214504000001088

CARROLL, R. J., RUPPERT, D., STEFANSKI, L. A. and CRAINICEANU, C. M. (2006). *Measurement Error in Nonlinear Models: A Modern Perspective*, 2nd ed. *Monographs on Statistics and Applied Probability* **105**. CRC Press/CRC, Boca Raton, FL. MR2243417 https://doi.org/10.1201/9781420010138

CHEN, X. and WHITE, H. (1999). Improved rates and asymptotic normality for nonparametric neural network estimators. *IEEE Trans. Inf. Theory* **45** 682–691. MR1677026 https://doi.org/10.1109/18.749011

COOK, J. R. and STEFANSKI, L. A. (1994). Simulation-extrapolation estimation in parametric measurement error models. *J. Amer. Statist. Assoc.* **89** 1314–1328.

CRAINICEANU, C. M., RUPPERT, D. and WAND, M. P. (2005). Bayesian analysis for penalized spline regression using WinBUGS. *J. Stat. Softw.* **14** 1–24.

CREMER, C., MORRIS, Q. and DUVENAUD, D. (2017). Reinterpreting importance-weighted autoencoders. Preprint. Available at arXiv:1704.02916.

CRESSIE, N. and KORNAK, J. (2003). Spatial statistics in the presence of location error with an application to remote sensing of the environment. *Statist. Sci.* **18** 436–456. MR2059325 https://doi.org/10.1214/ss/1081443228

DELAIGLE, A., FAN, J. and CARROLL, R. J. (2009). A design-adaptive local polynomial estimator for the errors-in-variables problem. *J. Amer. Statist. Assoc.* **104** 348–359. MR2504382 https://doi.org/10.1198/jasa.2009.0114

DELAIGLE, A. and HALL, P. (2008). Using SIMEX for smoothing-parameter choice in errors-in-variables problems. *J. Amer. Statist. Assoc.* **103** 280–287. MR2394636 https://doi.org/10.1198/016214507000001355

DELLAPORTAS, P. and STEPHENS, D. A. (1995). Bayesian analysis of errors-in-variables regression models. *Biometrics* **51** 1085–1095.

DEMPSTER, A. P., LAIRD, N. M. and RUBIN, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *J. Roy. Statist. Soc. Ser. B* **39** 1–38. With discussion. MR0501537

DINH, L., KRUEGER, D. and BENGIO, Y. (2015). NICE: Non-linear independent components estimation. In *Workshop in International Conference on Learning Representations*.

EILERS, P. H. C. and MARX, B. D. (1996). Flexible smoothing with B-splines and penalties. *Statist. Sci.* **11** 89–121. With comments and a rejoinder by the authors. MR1435485 https://doi.org/10.1214/ss/1038425655

FAN, J., MA, C. and ZHONG, Y. (2021). A selective overview of deep learning. *Statist. Sci.* **36** 264–290. MR4255194 https://doi.org/10.1214/20-sts783

FAN, J. and TRUONG, Y. K. (1993). Nonparametric regression with errors in variables. *Ann. Statist.* **21** 1900–1925. MR1245773 https://doi.org/10.1214/aos/1176349402

GANGULI, B., STAUDENMAYER, J. and WAND, M. P. (2005). Additive models with predictors subject to measurement error. *Aust. N. Z. J. Stat.* **47** 193–202. MR2155119 https://doi.org/10.1111/j.1467-842X.2005.00383.x

GILKS, W. R. and RICHARDSON, S. (1992). Analysis of disease risks using ancillary risk factors, with application to job–exposure matrices. *Stat. Med.* **11** 1443–1463.

GOLOWICH, N., RAKHLIN, A. and SHAMIR, O. (2018). Size-independent sample complexity of neural networks. In *Conference on Learning Theory* 297–299.

GUNASEKAR, S., LEE, J. D., SREBRO, N. and SOUDRY, D. (2018). Implicit bias of gradient descent on linear convolutional networks. *Adv. Neural Inf. Process. Syst.* **2018** 9461–9471.

GUSTAFSON, P. (2004). *Measurement Error and Misclassification in Statistics and Epidemiology: Impacts and Bayesian Adjustments. Interdisciplinary Statistics*. CRC Press/CRC, Boca Raton, FL. MR2005104

GUSTAFSON, P., LE, N. D. and VALLÉE, M. (2002). A Bayesian approach to case–control studies with errors in covariables. *Biostatistics* **3** 229–243.

HALL, P. and QIU, P. (2005). Discrete-transform approach to deconvolution problems. *Biometrika* **92** 135–148. MR2158615 https://doi.org/10.1093/biomet/92.1.135

HARDT, M., RECHT, B. and SINGER, Y. (2016). Train faster, generalize better: Stability of stochastic gradient descent. In *International Conference on Machine Learning* 1225–1234. PMLR.

HU, Z., KE, Z. T. and LIU, J. S. (2022). Supplement to "Measurement Error Models: From Nonparametric Methods to Deep Neural Networks." https://doi.org/10.1214/21-STS834SUPP

JIANG, F., MA, Y. and CARROLL, R. J. (2018). A spline-assisted semiparametric approach to non-parametric measurement error models. Preprint. Available at arXiv:1804.00793.

JORDAN, M. I., GHAHRAMANI, Z., JAAKKOLA, T. S. and SAUL, L. K. (1999). An introduction to variational methods for graphical models. *Mach. Learn.* **37** 183–233.

KEMP, A. C., HORTON, B. P., DONNELLY, J. P., MANN, M. E., VERMEER, M. and RAHMSTORF, S. (2011). Climate related sea-level variations over the past two millennia. *Proc. Natl. Acad. Sci. USA* **108** 11017–11022.

KINGMA, D. P. and BA, J. (2014). Adam: A method for stochastic optimization. In *International Conference on Learning Representations*.

KINGMA, D. P. and WELLING, M. (2014). Auto-encoding variational Bayes. In *International Conference on Learning Representations*.

LE, T. A., IGL, M., RAINFORTH, T., JIN, T. and WOOD, F. (2018). Auto-encoding sequential Monte Carlo. In *International Conference on Learning Representations*.

LIN, H. W., TEGMARK, M. and ROLNICK, D. (2017). Why does deep and cheap learning work so well? *J. Stat. Phys.* **168** 1223–1247. MR3691248 https://doi.org/10.1007/s10955-017-1836-5

MAIOROV, V. E. and MEIR, R. (2000). On the near optimality of the stochastic approximation of smooth functions by neural networks. *Adv. Comput. Math.* **13** 79–103. MR1759189 https://doi.org/10.1023/A:1018993908478

MALLICK, B., HOFFMAN, F. O. and CARROLL, R. J. (2002). Semiparametric regression modeling with mixtures of Berkson and classical error, with application to fallout from the Nevada test site. *Biometrics* **58** 13–20. MR1891038 https://doi.org/10.1111/j.0006-341X.2002.00013.x

MEI, S., MONTANARI, A. and NGUYEN, P.-M. (2018). A mean field view of the landscape of two-layer neural networks. *Proc. Natl. Acad. Sci. USA* **115** E7665–E7671. MR3845070 https://doi.org/10.1073/pnas.1806579115

MHASKAR, H. N. (1996). Neural networks for optimal approximation of smooth and analytic functions. *Neural Comput.* **8** 164–177.

MÜLLER, P. and ROEDER, K. (1997). A Bayesian semiparametric model for case-control studies with errors in variables. *Biometrika* **84** 523–537. MR1603977 https://doi.org/10.1093/biomet/84.3.523

PAPAMAKARIOS, G., PAVLAKOU, T. and MURRAY, I. (2017). Masked autoregressive flow for density estimation. In *Advances in Neural Information Processing Systems* 2338–2347.

RAINFORTH, T., KOSIOREK, A., LE, T. A., MADDISON, C., IGL, M., WOOD, F. and TEH, Y. W. (2018). Tighter variational bounds are not necessarily better. In *International Conference on Machine Learning* 4277–4285.

RASMUSSEN, C. E. (2003). Gaussian processes in machine learning. In *Summer School on Machine Learning* 63–71. Springer, Berlin.

REZENDE, D. and MOHAMED, S. (2015). Variational inference with normalizing flows. In *International Conference on Machine Learning* 1530–1538.

REZENDE, D. J., MOHAMED, S. and WIERSTRA, D. (2014). Stochastic backpropagation and approximate inference in deep generative models. In *International Conference on Machine Learning* 1278–1286.

RICHARDSON, S. and GILKS, W. R. (1993). Conditional independence models for epidemiological studies with covariate measurement error. *Stat. Med.* **12** 1703–1722.

ROEDER, G., WU, Y. and DUVENAUD, D. K. (2017). Sticking the landing: Simple, lower-variance gradient estimators for variational inference. In *Advances in Neural Information Processing Systems* 6925–6934.

ROLNICK, D. and TEGMARK, M. (2018). The power of deeper networks for expressing natural functions. In *International Conference on Learning Representations*.

RUPPERT, D., WAND, M. P. and CARROLL, R. J. (2003). *Semiparametric Regression*. *Cambridge Series in Statistical and Probabilistic Mathematics* **12**. Cambridge Univ. Press, Cambridge. MR1998720 https://doi.org/10.1017/CBO9780511755453

SCHENNACH, S. M. (2007). Instrumental variable estimation of nonlinear errors-in-variables models. *Econometrica* **75** 201–239. MR2284741 https://doi.org/10.1111/j.1468-0262.2007.00736.x

SCHMIDT-HIEBER, J. (2020). Nonparametric regression using deep neural networks with ReLU activation function. *Ann. Statist.* **48** 1875–1897. MR4134774 https://doi.org/10.1214/19-AOS1875

SOUDRY, D., HOFFER, E., NACSON, M. S., GUNASEKAR, S. and SREBRO, N. (2018). The implicit bias of gradient descent on separable data. *J. Mach. Learn. Res.* **19** Paper No. 70. MR3899772

STAUDENMAYER, J. and RUPPERT, D. (2004). Local polynomial regression and simulation-extrapolation. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **66** 17–30. MR2035756 https://doi.org/10.1046/j.1369-7412.2003.05282.x

STEFANSKI, L. and CARROLL, R. J. (1990). Deconvoluting kernel density estimators. *Statistics* **21** 169–184. MR1054861 https://doi.org/10.1080/02331889008802238

TABAK, E. G. and TURNER, C. V. (2013). A family of nonparametric density estimation algorithms. *Comm. Pure Appl. Math.* **66** 145–164. MR2999294 https://doi.org/10.1002/cpa.21423

TAUPIN, M.-L. (2001). Semi-parametric estimation in the nonlinear structural errors-in-variables model. *Ann. Statist.* **29** 66–93. MR1833959 https://doi.org/10.1214/aos/996986502

TUCKER, G., LAWSON, D., GU, S. and MADDISON, C. J. (2018). Doubly reparameterized gradient estimators for Monte Carlo objectives. In *International Conference on Learning Representations*.

WILLIAMS, R. J. (1992). Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Mach. Learn.* **8** 229–256.

Statistical Science
2022, Vol. 37, No. 4, 494–518
https://doi.org/10.1214/21-STS835

High-Performance Statistical Computing in the Computing Environments of the 2020s

Seyoon Ko, Hua Zhou, Jin J. Zhou and Joong-Ho Won

Abstract. Technological advances in the past decade, hardware and software alike, have made access to high-performance computing (HPC) easier than ever. We review these advances from a statistical computing perspective. Cloud computing makes access to supercomputers affordable. Deep learning software libraries make programming statistical algorithms easy and enable users to write code once and run it anywhere—from a laptop to a workstation with multiple graphics processing units (GPUs) or a supercomputer in a cloud. Highlighting how these developments benefit statisticians, we review recent optimization algorithms that are useful for high-dimensional models and can harness the power of HPC. Code snippets are provided to demonstrate the ease of programming. We also provide an easy-to-use distributed matrix data structure suitable for HPC. Employing this data structure, we illustrate various statistical applications including large-scale positron emission tomography and ℓ_1-regularized Cox regression. Our examples easily scale up to an 8-GPU workstation and a 720-CPU-core cluster in a cloud. As a case in point, we analyze the onset of type-2 diabetes from the UK Biobank with 200,000 subjects and about 500,000 single nucleotide polymorphisms using the HPC ℓ_1-regularized Cox regression. Fitting this half-million-variate model takes less than 45 minutes and reconfirms known associations. To our knowledge, this is the first demonstration of the feasibility of penalized regression of survival outcomes at this scale.

Key words and phrases: High-performance statistical computing, graphics processing units (GPUs), cloud computing, deep learning, MM algorithms, ADMM, PDHG, Cox regression.

REFERENCES

ABADI, M., AGARWAL, A., BARHAM, P., BREVDO, E., CHEN, Z., CITRO, C., CORRADO, G. S., DAVIS, A., DEAN, J. et al. (2016). TensorFlow: Large-scale machine learning on heterogeneous systems. Preprint. Available at arXiv:1603.04467. Software available from: https://tensorflow.org.

ARMBRUST, M., FOX, A., GRIFFITH, R., JOSEPH, A. D., KATZ, R., KONWINSKI, A., LEE, G., PATTERSON, D., RABKIN, A. et al. (2010). A view of cloud computing. *Commun. ACM* **53** 50–58.

ATCHADÉ, Y. F., FORT, G. and MOULINES, E. (2017). On perturbed proximal gradient algorithms. *J. Mach. Learn. Res.* **18** Paper No. 10, 33. MR3634877

BAHRAMPOUR, S., RAMAKRISHNAN, N., SCHOTT, L. and SHAH, M. (2016). Comparative study of deep learning software frameworks. Preprint. Available at arXiv:1511.06435.

BALLARD, G., DEMMEL, J., HOLTZ, O. and SCHWARTZ, O. (2011). Minimizing communication in numerical linear algebra. *SIAM J. Matrix Anal. Appl.* **32** 866–901. MR2837583 https://doi.org/10.1137/090769156

BAUER, B. and KOHLER, M. (2019). On deep learning as a remedy for the curse of dimensionality in nonparametric regression. *Ann. Statist.* **47** 2261–2285. MR3953451 https://doi.org/10.1214/18-AOS1747

BAYDIN, A. G., PEARLMUTTER, B. A., RADUL, A. A. and SISKIND, J. M. (2017). Automatic differentiation in machine learning: A survey. *J. Mach. Learn. Res.* **18** Paper No. 153, 43. MR3800512

BECK, A. (2017). *First-Order Methods in Optimization. MOS-SIAM Series on Optimization* **25**. SIAM, Philadelphia, PA. MR3719240 https://doi.org/10.1137/1.9781611974997.ch1

Seyoon Ko is Postdoctoral Scholar, Department of Biostatistics, UCLA Fielding School of Public Health, Los Angeles, California 90095, USA. Hua Zhou is Professor, Department of Biostatistics, UCLA Fielding School of Public Health, Los Angeles, California 90095, USA. Jin J. Zhou is Associate Professor, Department of Medicine, UCLA David Geffen School of Medicine, Los Angeles, California 90095, USA, and Department of Epidemiology and Biostatistics, Mel and Enid Zuckerman College of Public Health, University of Arizona, Tucson, Arizona 85724, USA. Joong-Ho Won is Associate Professor, Department of Statistics, Seoul National University, Seoul, Korea (e-mail: wonj@stats.snu.ac.kr).

BECK, A. and TEBOULLE, M. (2009). A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM J. Imaging Sci*. **2** 183–202. MR2486527 https://doi.org/10.1137/080716542

BESARD, T., FOKET, C. and DE SUTTER, B. (2019). Effective extensible programming: Unleashing Julia on GPUs. *IEEE Trans. Parallel Distrib. Syst*. **30** 827–841.

BEZANSON, J., EDELMAN, A., KARPINSKI, S. and SHAH, V. B. (2017). Julia: A fresh approach to numerical computing. *SIAM Rev.* **59** 65–98. MR3605826 https://doi.org/10.1137/141000671

BLACKFORD, L. S. et al. (2002). An updated set of basic linear algebra subprograms (BLAS). *ACM Trans. Math. Software* **28** 135–151. MR1928065 https://doi.org/10.1145/567806.567807

BOYD, S., PARIKH, N., CHU, E., PELEATO, B. and ECKSTEIN, J. (2011). Distributed optimization and statistical learning via the alternating direction method of multipliers. *Found. Trends Mach. Learn*. **3** 1–122.

BUCKNER, J., WILSON, J., SELIGMAN, M., ATHEY, B., WATSON, S. and MENG, F. (2009). The gputools package enables GPU computing in R. *Bioinformatics* **26** 134–135.

CHAMBOLLE, A. and POCK, T. (2011). A first-order primal-dual algorithm for convex problems with applications to imaging. *J. Math. Imaging Vision* **40** 120–145. MR2782122 https://doi.org/10.1007/s10851-010-0251-1

CHAMBOLLE, A., EHRHARDT, M. J., RICHTÁRIK, P. and SCHÖNLIEB, C.-B. (2018). Stochastic primal-dual hybrid gradient algorithm with arbitrary sampling and imaging applications. *SIAM J. Optim*. **28** 2783–2808. MR3860126 https://doi.org/10.1137/17M1134834

CHEN, Y., LAN, G. and OUYANG, Y. (2014). Optimal primal-dual methods for a class of saddle point problems. *SIAM J. Optim*. **24** 1779–1814. MR3272627 https://doi.org/10.1137/130919362

CHEN, T., LI, M., LI, Y., LIN, M., WANG, N., WANG, M., XIAO, T., XU, B., ZHANG et al. (2015). MXNet: A flexible and efficient machine learning library for heterogeneous distributed systems. Preprint. Available at arXiv:1512.01274.

CHI, E. C., ZHOU, H. and LANGE, K. (2014). Distance majorization and its applications. *Math. Program*. **146** 409–436. MR3232621 https://doi.org/10.1007/s10107-013-0697-1

CHU, D., ZHANG, C., SUN, S. and TAO, Q. (2020). Semismooth Newton algorithm for efficient projections onto $\ell_{1,\infty}$-norm ball. In *ICML* 2020. *Proc. Mach. Learn. Res*. **119** 1974–1983.

CHURCH, D. M., SCHNEIDER, V. A., GRAVES, T., AUGER, K., CUNNINGHAM, F., BOUK, N., CHEN, H.-C., AGARWALA, R., MCLAREN, W. M. et al. (2011). Modernizing reference genome assemblies. *PLoS Biol*. **9** e1001091.

COLLOBERT, R., KAVUKCUOGLU, K. and FARABET, C. (2011). Torch7: A Matlab-like environment for machine learning. In *BigLearn*, *NeurIPS Workshop*.

COMBETTES, P. L. (2018). Monotone operator theory in convex optimization. *Math. Program*. **170** 177–206. MR3816562 https://doi.org/10.1007/s10107-018-1303-3

COMBETTES, P. L. and PESQUET, J.-C. (2011). Proximal splitting methods in signal processing. In *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*. *Springer Optim. Appl.* **49** 185–212. Springer, New York. MR2858838 https://doi.org/10.1007/978-1-4419-9569-8_10

CONDAT, L. (2013). A primal-dual splitting method for convex optimization involving Lipschitzian, proximable and linear composite terms. *J. Optim. Theory Appl*. **158** 460–479. MR3084386 https://doi.org/10.1007/s10957-012-0245-9

THE WELLCOME TRUST CASE CONTROL CONSORTIUM (2007). Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* **447** 661–678.

COOK, A. L., CHEN, W., THURBER, A. E., SMIT, D. J., SMITH, A. G., BLADEN, T. G., BROWN, D. L., DUFFY, D. L., PASTORINO, L. et al. (2009). Analysis of cultured human melanocytes based on polymorphisms within the SLC45A2/MATP, SLC24A5/NCKX5, and OCA2/P loci. *J. Invest. Dermatol*. **129** 392–405.

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

DEAN, J. and GHEMAWAT, S. (2008). MapReduce: Simplified data processing on large clusters. *Commun. ACM* **51** 107–113.

DEFAZIO, A., BACH, F. and LACOSTE-JULIEN, S. (2014). SAGA: A fast incremental gradient method with support for non-strongly convex composite objectives. In *NeurIPS* 2014. *Adv. Neural Inform. Process. Syst*. **27** 1646–1654.

DEMPSTER, A. P., LAIRD, N. M. and RUBIN, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *J. Roy. Statist. Soc. Ser. B* **39** 1–38. MR0501537

DONOHO, D. (2017). 50 years of data science. *J. Comput. Graph. Statist*. **26** 745–766. MR3765335 https://doi.org/10.1080/10618600.2017.1384734

DUCHI, J. C., JORDAN, M. I., WAINWRIGHT, M. J. and ZHANG, Y. (2014). Optimality guarantees for distributed statistical estimation. Preprint. Available at arXiv:1405.0782.

DUPUIS, J., LANGENBERG, C., PROKOPENKO, I., SAXENA, R., SORANZO, N., JACKSON, A. U., WHEELER, E., GLAZER, N. L., BOUATIA-NAJI, N. et al. (2010). New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. *Nat. Genet*. **42** 105–116.

EDDELBUETTEL, D. (2021). Parallel computing with R: A brief review. *Wiley Interdiscip. Rev.: Comput. Stat*. **13** Paper No. e1515, 13. MR4218942 https://doi.org/10.1002/wics.1515

EIJKHOUT, V. (2016). Introduction to High Performance Scientific Computing, 2nd ed., Lulu.com.

ESSER, E., ZHANG, X. and CHAN, T. F. (2010). A general framework for a class of first order primal-dual algorithms for convex optimization in imaging science. *SIAM J. Imaging Sci*. **3** 1015–1046. MR2763706 https://doi.org/10.1137/09076934X

EVANGELINOS, C. and HILL, C. N. (2008). Cloud computing for parallel scientific HPC applications: Feasibility of running coupled atmosphere-ocean climate models on Amazon's EC2. In *CCA* 2008 ACM, New York.

FAN, J., GUO, Y. and WANG, K. (2019). Communication-efficient accurate statistical estimation. *J. Am. Stat. Assoc.* https://doi.org/10.1080/01621459.2021.1969238

APACHE SOFTWARE FOUNDATION (2021). 'Apache Hadoop', https://hadoop.apache.org. Version 3.3.1. Accessed: 2021-07-03.

FOX, A. (2011). Cloud computing—What's in it for me as a scientist?. *Science* **331** 406–407.

GABAY, D. and MERCIER, B. (1976). A dual algorithm for the solution of nonlinear variational problems via finite element approximation. *Comput. Math. Appl*. **2** 17–40.

GABRIEL, E., FAGG, G. E., BOSILCA, G., ANGSKUN, T., DONGARRA, J. J., SQUYRES, J. M., SAHAY, V., KAMBADUR, P., BARRETT, B. et al. (2004). Open MPI: Goals, concept, and design of a next generation MPI implementation. In *'Proceedings of the* 11*th European PVM/MPI Users' Group Meeting'* 97–104, Budapest, Hungary.

GENTZSCH, W. (2001). Sun Grid Engine: Towards creating a compute power grid. In *CCGRID* 2001 35–36. IEEE Comput. Soc., Los Alamitos, CA.

GITTENS, A., DEVARAKONDA, A., RACAH, E., RINGENBURG, M., GERHARDT, L., KOTTALAM, J., LIU, J., MASCHHOFF, K., CANON, S. et al. (2016). Matrix factorizations at scale: A comparison of scientific data analytics in Spark and $C+MPI$ using three case studies. In 2016 *IEEE BigData* 204–213. IEEE, New York.

GOLUB, G. H. and VAN LOAN, C. F. (2013). *Matrix Computations*, 4th ed. *Johns Hopkins Studies in the Mathematical Sciences*. Johns Hopkins Univ. Press, Baltimore, MD. MR3024913

GOWER, R. M., LOIZOU, N., QIAN, X., SAILANBAYEV, A., SHULGIN, E. and RICHTÁRIK, P. (2019). SGD: General analysis and improved rates. In *ICML* 2019. *Proc. Mach. Learn. Res.* **97** 5200–5209.

GRIEWANK, A. and WALTHER, A. (2008). *Evaluating Derivatives: Principles and Techniques of Algorithmic Differentiation*, 2nd ed. SIAM, Philadelphia, PA. MR2454953 https://doi.org/10.1137/1.9780898717761

GU, Y., FAN, J., KONG, L., MA, S. and ZOU, H. (2018). ADMM for high-dimensional sparse penalized quantile regression. *Technometrics* **60** 319–331. MR3847169 https://doi.org/10.1080/00401706.2017.1345703

HAGER, G. and WELLEIN, G. (2010). *Introduction to High Performance Computing for Scientists and Engineers*. CRC Press, Boca Raton, FL.

HASTIE, T. J. and TIBSHIRANI, R. J. (1990). *Generalized Additive Models*. *Monographs on Statistics and Applied Probability* **43**. CRC Press, London. MR1082147

HUANG, D. W., SHERMAN, B. T. and LEMPICKI, R. A. (2009a). Bioinformatics enrichment tools: Paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* **37** 1–13.

HUANG, D. W., SHERMAN, B. T. and LEMPICKI, R. A. (2009b). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4** 44–57.

HUANG, J., JIAO, Y., LIU, Y. and LU, X. (2018). A constructive approach to L_0 penalized regression. *J. Mach. Learn. Res.* **19** Paper No. 10, 37. MR3862417

HUANG, J., JIAO, Y., JIN, B., LIU, J., LU, X. and YANG, C. (2021). A unified primal dual active set algorithm for nonconvex sparse recovery. *Statist. Sci.* **36** 215–238. MR4255190 https://doi.org/10.1214/19-sts758

HUNTER, D. R. and LANGE, K. (2004). A tutorial on MM algorithms. *Amer. Statist.* **58** 30–37. MR2055509 https://doi.org/10.1198/0003130042836

HUNTER, D. R. and LI, R. (2005). Variable selection using MM algorithms. *Ann. Statist.* **33** 1617–1642. MR2166557 https://doi.org/10.1214/009053605000000200

IMAIZUMI, M. and FUKUMIZU, K. (2019). Deep neural networks learn non-smooth functions effectively. In *AISTATS* 2019. *Proc. Mach. Learn. Res.* **89** 869–878.

INCUBATOR, F. (2021). Gloo: Collective communications library with various primitives for multi-machine training. https://github.com/facebookincubator/gloo. Accessed: 2021-07-03.

JANSSENS, B. (2021). MPIArrays.jl: Distributed arrays based on MPI one-sided communication. https://github.com/barche/MPIArrays.jl. Accessed: 2021-07-03.

JHA, S., QIU, J., LUCKOW, A., MANTHA, P. and FOX, G. C. (2014). A tale of two data-intensive paradigms: Applications, abstractions, and architectures. In 2014 *IEEE BigData* 645–652. IEEE, New York.

JIA, Y., SHELHAMER, E., DONAHUE, J., KARAYEV, S., LONG, J., GIRSHICK, R., GUADARRAMA, S. and DARRELL, T. (2014). Caffe: Convolutional architecture for fast feature embedding. In *MM* 2014 675–678. ACM, New York.

JOHNSON, R. and ZHANG, T. (2013). Accelerating stochastic gradient descent using predictive variance reduction. In *NeurIPS* 2013. *Adv. Neural Inform. Process. Syst.* **26** 315–323.

JORDAN, M. I., LEE, J. D. and YANG, Y. (2019). Communication-efficient distributed statistical inference. *J. Amer. Statist. Assoc.* **114** 668–681. MR3963171 https://doi.org/10.1080/01621459.2018.1429274

KEYS, K. L., ZHOU, H. and LANGE, K. (2019). Proximal distance algorithms: Theory and practice. *J. Mach. Learn. Res.* **20** Paper No. 66, 38. MR3960920

KIRK, D. (2007). NVIDIA CUDA software and GPU parallel computing architecture. In *ISMM* **7** 103–104.

KLÖCKNER, A., PINTO, N., LEE, Y., CATANZARO, B., IVANOV, P. and FASIH, A. (2012). PyCUDA and PyOpenCL: A scripting-based approach to GPU run-time code generation. *Parallel Comput.* **38** 157–174.

KO, S. (2020). Easily parallelizable statistical computing methods and their applications in modern high-performance computing environments. Ph.D. thesis, Seoul National Univ.

KO, S. and WON, J.-H. (2019). Optimal minimization of the sum of three convex functions with a linear operator. In *AISTATS* 2019. *Proc. Mach. Learn. Res.* **89** 1185–1194.

KO, S., YU, D. and WON, J.-H. (2019). Easily parallelizable and distributable class of algorithms for structured sparsity, with optimal acceleration. *J. Comput. Graph. Statist.* **28** 821–833. MR4045851 https://doi.org/10.1080/10618600.2019.1592757

KO S., ZHOU H., ZHOU J. J. and WON J.-H. (2022). Supplement to "High-Performance Statistical Computing in the Computing Environments of the 2020s." https://doi.org/10.1214/21-STS835SUPP

KOANANTAKOOL, P., AZAD, A., BULUÇ, A., MOROZOV, D., OH, S.-Y., OLIKER, L. and YELICK, L. (2016). Communication-avoiding parallel sparse-dense matrix-matrix multiplication. In 2016 *IEEE IPDPS* 842–853. IEEE, New York.

KOANANTAKOOL, P., ALI, A., AZAD, A., BULUC, A., MOROZOV, D., OLIKER, L., YELICK, K. and OH, S.-Y. (2018). Communication-avoiding optimization methods for distributed massive-scale sparse inverse covariance estimation. In *AISTATS* 2018. *Proc. Mach. Learn. Res.* **84** 1376–1386.

KUMMER, B. (1988). Newton's method for non-differentiable functions. In *Advances in Mathematical Optimization* (Guddat, J., Bank, B., Hollatz, H., Kall, P., Klatte, D., Kummer, B., Lommatzsch, K., Tammer, K., Vlach, M. et al., eds.). *Mathematical Research* **45** 114–125. Akademie-Verlag, Berlin. MR0953317

LAI, T. L. and YUAN, H. (2021). Stochastic approximation: From statistical origin to big-data, multidisciplinary applications. *Statist. Sci.* **36** 291–302. MR4255195 https://doi.org/10.1214/20-sts784

LAM, S. K., PITROU, A. and SEIBERT, S. (2015). Numba: A LLVM-based Python JIT compiler. In *LLVM* 2015 **7**, ACM, New York, 1–6.

LANGE, K. (2016). *MM Optimization Algorithms*. SIAM, Philadelphia, PA. MR3522165 https://doi.org/10.1137/1.9781611974409.ch1

LANGE, K., HUNTER, D. R. and YANG, I. (2000). Optimization transfer using surrogate objective functions. *J. Comput. Graph. Statist.* **9** 1–59. MR1819865 https://doi.org/10.2307/1390605

LECUN, Y., BENGIO, Y. and HINTON, G. (2015). Deep learning. *Nature* **521** 436–444.

LEE, D. D. and SEUNG, H. S. (1999). Learning the parts of objects by non-negative matrix factorization. *Nature* **401** 788–791.

LEE, D. D. and SEUNG, H. S. (2001). Algorithms for non-negative matrix factorization. In *NeurIPS* 2001. *Adv. Neural Inform. Process. Syst.* **14** 556–562.

LEE, J. D., LIU, Q., SUN, Y. and TAYLOR, J. E. (2017a). Communication-efficient sparse regression. *J. Mach. Learn. Res.* **18** Paper No. 5, 30. MR3625709

LEE, T., WON, J.-H., LIM, J. and YOON, S. (2017b). Large-scale structured sparsity via parallel fused lasso on multiple GPUs. *J. Comput. Graph. Statist.* **26** 851–864. MR3765349 https://doi.org/10.1080/10618600.2017.1328363

LI, X., SUN, D. and TOH, K.-C. (2018). A highly efficient semismooth Newton augmented Lagrangian method for solving lasso problems. *SIAM J. Optim.* **28** 433–458. MR3763098 https://doi.org/10.1137/16M1097572

LIM, H., DEWARAJA, Y. K. and FESSLER, J. A. (2018). A PET reconstruction formulation that enforces non-negativity in projection space for bias reduction in Y-90 imaging. *Phys. Med. Biol.* **63** 035042.

LIN, C.-J. (2007). Projected gradient methods for nonnegative matrix factorization. *Neural Comput.* **19** 2756–2779. MR2348161 https://doi.org/10.1162/neco.2007.19.10.2756

LIU, X., LI, Y., TANG, J. and YAN, M. (2020). A double residual compression algorithm for efficient distributed learning. In *AISTATS* 2020. *Proc. Mach. Learn. Res.* **108** 133–143.

MAHAJAN, A., TALIUN, D., THURNER, M., ROBERTSON, N. R., TORRES, J. M., RAYNER, N. W., PAYNE, A. J., STEINTHORSDOTTIR, V., SCOTT, R. A. et al. (2018). Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps. *Nat. Genet.* **50** 1505–1513.

MCLAREN, W., GIL, L., HUNT, S. E., RIAT, H. S., RITCHIE, G. R., THORMANN, A., FLICEK, P. and CUNNINGHAM, F. (2016). The Ensembl variant effect predictor. *Genome Biol.* **17** 122.

MITTAL, S., MADIGAN, D., BURD, R. S. and SUCHARD, M. A. (2014). High-dimensional, massive sample-size Cox proportional hazards regression for survival analysis. *Biostatistics* **15** 207–221.

MUNSHI, A. (2009). The OpenCL specification. In 2009 *IEEE HCS* 1–314. IEEE, New York.

NAKANO, J. (2012). Parallel computing techniques. In *Handbook of Computational Statistics—Concepts and Methods*. 1, 2. *Springer Handb. Comput. Stat.* 243–271. Springer, Heidelberg. MR2985401 https://doi.org/10.1007/978-3-642-21551-3_9

NEGAHBAN, S. N., RAVIKUMAR, P., WAINWRIGHT, M. J. and YU, B. (2012). A unified framework for high-dimensional analysis of M-estimators with decomposable regularizers. *Statist. Sci.* **27** 538–557. MR3025133 https://doi.org/10.1214/12-STS400

NERSC (2021). Distributed TensorFlow. https://docs.nersc.gov/machinelearning/tensorflow/#distributed-tensorflow. Accessed: 2021-07-03.

NG, M. C., SHRINER, D., CHEN, B. H., LI, J., CHEN, W.-M., GUO, X., LIU, J., BIELINSKI, S. J., YANEK, L. R. et al. (2014). Meta-analysis of genome-wide association studies in African Americans provides insights into the genetic architecture of type 2 diabetes. *PLoS Genet.* **10** e1004517.

NITANDA, A. (2014). Stochastic proximal gradient descent with acceleration techniques. In *NeurIPS* 2014. *Adv. Neural Inform. Process. Syst.* **27** 1574–1582.

NVIDIA (2021a). Basic linear algebra subroutines (cuBLAS) library. http://docs.nvidia.com/cuda/cublas. Accessed: 2021-07-03.

NVIDIA (2021b). Sparse matrix library (cuSPARSE). http://docs.nvidia.com/cuda/cusparse. Accessed: 2021-07-03.

O'HARA, R. B. and SILLANPÄÄ, M. J. (2009). A review of Bayesian variable selection methods: What, how and which. *Bayesian Anal.* **4** 85–117. MR2486240 https://doi.org/10.1214/09-BA403

OHN, I. and KIM, Y. (2019). Smooth function approximation by deep neural networks with general activation functions. *Entropy* **21** Paper No. 627, 21. MR3988437 https://doi.org/10.3390/e21070627

OWENS, J. D., LUEBKE, D., GOVINDARAJU, N., HARRIS, M., KRÜGER, J., LEFOHN, A. E. and PURCELL, T. J. (2007). A survey of general-purpose computation on graphics hardware. In *Computer Graphics Forum* **26** 80–113. Wiley, New York.

PARIKH, N. and BOYD, S. (2014). Proximal algorithms. *Found. Trends Optim.* **1** 127–239.

PASZKE, A., GROSS, S., MASSA, F., LERER, A., BRADBURY, J., CHANAN, G., KILLEEN, T., LIN, Z., GIMELSHEIN, N. et al. (2019). PyTorch: An imperative style, high-performance deep learning library. In *NeurIPS* 2019. *Adv. Neural Inform. Process. Syst.* **32** 8026–8037.

POLSON, N. G., SCOTT, J. G. and WILLARD, B. T. (2015). Proximal algorithms in statistics and machine learning. *Statist. Sci.* **30** 559–581. MR3432841 https://doi.org/10.1214/15-STS530

QI, L. Q. and SUN, J. (1993). A nonsmooth version of Newton's method. *Math. Program.* **58** 353–367. MR1216791 https://doi.org/10.1007/BF01581275

QIAN, X., QU, Z. and RICHTÁRIK, P. (2019). SAGA with arbitrary sampling. In *ICML* 2019. *Proc. Mach. Learn. Res.* **97** 5190–5199.

RAINA, R., MADHAVAN, A. and NG, A. Y. (2009). Large-scale deep unsupervised learning using graphics processors. In *ICML* 2009 873–880. ACM, New York.

RAMDAS, A. and TIBSHIRANI, R. J. (2016). Fast and flexible ADMM algorithms for trend filtering. *J. Comput. Graph. Statist.* **25** 839–858. MR3533641 https://doi.org/10.1080/10618600.2015.1054033

HYPERION RESEARCH (2019). HPC market update from ISC 2019, Technical report, Hyperion Research.

REYES, A. R. (2021). rTorch. https://f0nzie.github.io/rTorch/. Accessed: 2021-07-03.

REYES-ORTIZ, J. L., ONETO, L. and ANGUITA, D. (2015). Big data analytics in the cloud: Spark on Hadoop vs MPI/OpenMP on Beowulf. In *INNS Conference on Big Data* **8** 121.

RICHTÁRIK, P. and TAKÁČ, M. (2016a). On optimal probabilities in stochastic coordinate descent methods. *Optim. Lett.* **10** 1233–1243. MR3529843 https://doi.org/10.1007/s11590-015-0916-1

RICHTÁRIK, P. and TAKÁČ, M. (2016b). Parallel coordinate descent methods for big data optimization. *Math. Program.* **156** 433–484. MR3459207 https://doi.org/10.1007/s10107-015-0901-6

ROBBINS, H. and MONRO, S. (1951). A stochastic approximation method. *Ann. Math. Stat.* **22** 400–407. MR0042668 https://doi.org/10.1214/aoms/1177729586

ROLAND, CH., VARADHAN, R. and FRANGAKIS, C. E. (2007). Squared polynomial extrapolation methods with cycling: An application to the positron emission tomography problem. *Numer. Algorithms* **44** 159–172. MR2334694 https://doi.org/10.1007/s11075-007-9094-2

ROSASCO, L., VILLA, S. and VŨ, B. C. (2020). Convergence of stochastic proximal gradient algorithm. *Appl. Math. Optim.* **82** 891–917. MR4167686 https://doi.org/10.1007/s00245-019-09617-7

RSTUDIO (2021). 'R interface to TensorFlow', https://tensorflow.rstudio.com/. Version 2.5.0. Accessed: 2021-07-03.

RUDIN, L. I., OSHER, S. and FATEMI, E. (1992). Nonlinear total variation based noise removal algorithms. *Phys. D* **60** 259–268. MR3363401 https://doi.org/10.1016/0167-2789(92)90242-F

RUMELHART, D. E., HINTON, G. E. and WILLIAMS, R. J. (1986). Learning representations by back-propagating errors. *Nature* **323** 533–536.

RYU, E. K., KO, S. and WON, J.-H. (2020). Splitting with near-circulant linear systems: Applications to total variation CT and PET. *SIAM J. Sci. Comput.* **42** B185–B206. MR4059383 https://doi.org/10.1137/18M1224003

SCHMIDT-HIEBER, J. (2020). Nonparametric regression using deep neural networks with ReLU activation function. *Ann. Statist.* **48** 1875–1897. MR4134774 https://doi.org/10.1214/19-AOS1875

SCOTT, L. J., MOHLKE, K. L., BONNYCASTLE, L. L., WILLER, C. J., LI, Y., DUREN, W. L., ERDOS, M. R., STRINGHAM, H. M., CHINES, P. S. et al. (2007). A genome-wide association study of type 2 diabetes in finns detects multiple susceptibility variants. *Science* **316** 1341–1345.

SEIDE, F. and AGARWAL, A. (2016). CNTK: Microsoft's open-source deep-learning toolkit. In *SIGKDD* 2016 2135–2135. ACM, New York.

SERGEEV, A. and DEL BALSO, M. (2018). Horovod: Fast and easy distributed deep learning in tensorflow. Preprint. Available at arXiv:1802.05799.

SERVICES, A. W. (2021). AWS ParallelCluster. https://aws.amazon.com/ko/hpc/parallelcluster/. Version 2.11.0. Accessed: 2021-07-03.

STAPLES, G. (2006). Torque resource manager. In *SC* 2006 8. ACM, New York.

SUCHARD, M. A., WANG, Q., CHAN, C., FRELINGER, J., CRON, A. and WEST, M. (2010). Understanding GPU programming for statistical computation: Studies in massively parallel massive mixtures. *J. Comput. Graph. Statist.* **19** 419–438. MR2758309 https://doi.org/10.1198/jcgs.2010.10016

SUCHARD, M. A., SIMPSON, S. E., ZORYCH, I., RYAN, P. and MADIGAN, D. (2013). Massive parallelization of serial inference algorithms for a complex generalized linear model. *ACM Trans. Model. Comput. Simul.* **23** Art. 10, 17. MR3034220 https://doi.org/10.1145/2414416.2414791

SUDLOW, C., GALLACHER, J., ALLEN, N., BERAL, V., BURTON, P., DANESH, J., DOWNEY, P., ELLIOTT, P., GREEN, J. et al. (2015). UK Biobank: An open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med.* **12** e1001779.

SUZUKI, T. (2019). Adaptivity of deep ReLU network for learning in Besov and mixed smooth Besov spaces: Optimal rate and curse of dimensionality. In *ICLR* 2019.

TANG, H., YU, C., LIAN, X., ZHANG, T. and LIU, J. (2019). `DoubleSqueeze`: Parallel stochastic gradient descent with double-pass error-compensated compression. In *ICML* 2019. *Proc. Mach. Learn. Res.* **97** 6155–6165.

THEANO DEVELOPMENT TEAM (2016). Theano: A Python framework for fast computation of mathematical expressions. Preprint. Available at arXiv:1605.02688.

JULIAPARALLEL TEAM (2021). DistributedArrays.jl: Distributed arrays in Julia. https://github.com/JuliaParallel/DistributedArrays.jl. Accessed: 2021-07-03.

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

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

TIBSHIRANI, R. J. and TAYLOR, J. (2011). The solution path of the generalized lasso. *Ann. Statist.* **39** 1335–1371. MR2850205 https://doi.org/10.1214/11-AOS878

TIELEMAN, T. (2010). Gnumpy: An easy way to use GPU boards in Python Technical Report UTML TR 2010–002, Department of Computer Science, Univ. Toronto.

TSENG, P. and YUN, S. (2009). A coordinate gradient descent method for nonsmooth separable minimization. *Math. Program.* **117** 387–423. MR2421312 https://doi.org/10.1007/s10107-007-0170-0

. UNIVERSITY OF ZURICH (2021). ElastiCluster. https://elasticluster.readthedocs.io/en/latest/. Accessed: 2021-07-03.

USHEY, K., ALLAIRE, J. and TANG, Y. (2021). reticulate: Interface to 'Python'. https://cran.r-project.org/package=reticulate. Version 1.20. Accessed: 2021-07-03.

VAN ROSSUM, G. (1995). Python tutorial Technical Report CS-R9526, Centrum voor Wiskunde en Informatica (CWI), Amsterdam.

VAN DE GEIJN, R. A. and WATTS, J. (1997). SUMMA: Scalable universal matrix multiplication algorithm. *Concurrency: Practice and Experience* **9** 255–274.

VOIGHT, B. F., SCOTT, L. J., STEINTHORSDOTTIR, V., MORRIS, A. P., DINA, C., WELCH, R. P., ZEGGINI, E., HUTH, C., AULCHENKO, Y. S. et al. (2010). Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. *Nat. Genet.* **42** 579.

VŨ, B. C. (2013). A splitting algorithm for dual monotone inclusions involving cocoercive operators. *Adv. Comput. Math.* **38** 667–681. MR3037034 https://doi.org/10.1007/s10444-011-9254-8

WALKER, E. (2008). Benchmarking Amazon EC2 for higperformance scientific computing. *Login:: The Magazine of USENIX & SAGE* **33** 18–23.

WANG, E., ZHANG, Q., SHEN, B., ZHANG, G., LU, X., WU, Q. and WANG, Y. (2014). Intel Math Kernel library. In *High-Performance Computing on the Intel® Xeon Phi*™ 167–188. Springer, Berlin.

WANG, J., KOLAR, M., SREBRO, N. and ZHANG, T. (2017). Efficient distributed learning with sparsity. In *ICML* 2017. *Proc. Mach. Learn. Res.* **70** 3636–3645.

WON, J.-H. (2020). Proximity operator of the matrix perspective function and its applications. In *NeurIPS* 2020. *Adv. Neural Inform. Process. Syst.* **33**.

WRIGHT, S. J. (2015). Coordinate descent algorithms. *Math. Program.* **151** 3–34. MR3347548 https://doi.org/10.1007/s10107-015-0892-3

WU, T. T. and LANGE, K. (2010). The MM alternative to EM. *Statist. Sci.* **25** 492–505. MR2807766 https://doi.org/10.1214/08-STS264

XIAO, L. and ZHANG, T. (2014). A proximal stochastic gradient method with progressive variance reduction. *SIAM J. Optim.* **24** 2057–2075. MR3285905 https://doi.org/10.1137/140961791

XUE, L., MA, S. and ZOU, H. (2012). Positive-definite ℓ_1-penalized estimation of large covariance matrices. *J. Amer. Statist. Assoc.* **107** 1480–1491. MR3036409 https://doi.org/10.1080/01621459.2012.725386

YOO, A. B., JETTE, M. A. and GRONDONA, M. (2003). Slurm: Simple Linux utility for resource management. In *JSSPP* 2003 44–60. Springer, Berlin.

YU, D., WON, J.-H., LEE, T., LIM, J. and YOON, S. (2015). High-dimensional fused lasso regression using majorization-minimization and parallel processing. *J. Comput. Graph. Statist.* **24** 121–153. MR3328251 https://doi.org/10.1080/10618600.2013.878662

ZAHARIA, M., CHOWDHURY, M., FRANKLIN, M. J., SHENKER, S., STOICA, I. et al. (2010). Spark: Cluster computing with working sets. *HotCloud* **10** 95.

ZHANG, Y., DUCHI, J. C. and WAINWRIGHT, M. J. (2013). Communication-efficient algorithms for statistical optimization. *J. Mach. Learn. Res.* **14** 3321–3363. MR3144464

ZHANG, X., WANG, Q. and CHOTHIA, Z. (2021). OpenBLAS: An optimized BLAS library. https://www.openblas.net/. Accessed: 2021-07-03.

ZHOU, H., LANGE, K. and SUCHARD, M. A. (2010). Graphics processing units and high-dimensional optimization. *Statist. Sci.* **25** 311–324. MR2791670 https://doi.org/10.1214/10-STS336

ZHU, Y. (2017). An augmented ADMM algorithm with application to the generalized lasso problem. *J. Comput. Graph. Statist.* **26** 195–204. MR3610420 https://doi.org/10.1080/10618600.2015.1114491

ZHU, M. and CHAN, T. (2008). An efficient primal-dual hybrid gradient algorithm for total variation image restoration Technical Report 08-34, UCLA CAM.

Statistical Science
2022, Vol. 37, No. 4, 519–540
https://doi.org/10.1214/21-STS838

The SPDE Approach to Matérn Fields: Graph Representations

Daniel Sanz-Alonso and Ruiyi Yang

Abstract. This paper investigates Gaussian Markov random field approximations to nonstationary Gaussian fields using graph representations of stochastic partial differential equations. We establish approximation error guarantees building on the theory of spectral convergence of graph Laplacians. The proposed graph representations provide a generalization of the Matérn model to unstructured point clouds, and facilitate inference and sampling using linear algebra methods for sparse matrices. In addition, they bridge and unify several models in Bayesian inverse problems, spatial statistics and graph-based machine learning. We demonstrate through examples in these three disciplines that the unity revealed by graph representations facilitates the exchange of ideas across them.

Key words and phrases: Gauss Matérn fields, Gaussian Markov random fields, graph Laplacians, latent Gaussian models, stochastic partial differential equations.

REFERENCES

[1] ABRAHAMSEN, P. (1997). A review of Gaussian random fields and correlation functions. Norsk Regnesentral/Norwegian Computing Center Oslo.

[2] ADLER, R. J. (2010). *The Geometry of Random Fields*. *Classics in Applied Mathematics* **62**. SIAM, Philadelphia, PA. MR3396215 https://doi.org/10.1137/1.9780898718980.ch1

[3] ANDERES, E. B. and STEIN, M. L. (2008). Estimating deformations of isotropic Gaussian random fields on the plane. *Ann. Statist.* **36** 719–741. MR2396813 https://doi.org/10.1214/009053607000000893

[4] BAKKA, H., RUE, H., FUGLSTAD, G.-A., RIEBLER, A., BOLIN, D., ILLIAN, J., KRAINSKI, E., SIMPSON, D. and LINDGREN, F. (2018). Spatial modeling with R-INLA: A review. *Wiley Interdiscip. Rev.: Comput. Stat.* **10** e1443. MR3873676 https://doi.org/10.1002/wics.1443

[5] BANERJEE, S., CARLIN, B. P. and GELFAND, A. E. (2015). *Hierarchical Modeling and Analysis for Spatial Data*, 2nd ed. *Monographs on Statistics and Applied Probability* **135**. CRC Press, Boca Raton, FL. MR3362184

[6] BANERJEE, S., GELFAND, A. E., FINLEY, A. O. and SANG, H. (2008). Gaussian predictive process models for large spatial data sets. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **70** 825–848. MR2523906 https://doi.org/10.1111/j.1467-9868.2008.00663.x

[7] BARDEEN, J. M., SZALAY, A. S., KAISER, N. and BOND, J. R. (1985). The statistics of peaks of Gaussian random fields. *Astrophys. J.* **304** 15–61.

[8] BARDSLEY, J. M. (2013). Gaussian Markov random field priors for inverse problems. *Inverse Probl. Imaging* **7** 397–416. MR3063540 https://doi.org/10.3934/ipi.2013.7.397

[9] BELKIN, M., MATVEEVA, I. and NIYOGI, P. (2004). Regularization and semi-supervised learning on large graphs. In *Learning Theory*. *Lecture Notes in Computer Science* **3120** 624–638. Springer, Berlin. MR2177939 https://doi.org/10.1007/978-3-540-27819-1_43

[10] BELKIN, M. and NIYOGI, P. (2001). Semi-supervised learning on Riemannian manifolds. *Mach. Learn.* **56** 209–239.

[11] BELKIN, M. and NIYOGI, P. (2005). Towards a theoretical foundation for Laplacian-based manifold methods. In *Learning Theory*. *Lecture Notes in Computer Science* **3559** 486–500. Springer, Berlin. MR2203282 https://doi.org/10.1007/11503415_33

[12] BELKIN, M. and NIYOGI, P. (2007). Convergence of Laplacian eigenmaps. *Adv. Neural Inf. Process. Syst.* **19** 129.

[13] BERTOZZI, A. L., LUO, X., STUART, A. M. and ZYGALAKIS, K. C. (2018). Uncertainty quantification in graph-based classification of high dimensional data. *SIAM/ASA J. Uncertain. Quantificat.* **6** 568–595. MR3794350 https://doi.org/10.1137/17M1134214

[14] BESAG, J. (1974). Spatial interaction and the statistical analysis of lattice systems. *J. Roy. Statist. Soc. Ser. B* **36** 192–236. MR0373208

[15] BESAG, J. (1975). Statistical analysis of non-lattice data. *J. R. Stat. Soc., Ser. D, Stat.* **24** 179–195.

[16] BEVILACQUA, M., FAOUZI, T., FURRER, R. and PORCU, E. (2019). Estimation and prediction using generalized Wendland covariance functions under fixed domain asymptotics.

Daniel Sanz-Alonso is Assistant Professor, Department of Statistics and Committee on Computational and Applied Mathematics, University of Chicago, Chicago, Illinois 60637, USA (e-mail: sanzalonso@uchicago.edu). Ruiyi Yang is Ph.D. Candidate, Committee on Computational and Applied Mathematics, University of Chicago, Chicago, Illinois 60637, USA (e-mail: yry@uchicago.edu).

Ann. Statist. **47** 828–856. MR3909952 https://doi.org/10.1214/17-AOS1652

[17] BIGONI, D., CHEN, Y., GARCIA TRILLOS, N., MARZOUK, Y. and SANZ-ALONSO, D. (2020). Data-driven forward discretizations for Bayesian inversion. *Inverse Probl.* **36** 105008. MR4162003 https://doi.org/10.1088/1361-6420/abb2fa

[18] BOGACHEV, V. I. (1998). *Gaussian Measures*. *Mathematical Surveys and Monographs* **62**. Amer. Math. Soc., Providence, RI. MR1642391 https://doi.org/10.1090/surv/062

[19] BOLIN, D. (2014). Spatial Matérn fields driven by non-Gaussian noise. *Scand. J. Stat.* **41** 557–579. MR3249417 https://doi.org/10.1111/sjos.12046

[20] BOLIN, D. and KIRCHNER, K. (2020). The rational SPDE approach for Gaussian random fields with general smoothness. *J. Comput. Graph. Statist.* **29** 274–285. MR4116041 https://doi.org/10.1080/10618600.2019.1665537

[21] BOLIN, D., KIRCHNER, K. and KOVÁCS, M. (2018). Weak convergence of Galerkin approximations for fractional elliptic stochastic PDEs with spatial white noise. *BIT* **58** 881–906. MR3882975 https://doi.org/10.1007/s10543-018-0719-8

[22] BOLIN, D., KIRCHNER, K. and KOVÁCS, M. (2020). Numerical solution of fractional elliptic stochastic PDEs with spatial white noise. *IMA J. Numer. Anal.* **40** 1051–1073. MR4092278 https://doi.org/10.1093/imanum/dry091

[23] BOLIN, D. and LINDGREN, F. (2011). Spatial models generated by nested stochastic partial differential equations, with an application to global ozone mapping. *Ann. Appl. Stat.* **5** 523–550. MR2810408 https://doi.org/10.1214/10-AOAS383

[24] BROCHU, E., CORA, V. M. and DE FREITAS, N. (2010). A tutorial on Bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning. Preprint. Available at arXiv:1012.2599.

[25] BUI-THANH, T., GHATTAS, O., MARTIN, J. and STADLER, G. (2013). A computational framework for infinite-dimensional Bayesian inverse problems Part I: The linearized case, with application to global seismic inversion. *SIAM J. Sci. Comput.* **35** A2494–A2523. MR3126997 https://doi.org/10.1137/12089586X

[26] BURAGO, D., IVANOV, S. and KURYLEV, Y. (2014). A graph discretization of the Laplace–Beltrami operator. *J. Spectr. Theory* **4** 675–714. MR3299811 https://doi.org/10.4171/JST/83

[27] CALDER, J. and GARCIA TRILLOS, N. (2019). Improved spectral convergence rates for graph Laplacians on epsilon-graphs and k-NN graphs. Preprint. Available at arXiv:1910.13476.

[28] CALVETTI, D. and SOMERSALO, E. (2007). *Introduction to Bayesian Scientific Computing*: *Ten Lectures on Subjective Computing*. *Surveys and Tutorials in the Applied Mathematical Sciences* **2**. Springer, New York. MR2351679

[29] CAMELETTI, M., LINDGREN, F., SIMPSON, D. and RUE, H. (2013). Spatio-temporal modeling of particulate matter concentration through the SPDE approach. *AStA Adv. Stat. Anal.* **97** 109–131. MR3045763 https://doi.org/10.1007/s10182-012-0196-3

[30] CANZANI, Y. (2013). Analysis on manifolds via the Laplacian. Lecture Notes. Available at http://www.math.harvard.edu/canzani/docs/Laplacian.pdf.

[31] CHUNG, F. R. K. (1997). *Spectral Graph Theory*. *CBMS Regional Conference Series in Mathematics* **92**. Amer. Math. Soc., Providence, RI. MR1421568

[32] COHEN, F. S., FAN, Z. and PATEL, M. A. (1991). Classification of rotated and scaled textured images using Gaussian Markov random field models. *IEEE Trans. Pattern Anal. Mach. Intell.* **2** 192–202.

[33] DUNLOP, M. M., IGLESIAS, M. A. and STUART, A. M. (2017). Hierarchical Bayesian level set inversion. *Stat. Comput.* **27** 1555–1584. MR3687326 https://doi.org/10.1007/s11222-016-9704-8

[34] DUNLOP, M. M. and STUART, A. M. (2016). The Bayesian formulation of EIT: Analysis and algorithms. *Inverse Probl. Imaging* **10** 1007–1036. MR3610749 https://doi.org/10.3934/ipi.2016030

[35] FRAZIER, P. I. (2018). A tutorial on Bayesian optimization. *Stat* **1050**.

[36] FUGLSTAD, G.-A., LINDGREN, F., SIMPSON, D. and RUE, H. (2015). Exploring a new class of non-stationary spatial Gaussian random fields with varying local anisotropy. *Statist. Sinica* **25** 115–133. MR3328806

[37] FUGLSTAD, G.-A., SIMPSON, D., LINDGREN, F. and RUE, H. (2015). Does non-stationary spatial data always require non-stationary random fields? *Spat. Stat.* **14** 505–531. MR3431054 https://doi.org/10.1016/j.spasta.2015.10.001

[38] FUGLSTAD, G.-A., SIMPSON, D., LINDGREN, F. and RUE, H. (2019). Constructing priors that penalize the complexity of Gaussian random fields. *J. Amer. Statist. Assoc.* **114** 445–452. MR3941267 https://doi.org/10.1080/01621459.2017.1415907

[39] FURRER, R., GENTON, M. G. and NYCHKA, D. (2006). Covariance tapering for interpolation of large spatial datasets. *J. Comput. Graph. Statist.* **15** 502–523. MR2291261 https://doi.org/10.1198/106186006X132178

[40] GARCÍA TRILLOS, N., GERLACH, M., HEIN, M. and SLEPČEV, D. (2020). Error estimates for spectral convergence of the graph Laplacian on random geometric graphs toward the Laplace–Beltrami operator. *Found. Comput. Math.* **20** 827–887. MR4130541 https://doi.org/10.1007/s10208-019-09436-w

[41] GARCÍA TRILLOS, N., KAPLAN, Z., SAMAKHOANA, T. and SANZ-ALONSO, D. (2020). On the consistency of graph-based Bayesian semi-supervised learning and the scalability of sampling algorithms. *J. Mach. Learn. Res.* **21** Paper No. 28. MR4073761

[42] GARCÍA TRILLOS, N. and SANZ-ALONSO, D. (2017). The Bayesian formulation and well-posedness of fractional elliptic inverse problems. *Inverse Probl.* **33** 065006. MR3656667 https://doi.org/10.1088/1361-6420/aa711e

[43] GARCÍA TRILLOS, N. and SANZ-ALONSO, D. (2018). Continuum limits of posteriors in graph Bayesian inverse problems. *SIAM J. Math. Anal.* **50** 4020–4040. MR3829512 https://doi.org/10.1137/17M1138005

[44] GARCÍA TRILLOS, N., SANZ-ALONSO, D. and YANG, R. (2019). Local regularization of noisy point clouds: Improved global geometric estimates and data analysis. *J. Mach. Learn. Res.* **20** Paper No. 136. MR4002890

[45] GELFAND, A. E., DIGGLE, P., GUTTORP, P. and FUENTES, M. (2010). *Handbook of Spatial Statistics*. CRC Press, Boca Raton, FL.

[46] GINÉ, E. and KOLTCHINSKII, V. (2006). Empirical graph Laplacian approximation of Laplace–Beltrami operators: Large sample results. In *High Dimensional Probability*. *Institute of Mathematical Statistics Lecture Notes—Monograph Series* **51** 238–259. IMS, Beachwood, OH. MR2387773 https://doi.org/10.1214/074921706000000888

[47] GNEITING, T., RAFTERY, A. E., WESTVELD III, A. H. and GOLDMAN, T. (2005). Calibrated probabilistic forecasting using ensemble model output statistics and minimum CRPS estimation. *Mon. Weather Rev.* **133** 1098–1118.

[48] 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

[49] GUTTORP, P. and GNEITING, T. (2006). Studies in the history of probability and statistics. XLIX. On the Matérn correlation family. *Biometrika* **93** 989–995. MR2285084 https://doi.org/10.1093/biomet/93.4.989

[50] HARIZANOV, S., LAZAROV, R., MARGENOV, S., MARINOV, P. and VUTOV, Y. (2018). Optimal solvers for linear systems with fractional powers of sparse SPD matrices. *Numer. Linear Algebra Appl*. **25** e2167. MR3859150 https://doi.org/10.1002/nla.2167

[51] HARLIM, J., SANZ-ALONSO, D. and YANG, R. (2020). Kernel methods for Bayesian elliptic inverse problems on manifolds. *SIAM/ASA J. Uncertain. Quantificat*. **8** 1414–1445. MR4178385 https://doi.org/10.1137/19M1295222

[52] HEATON, M. J., DATTA, A., FINLEY, A. O. et al. (2019). A case study competition among methods for analyzing large spatial data. *J. Agric. Biol. Environ. Stat*. **24** 398–425. MR3996451 https://doi.org/10.1007/s13253-018-00348-w

[53] HEIN, M. (2006). Uniform convergence of adaptive graph-based regularization. In *Proc. of the* 19*th Annual Conference on Learning Theory* (*COLT*). (G. Lugosi and H. U. Simon, eds.) 50–64. Springer, Berlin.

[54] HEIN, M., AUDIBERT, J.-Y. and VON LUXBURG, U. (2005). From graphs to manifolds—weak and strong pointwise consistency of graph Laplacians. In *Learning Theory*. *Lecture Notes in Computer Science* **3559** 470–485. Springer, Berlin. MR2203281 https://doi.org/10.1007/11503415_32

[55] HENNIG, P., OSBORNE, M. A. and GIROLAMI, M. (2015). Probabilistic numerics and uncertainty in computations. *Proc. A*. **471** 20150142. MR3378744 https://doi.org/10.1098/rspa.2015.0142

[56] ISAAC, T., PETRA, N., STADLER, G. and GHATTAS, O. (2015). Scalable and efficient algorithms for the propagation of uncertainty from data through inference to prediction for large-scale problems, with application to flow of the Antarctic ice sheet. *J. Comput. Phys*. **296** 348–368. MR3348506 https://doi.org/10.1016/j.jcp.2015.04.047

[57] KAIPIO, J. and SOMERSALO, E. (2005). *Statistical and Computational Inverse Problems*. *Applied Mathematical Sciences* **160**. Springer, New York. MR2102218

[58] 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

[59] KERSTING, H. and HENNIG, P. (2016). Active uncertainty calibration in Bayesian ode solvers. In *Proceedings of the Thirty-Second Conference on Uncertainty in Artificial Intelligence* 309–318.

[60] KHRISTENKO, U., SCARABOSIO, L., SWIERCZYNSKI, P., ULLMANN, E. and WOHLMUTH, B. (2019). Analysis of boundary effects on PDE-based sampling of Whittle–Matérn random fields. *SIAM/ASA J. Uncertain. Quantificat*. **7** 948–974. MR3984316 https://doi.org/10.1137/18M1215700

[61] 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

[62] KONDOR, R. I. and LAFFERTY, J. (2002). Diffusion kernels on graphs and other discrete structures. In *Proceedings of the* 19*th International Conference on Machine Learning* **2002** 315–322.

[63] LI, Y., MARK, B., RASKUTTI, G. and WILLETT, R. (2018). Graph-based regularization for regression problems with highly-correlated designs. In 2018 *IEEE Global Conference on Signal and Information Processing* (*GlobalSIP*) 740–742. IEEE, New York.

[64] LINDGREN, F., RUE, H. and LINDSTRÖM, J. (2011). An explicit link between Gaussian fields and Gaussian Markov random fields: The stochastic partial differential equation approach. *J. R. Stat. Soc. Ser. B. Stat. Methodol*. **73** 423–498. MR2853727 https://doi.org/10.1111/j.1467-9868.2011.00777.x

[65] LISCHKE, A., PANG, G., GULIAN, M. et al. (2020). What is the fractional Laplacian? A comparative review with new results. *J. Comput. Phys*. **404** 109009. MR4043885 https://doi.org/10.1016/j.jcp.2019.109009

[66] LIU, F., CHAKRABORTY, S., LI, F., LIU, Y. and LOZANO, A. C. (2014). Bayesian regularization via graph Laplacian. *Bayesian Anal*. **9** 449–474. MR3217003 https://doi.org/10.1214/14-BA860

[67] MACKAY, D. J. C. Gaussian processes-a replacement for supervised neural networks? NIPS tutorial.

[68] MARTIN, J. D. and SIMPSON, T. W. (2005). Use of kriging models to approximate deterministic computer models. *AIAA J*. **43** 853–863.

[69] MATÉRN, B. (2013). *Spatial Variation*, 2nd ed. *Lecture Notes in Statistics* **36**. Springer, Berlin.

[70] MONTAGNA, S. and TOKDAR, S. T. (2016). Computer emulation with nonstationary Gaussian processes. *SIAM/ASA J. Uncertain. Quantificat*. **4** 26–47. MR3659684 https://doi.org/10.1137/141001512

[71] MONTERRUBIO-GÓMEZ, K., ROININEN, L., WADE, S., DAMOULAS, T. and GIROLAMI, M. (2020). Posterior inference for sparse hierarchical non-stationary models. *Comput. Statist. Data Anal*. **148** 106954. MR4082506 https://doi.org/10.1016/j.csda.2020.106954

[72] NG, Y. C., COLOMBO, N. and SILVA, R. (2018). Bayesian semi-supervised learning with graph Gaussian processes. In *Advances in Neural Information Processing Systems* 1683–1694.

[73] PEYRÉ, G., CUTURI, M. et al. (2019). Computational optimal transport: With applications to data science. *Found. Trends Mach. Learn*. **11** 355–607.

[74] RAISSI, M., PERDIKARIS, P. and KARNIADAKIS, G. E. (2017). Machine learning of linear differential equations using Gaussian processes. *J. Comput. Phys*. **348** 683–693. MR3689653 https://doi.org/10.1016/j.jcp.2017.07.050

[75] RAISSI, M., PERDIKARIS, P. and KARNIADAKIS, G. E. (2018). Numerical Gaussian processes for time-dependent and nonlinear partial differential equations. *SIAM J. Sci. Comput*. **40** A172–A198. MR3747469 https://doi.org/10.1137/17M1120762

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

[77] ROININEN, L., GIROLAMI, M., LASANEN, S. and MARKKANEN, M. (2019). Hyperpriors for Matérn fields with applications in Bayesian inversion. *Inverse Probl. Imaging* **13** 1–29. MR3917848 https://doi.org/10.3934/ipi.2019001

[78] ROININEN, L., HUTTUNEN, J. M. J. and LASANEN, S. (2014). Whittle–Matérn priors for Bayesian statistical inversion with applications in electrical impedance tomography. *Inverse Probl. Imaging* **8** 561–586. MR3209311 https://doi.org/10.3934/ipi.2014.8.561

[79] RUE, H. and HELD, L. (2005). *Gaussian Markov Random Fields*: *Theory and Applications*. *Monographs on Statistics and Applied Probability* **104**. CRC Press/CRC, Boca Raton, FL. MR2130347 https://doi.org/10.1201/9780203492024

[80] RUE, H., MARTINO, S. and CHOPIN, N. (2009). Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *J. R. Stat. Soc. Ser.*

B. Stat. Methodol. **71** 319–392. MR2649602 https://doi.org/10.1111/j.1467-9868.2008.00700.x
[81] SAMPSON, P. D., DAMIAN, D. and GUTTORP, P. (2001). Advances in modeling and inference for environmental processes with nonstationary spatial covariance. In *GeoENV III—Geostatistics for Environmental Applications* 17–32. Springer, Berlin.
[82] SANCHEZ-VILA, X., GUADAGNINI, A. and CARRERA, J. (2006). Representative hydraulic conductivities in saturated groundwater flow. *Reviews of Geophysics* **44**.
[83] SANZ-ALONSO, D., STUART, A. M. and TAEB, A. (2019). Inverse problems and data assimilation. Preprint. Available at arXiv:1810.06191.
[84] SANZ-ALONSO, D. and YANG, R. (2022). Supplement to "The SPDE Approach to Matérn Fields: Graph Representations." https://doi.org/10.1214/21-STS838SUPPA, https://doi.org/10.1214/21-STS838SUPPB
[85] SEEGER, M. (2000). Relationships between Gaussian processes, support vector machines and smoothing splines. *Mach. Learn.*.
[86] SIMPSON, D., LINDGREN, F. and RUE, H. (2012). Think continuous: Markovian Gaussian models in spatial statistics. *Spat. Stat.* **1** 16–29.
[87] SINGER, A. (2006). From graph to manifold Laplacian: The convergence rate. *Appl. Comput. Harmon. Anal.* **21** 128–134. MR2238670 https://doi.org/10.1016/j.acha.2006.03.004
[88] SINGER, A. and WU, H.-T. (2017). Spectral convergence of the connection Laplacian from random samples. *Inf. Inference* **6** 58–123. MR3636868 https://doi.org/10.1093/imaiai/iaw016
[89] SOLLICH, P. (2002). Bayesian methods for support vector machines: Evidence and predictive class probabilities. *Mach. Learn.* **46** 21–52.
[90] SOMERSALO, E., CHENEY, M. and ISAACSON, D. (1992). Existence and uniqueness for electrode models for electric current computed tomography. *SIAM J. Appl. Math.* **52** 1023–1040. MR1174044 https://doi.org/10.1137/0152060
[91] STATHOPOULOS, V., ZAMORA-GUTIERREZ, V., JONES, K. and GIROLAMI, M. (2014). Bat call identification with Gaussian process multinomial probit regression and a dynamic time warping kernel. In *Artificial Intelligence and Statistics* 913–921.
[92] STEIN, M. L. (1999). *Interpolation of Spatial Data: Some Theory for Kriging. Springer Series in Statistics*. Springer, New York. MR1697409 https://doi.org/10.1007/978-1-4612-1494-6
[93] STUART, A. M. (2010). Inverse problems: A Bayesian perspective. *Acta Numer.* **19** 451–559. MR2652785 https://doi.org/10.1017/S0962492910000061
[94] STUART, A. M. and TECKENTRUP, A. L. (2018). Posterior consistency for Gaussian process approximations of Bayesian posterior distributions. *Math. Comp.* **87** 721–753. MR3739215 https://doi.org/10.1090/mcom/3244
[95] 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
[96] TAO, W. and SHI, Z. (2020). Convergence of Laplacian spectra from random samples. *J. Comput. Math.* **38** 952–984.
[97] TAYLOR, J. E. and WORSLEY, K. J. (2007). Detecting sparse signals in random fields, with an application to brain mapping. *J. Amer. Statist. Assoc.* **102** 913–928. MR2354405 https://doi.org/10.1198/016214507000000815
[98] TING, D., HUANG, L. and JORDAN, M. I. (2010). An analysis of the convergence of graph Laplacians. In *Proc. of the* 27*th Int. Conference on Machine Learning* (*ICML*).
[99] VAN DER VAART, A. W. and VAN ZANTEN, J. H. (2008). Rates of contraction of posterior distributions based on Gaussian process priors. *Ann. Statist.* **36** 1435–1463. MR2418663 https://doi.org/10.1214/009053607000000613
[100] VON LUXBURG, U. (2007). A tutorial on spectral clustering. *Stat. Comput.* **17** 395–416. MR2409803 https://doi.org/10.1007/s11222-007-9033-z
[101] WHITTLE, P. (1954). On stationary processes in the plane. *Biometrika* **41** 434–449. MR0067450 https://doi.org/10.1093/biomet/41.3-4.434
[102] WHITTLE, P. (1963). Stochastic processes in several dimensions. *Bull. Int. Stat. Inst.* **40** 974–994. MR0173287
[103] WIENS, A., NYCHKA, D. and KLEIBER, W. (2020). Modeling spatial data using local likelihood estimation and a Matérn to spatial autoregressive translation. *Environmetrics* **31** e2652. MR4151871 https://doi.org/10.1002/env.2652
[104] WILLIAMS, C. K. I. and RASMUSSEN, C. E. (1996). Gaussian processes for regression. In *Advances in Neural Information Processing Systems* 514–520.
[105] ZELNIK-MANOR, L. and PERONA, P. (2005). Self-tuning spectral clustering. In *Advances in Neural Information Processing Systems* 1601–1608.
[106] ZHU, X., LAFFERTY, J. and GHAHRAMANI, Z. (2003). Semi-supervised learning: From Gaussian fields to Gaussian processes. In *School of CS, CMU* Citeseer.

Statistical Science
2022, Vol. 37, No. 4, 541–561
https://doi.org/10.1214/21-STS839

The Covariate-Adjusted ROC Curve: The Concept and Its Importance, Review of Inferential Methods, and a New Bayesian Estimator

Vanda Inácio and María Xosé Rodríguez-Álvarez

Abstract. Accurate diagnosis of disease is of fundamental importance in clinical practice and medical research. Before a medical diagnostic test is routinely used in practice, its ability to distinguish between diseased and nondiseased states must be rigorously assessed. The receiver operating characteristic (ROC) curve is the most popular used tool for evaluating the diagnostic accuracy of continuous-outcome tests. It has been acknowledged that several factors (e.g., subject-specific characteristics such as age and/or gender) can affect the test outcomes and accuracy beyond disease status. Recently, the covariate-adjusted ROC curve has been proposed and successfully applied as a global summary measure of diagnostic accuracy that takes covariate information into account. The aim of this paper is three-fold. First, we motivate the importance of including covariate-information, whenever available, in ROC analysis and, in particular, how the covariate-adjusted ROC curve is an important tool in this context. Second, we review and provide insight on the existing approaches for estimating the covariate-adjusted ROC curve. Third, we develop a highly flexible Bayesian method, based on the combination of a Dirichlet process mixture of additive normal models and the Bayesian bootstrap, for conducting inference about the covariate-adjusted ROC curve. A simulation study is conducted to assess the performance of the different methods and it also demonstrates the ability of our proposed Bayesian model to successfully recover the true covariate-adjusted ROC curve and to produce valid inferences in a variety of complex scenarios. The methods are applied to an endocrine study where the goal is to assess the accuracy of the body mass index, adjusted for age and gender, for detecting clusters of cardiovascular disease risk factors.

Key words and phrases: Classification accuracy, covariate-adjustment, decision threshold, diagnostic test, Dirichlet process (mixture) model, receiver operating characteristic curve.

REFERENCES

ALONZO, T. A. and PEPE, M. S. (2002). Distibution–free ROC analysis using binary regression techniques. *Biostatistics* **3** 421–432.

ALTMAN, D. G. and ROYSTON, P. (2006). The cost of dichotomising continuous variables. *BMJ* **332** 1080. https://doi.org/10.1136/bmj.332.7549.1080

BRANSCUM, A. J., JOHNSON, W. O., HANSON, T. E. and BARON, A. T. (2015). Flexible regression models for ROC and risk analysis, with or without a gold standard. *Stat. Med.* **34** 3997–4015. MR3431318 https://doi.org/10.1002/sim.6610

CAI, T. (2004). Semi-parametric ROC regression analysis with placement values. *Biostatistics* **5** 45–60. https://doi.org/10.1093/biostatistics/5.1.45

DAVISON, A. C. and HINKLEY, D. V. (1997). *Bootstrap Methods and Their Application*. *Cambridge Series in Statistical and Probabilistic Mathematics* **1**. Cambridge Univ. Press, Cambridge. MR1478673 https://doi.org/10.1017/CBO9780511802843

DAWSON, N. V. and WEISS, R. (2012). Dichotomizing continuous

Vanda Inácio is Lecturer in Statistics, School of Mathematics, University of Edinburgh, Edinburgh, UK (e-mail: vanda.inacio@ed.ac.uk). María Xosé Rodríguez-Álvarez is a Ramón y Cajal Fellow, Department of Statistics and Operations Research, Universidade de Vigo, Vigo, Spain (e-mail: mxrodriguez@uvigo.es).

variables in statistical analysis: A practice to avoid. *Med. Decis. Mak.* **32** 225–226.

DE BOOR, C. (1978). *A Practical Guide to Splines*. *Applied Mathematical Sciences* **27**. Springer, New York–Berlin. MR0507062

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

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

FAN, J. and GIJBELS, I. (1996). *Local Polynomial Modelling and Its Applications*. *Monographs on Statistics and Applied Probability* **66**. CRC Press, London. MR1383587

FARAGGI, D. (2003). Adjusting receiver operating characteristic curves and related indices for covariates. *Statistician* **52** 179–192. MR1977259 https://doi.org/10.1111/1467-9884.00350

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

GABRY, J., SIMPSON, D., VEHTARI, A., BETANCOURT, M. and GELMAN, A. (2019). Visualization in Bayesian workflow. *J. Roy. Statist. Soc. Ser. A* **182** 389–402. MR3902665 https://doi.org/10.1111/rssa.12378

GEISSER, S. and EDDY, W. F. (1979). A predictive approach to model selection. *J. Amer. Statist. Assoc.* **74** 153–160. MR0529531

GELMAN, A., CARLIN, J. B., STERN, H. S., DUNSON, D. B., VEHTARI, A. and RUBIN, D. B. (2013). *Bayesian Data Analysis*. CRC Press, Boca Raton.

GELMAN, A., HWANG, J. and VEHTARI, A. (2014). Understanding predictive information criteria for Bayesian models. *Stat. Comput.* **24** 997–1016. MR3253850 https://doi.org/10.1007/s11222-013-9416-2

GNEITING, T. and VOGEL, P. (2018). Receiver Operating Characteristic (ROC) Curves.

GONZÁLEZ-MANTEIGA, W., PARDO-FERNÁNDEZ, J. C. and VAN KEILEGOM, I. (2011). ROC curves in non-parametric location-scale regression models. *Scand. J. Stat.* **38** 169–184. MR2760145 https://doi.org/10.1111/j.1467-9469.2010.00693.x

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

GU, J., GHOSAL, S. and ROY, A. (2008). Bayesian bootstrap estimation of ROC curve. *Stat. Med.* **27** 5407–5420. MR2542360 https://doi.org/10.1002/sim.3366

GUAN, Z., QIN, J. and ZHANG, B. (2012). Information borrowing methods for covariate-adjusted ROC curve. *Canad. J. Statist.* **40** 569–587. MR2968399 https://doi.org/10.1002/cjs.11145

HANSON, T. E. (2006). Inference for mixtures of finite Polya tree models. *J. Amer. Statist. Assoc.* **101** 1548–1565. MR2279479 https://doi.org/10.1198/016214506000000384

HAYFIELD, T. and RACINE, J. S. (2008). Nonparametric econometrics: The np package. *J. Stat. Softw.* **27**.

INÁCIO DE CARVALHO, V., DE CARVALHO, M. and BRANSCUM, A. J. (2017). Nonparametric Bayesian covariate-adjusted estimation of the Youden index. *Biometrics* **73** 1279–1288. MR3744541 https://doi.org/10.1111/biom.12686

INÁCIO DE CARVALHO, V., JARA, A., HANSON, T. E. and DE CARVALHO, M. (2013). Bayesian nonparametric ROC regression modeling. *Bayesian Anal.* **8** 623–645. MR3102228 https://doi.org/10.1214/13-BA825

INÁCIO, V. and RODRÍGUEZ-ÁLVAREZ, M. X (2022). Supplement to "The covariate-adjusted ROC curve: The concept and its importance, review of inferential methods, and a new Bayesian estimator." https://doi.org/10.1214/21-STS839SUPP

INÁCIO, V., RODRÍGUEZ-ÁLVAREZ, M. X. and GAYOSO-DIZ, P. (2021). Statistical evaluation of medical tests. *Annu. Rev. Stat. Appl.* **8** 41–67. MR4243538 https://doi.org/10.1146/annurev-statistics-040720-022432

ISHWARAN, H. and JAMES, L. F. (2001). Gibbs sampling methods for stick-breaking priors. *J. Amer. Statist. Assoc.* **96** 161–173. MR1952729 https://doi.org/10.1198/016214501750332758

JANES, H. and PEPE, M. S. (2008a). Adjusting for covariates in studies of diagnostic, screening, or prognostic markers: An old concept in a new setting. *Am. J. Epidemiol.* **168** 89–97.

JANES, H. and PEPE, M. S. (2008b). Matching in studies of classification accuracy: Implications for analysis, efficiency, and assessment of incremental value. *Biometrics* **64** 1–9. MR2422813 https://doi.org/10.1111/j.1541-0420.2007.00823.x

JANES, H. and PEPE, M. S. (2009). Adjusting for covariate effects on classification accuracy using the covariate-adjusted receiver operating characteristic curve. *Biometrika* **96** 371–382. MR2507149 https://doi.org/10.1093/biomet/asp002

KERR, K. F. and PEPE, M. S. (2011). Joint modeling, covariate adjustment, and interaction: Contrasting notions in risk prediction models and risk prediction performance. *Epidemiology* **22** 805–812.

KIM, S. and HUANG, Y. (2017). Combining biomarkers for classification with covariate adjustment. *Stat. Med.* **36** 2347–2362. MR3660136 https://doi.org/10.1002/sim.7274

LIN, H., ZHOU, L., PENG, H. and ZHOU, X.-H. (2011). Selection and combination of biomarkers using ROC method for disease classification and prediction. *Canad. J. Statist.* **39** 324–343. MR2839483 https://doi.org/10.1002/cjs.10107

LIU, C., LIU, A. and HALABI, S. (2011). A min-max combination of biomarkers to improve diagnostic accuracy. *Stat. Med.* **30** 2005–2014. MR2829061 https://doi.org/10.1002/sim.4238

LIU, D. and ZHOU, X. H. (2013). ROC analysis in biomarker combination with covariate adjustment. *Acad. Radiol.* **20** 874–882.

LÓPEZ-RATÓN, M., RODRÍGUEZ-ÁLVAREZ, M. X., CADARSO-SUÁREZ, C. and GUDE-SAMPEDRO, F. (2014). OptimalCutpoints: An R package for selecting optimal cutpoints in diagnostic tests. *J. Stat. Softw.* **61** 1–36.

MEISNER, A., PARIKH, C. R. and KERR, K. F. (2020). Developing biomarker combinations in multicenter studies via direct maximization and penalization. *Stat. Med.* **39** 3412–3426. MR4157847 https://doi.org/10.1002/sim.8673

PARDO-FERNÁNDEZ, J. C., RODRÍGUEZ-ÁLVAREZ, M. X. and VAN KEILEGOM, I. (2014). A review on ROC curves in the presence of covariates. *REVSTAT* **12** 21–41.

PEPE, M. S. (1998). Three approaches to regression analysis of receiver operating characteristic curves for continuous test results. *Biometrics* **54** 124–135.

PEPE, M. S. (2003). *The Statistical Evaluation of Medical Tests for Classification and Prediction*. *Oxford Statistical Science Series* **28**. Oxford Univ. Press, Oxford. MR2260483

PEPE, M. S. and CAI, T. (2004). The analysis of placement values for evaluating discriminatory measures. *Biometrics* **60** 528–535. MR2067011 https://doi.org/10.1111/j.0006-341X.2004.00200.x

PEPE, M. S., CAI, T. and LONGTON, G. (2006). Combining predictors for classification using the area under the receiver operating characteristic curve. *Biometrics* **62** 221–229. MR2226576 https://doi.org/10.1111/j.1541-0420.2005.00420.x

PEPE, M., FAN, J. and SEYMOUR, C. (2013). Estimating the receiver operating characteristic curve in studies that match controls to cases on covariates. *Acad. Radiol.* **20** 863–873.

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

RODRÍGUEZ, A. and MARTÍNEZ, J. C. (2013). Bayesian semiparametric estimation of covariate-dependent ROC curves. *Biostatistics* **15** 353–369.

RODRÍGUEZ-ÁLVAREZ, M. X. and INÁCIO, V. (2021). ROCnReg: An R package for receiver operating characteristic curve inference with and without covariates. *R J.* **13** 525–555. https://doi.org/10.32614/RJ-2021-066

RODRÍGUEZ-ÁLVAREZ, M. X., ROCA-PARDIÑAS, J. and CADARSO-SUÁREZ, C. (2011a). ROC curve and covariates: Extending induced methodology to the non-parametric framework. *Stat. Comput.* **21** 483–499. MR2826687 https://doi.org/10.1007/s11222-010-9184-1

RODRÍGUEZ-ÁLVAREZ, M. X., ROCA-PARDIÑAS, J. and CADARSO-SUÁREZ, C. (2011b). A new flexible direct ROC regression model: Application to the detection of cardiovascular risk factors by anthropometric measures. *Comput. Statist. Data Anal.* **55** 3257–3270. MR2825408 https://doi.org/10.1016/j.csda.2011.06.008

RODRÍGUEZ-ÁLVAREZ, M. X., TAHOCES, P. G., CADARSO-SUÁREZ, C. and LADO, M. J. (2011). Comparative study of ROC regression techniques—applications for the computer-aided diagnostic system in breast cancer detection. *Comput. Statist. Data Anal.* **55** 888–902. MR2736605 https://doi.org/10.1016/j.csda.2010.07.018

ROSENBERG, P. S. (1995). Hazard function estimation using B-splines. *Biometrics* **51** 874–887.

RUBIN, D. B. (1981). The Bayesian bootstrap. *Ann. Statist.* **9** 130–134. MR0600538

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

TAYLOR, J. M. G. and YU, M. (2002). Bias and efficiency loss due to categorizing an explanatory variable. *J. Multivariate Anal.* **83** 248–263. MR1934983 https://doi.org/10.1006/jmva.2001.2045

TOMÉ MARTÍNEZ DE RITUERTO, M. A., BOTANA, M. A., CADARSO-SUÁREZ, C., REGO-IRAETA, A., FERNÁNDEZ-MARIÑO, A., MATO, J. A., SOLACHE, I. and PEREZ-FERNANDEZ, R. (2009). Prevalence of metabolic syndrome in Galicia (NW Spain) on four alternative definitions and association with insulin resistance. *J. Endocrinol. Invest.* **32** 505–511.

WICKHAM, H. (2009). *Ggplot2: Elegant Graphics for Data Analysis*. Springer, New York.

WOOD, S. N. (2017). *Generalized Additive Models: An Introduction with R. Texts in Statistical Science Series*. CRC Press, Boca Raton, FL. MR3726911

ZHOU, X.-H., OBUCHOWSKI, N. A. and MCCLISH, D. K. (2011). *Statistical Methods in Diagnostic Medicine. Wiley Series in Probability and Statistics*. Wiley, Hoboken, NJ. MR2816760 https://doi.org/10.1002/9780470906514

ZWEIG, M. H. and CAMPBELL, G. (1993). Receiver-operating characteristic (ROC) plots: A fundamental evaluation tool in clinical medicine. *Clin. Chem.* **39** 561–577.

Statistical Science
2022, Vol. 37, No. 4, 562–579
https://doi.org/10.1214/21-STS841

A Regression Perspective on Generalized Distance Covariance and the Hilbert–Schmidt Independence Criterion

Dominic Edelmann and Jelle Goeman

Abstract. In a seminal paper, Sejdinovic et al. (*Ann. Statist.* **41** (2013) 2263–2291) showed the equivalence of the Hilbert–Schmidt Independence Criterion (HSIC) and a generalization of distance covariance. In this paper, the two notions of dependence are unified with a third prominent concept for independence testing, the "global test" introduced in (*J. R. Stat. Soc. Ser. B. Stat. Methodol.* **68** (2006) 477–493). The new viewpoint provides novel insights into all three test traditions, as well as a unified overall view of the way all three tests contrast with classical association tests. As our main result, a regression perspective on HSIC and generalized distance covariance is obtained, allowing such tests to be used with nuisance covariates or for survival data. Several more examples of cross-fertilization of the three traditions are provided, involving theoretical results and novel methodology. To illustrate the difference between classical statistical tests and the unified HSIC/distance covariance/global tests we investigate the case of association between two categorical variables in depth.

Key words and phrases: Distance covariance, distance correlation, Hilbert–Schmidt Independence Criterion, global test, equivalence, locally most powerful.

REFERENCES

[1] BARINGHAUS, L. and FRANZ, C. (2004). On a new multivariate two-sample test. *J. Multivariate Anal.* **88** 190–206. MR2021870 https://doi.org/10.1016/S0047-259X(03)00079-4

[2] BERG, C., CHRISTENSEN, J. P. R. and RESSEL, P. (1984). *Harmonic Analysis on Semigroups: Theory of Positive Definite and Related Functions. Graduate Texts in Mathematics* **100**. Springer, New York. MR0747302 https://doi.org/10.1007/978-1-4612-1128-0

[3] BERSCHNEIDER, G. and BÖTTCHER, B. (2019). On complex Gaussian random fields, Gaussian quadratic forms and sample distance multivariance. Preprint. arXiv:1808.07280.

[4] BÖTTCHER, B., KELLER-RESSEL, M. and SCHILLING, R. L. (2018). Detecting independence of random vectors: Generalized distance covariance and Gaussian covariance. *Modern Stoch. Theory Appl.* **5** 353–383. MR3868546 https://doi.org/10.15559/18-vmsta116

[5] CAI, T., TONINI, G. and LIN, X. (2011). Kernel machine approach to testing the significance of multiple genetic markers for risk prediction. *Biometrics* **67** 975–986. MR2829272 https://doi.org/10.1111/j.1541-0420.2010.01544.x

[6] CHATURVEDI, N., DE MENEZES, R. X. and GOEMAN, J. J. (2017). A global $\times$ global test for testing associations between two large sets of variables. *Biom. J.* **59** 145–158. MR3593726 https://doi.org/10.1002/bimj.201500106

[7] 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

[8] DUECK, J., EDELMANN, D. and RICHARDS, D. (2017). Distance correlation coefficients for Lancaster distributions. *J. Multivariate Anal.* **154** 19–39. MR3588555 https://doi.org/10.1016/j.jmva.2016.10.012

[9] EDELMANN, D., FOKIANOS, K. and PITSILLOU, M. (2019). An updated literature review of distance correlation and its applications to time series. *Int. Stat. Rev.* **87** 237–262. MR3994758 https://doi.org/10.1111/insr.12294

[10] EDELMANN, D. and GOEMAN, J. (2022). Supplement to "A Regression Perspective on Generalized Distance Covariance and the Hilbert–Schmidt Independence Criterion." https://doi.org/10.1214/21-STS841SUPP

[11] EDELMANN, D., RICHARDS, D. and VOGEL, D. (2020). The distance standard deviation. *Ann. Statist.* **48** 3395–3416. MR4185813 https://doi.org/10.1214/19-AOS1935

[12] EDELMANN, D., SAADATI, M., PUTTER, H. and GOEMAN, J. (2020). A global test for competing risks survival analysis. *Stat. Methods Med. Res.* **29** 3666–3683. MR4159219 https://doi.org/10.1177/0962280220938402

Dominic Edelmann is Professor, German Cancer Research Center, Im Neuenheimer Feld 280, 69120 Heidelberg, Germany (e-mail: dominic.edelmann@dkfz.de). Jelle Goeman is Postdoctoral Researcher, Leiden University Medical Center, Albinusdreef 2, 2333 ZA Leiden, Netherlands (e-mail: J.J.Goeman@lumc.nl).

[13] EFRON, B., HASTIE, T., JOHNSTONE, I. and TIBSHIRANI, R. (2004). Least angle regression. *Ann. Statist.* **32** 407–499. MR2060166 https://doi.org/10.1214/009053604000000067

[14] GOEMAN, J. J. and LE CESSIE, S. (2006). A goodness-of-fit test for multinomial logistic regression. *Biometrics* **62** 980–985. MR2297668 https://doi.org/10.1111/j.1541-0420.2006.00581.x

[15] GOEMAN, J. J., OOSTING, J., CLETON-JANSEN, A.-M., ANNINGA, J. K. and VAN HOUWELINGEN, H. C. (2005). Testing association of a pathway with survival using gene expression data. *Bioinformatics* **21** 1950–1957.

[16] GOEMAN, J. J., VAN DE GEER, S. A., DE KORT, F. and VAN HOUWELINGEN, H. C. (2004). A global test for groups of genes: Testing association with a clinical outcome. *Bioinformatics* **20** 93–99.

[17] GOEMAN, J. J., VAN HOUWELINGEN, H. C. and FINOS, L. (2011). Testing against a high-dimensional alternative in the generalized linear model: Asymptotic type I error control. *Biometrika* **98** 381–390. MR2806435 https://doi.org/10.1093/biomet/asr016

[18] GOEMAN, J. J., VAN DE GEER, S. A. and VAN HOUWELINGEN, H. C. (2006). Testing against a high dimensional alternative. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **68** 477–493. MR2278336 https://doi.org/10.1111/j.1467-9868.2006.00551.x

[19] GRETTON, A., BORGWARDT, K. M., RASCH, M. J., SCHÖLKOPF, B. and SMOLA, A. (2007). A kernel method for the two-sample problem. *Adv. Neural Inf. Process. Syst.* **20** 513–520.

[20] GRETTON, A., BORGWARDT, K. M., RASCH, M. J., SCHÖLKOPF, B. and SMOLA, A. (2012). A kernel two-sample test. *J. Mach. Learn. Res.* **13** 723–773. MR2913716

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

[22] GRETTON, A., FUKUMIZU, K. and SRIPERUMBUDUR, B. K. (2009). Discussion of: Brownian distance covariance [MR2752127]. *Ann. Appl. Stat.* **3** 1285–1294. MR2752132 https://doi.org/10.1214/09-AOAS312E

[23] GRETTON, A., FUKUMIZU, K., TEO, C. H., SONG, L., SCHÖLKOPF, B. and SMOLA, A. J. (2008). A kernel statistical test of independence. In *Adv. Neur. Inf. Proc. Sys.* **21** 585–592.

[24] GUO, B. and CHEN, S. X. (2016). Tests for high dimensional generalized linear models. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **78** 1079–1102. MR3557190 https://doi.org/10.1111/rssb.12152

[25] HEMERIK, J., GOEMAN, J. J. and FINOS, L. (2020). Robust testing in generalized linear models by sign flipping score contributions. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **82** 841–864. MR4112787 https://doi.org/10.1111/rssb.12369

[26] HOFMANN, T., SCHÖLKOPF, B. and SMOLA, A. J. (2008). Kernel methods in machine learning. *Ann. Statist.* **36** 1171–1220. MR2418654 https://doi.org/10.1214/009053607000000677

[27] HONEINE, P. (2014). An eigenanalysis of data centering in machine learning. Preprint. arXiv:1407.2904.

[28] HUA, W.-Y. and GHOSH, D. (2015). Equivalence of kernel machine regression and kernel distance covariance for multidimensional phenotype association studies. *Biometrics* **71** 812–820. MR3402617 https://doi.org/10.1111/biom.12314

[29] HUANG, C. and HUO, X. (2017). A statistically and numerically efficient independence test based on random projections and distance covariance. Preprint. arXiv:1701.06054.

[30] JAKOBSEN, M. E. (2017). Distance Covariance in Metric Spaces: Non-Parametric Independence Testing in Metric Spaces. Master's thesis. University of Copenhagen. Available on arXiv:1706.03490.

[31] JELIZAROW, M., MANSMANN, U. and GOEMAN, J. J. (2016). A Cochran-Armitage-type and a score-free global test for multivariate ordinal data. *Stat. Med.* **35** 2754–2769. MR3513716 https://doi.org/10.1002/sim.6898

[32] KONG, J., KLEIN, B. E., KLEIN, R., LEE, K. E. and WAHBA, G. (2012). Using distance correlation and SS-ANOVA to assess associations of familial relationships, lifestyle factors, diseases, and mortality. *Proc. Natl. Acad. Sci. USA* **109** 20352–20357.

[33] KONG, J., KLEIN, B. E., KLEIN, R., LEE, K. E. and WAHBA, G. (2013). Correction for Kong et al., Using distance correlation and SS-ANOVA to assess associations of familial relationships, lifestyle factors, diseases, and mortality. *Proc. Natl. Acad. Sci. USA* **110** 13691–13691.

[34] LEE, A. J. (2019). *U-Statistics: Theory and Practice*. Routledge, London.

[35] LE CESSIE, S. and VAN HOUWELINGEN, H. C. (1995). Testing the fit of a regression model via score tests in random effects models. *Biometrics* 600–614.

[36] LIU, C. and RUBIN, D. B. (1995). ML estimation of the t distribution using EM and its extensions, ECM and ECME. *Statist. Sinica* **5** 19–39. MR1329287

[37] LOÈVE, M. (1978). *Probability Theory. II*, 4th ed. *Graduate Texts in Mathematics*, *Vol.* 46. Springer, New York. MR0651018

[38] LYONS, R. (2013). Distance covariance in metric spaces. *Ann. Probab.* **41** 3284–3305. MR3127883 https://doi.org/10.1214/12-AOP803

[39] MARTÍNEZ-GÓMEZ, E., RICHARDS, M. T. and RICHARDS, D. S. P. (2014). Distance correlation methods for discovering associations in large astrophysical databases. *Astrophys. J.* **781** 39 (11pp).

[40] MINH, H. Q., NIYOGI, P. and YAO, Y. (2006). Mercer's theorem, feature maps, and smoothing. In *Learning Theory. Lecture Notes in Computer Science* **4005** 154–168. Springer, Berlin. MR2280604 https://doi.org/10.1007/11776420_14

[41] PAN, W. (2009). Asymptotic tests of association with multiple SNPs in linkage disequilibrium. *Genet. Epidemiol.* **33** 497–507.

[42] 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

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

[44] RICHARDS, M. T., RICHARDS, D. S. P. and MARTÍNEZ-GÓMEZ, E. (2014). Interpreting the distance correlation results for the COMBO-17 survey. *Astrophys. J. Lett.* **784** L34 (5 pp.).

[45] RIZZO, M. L. and SZÉKELY, G. J. (2010). DISCO analysis: A nonparametric extension of analysis of variance. *Ann. Appl. Stat.* **4** 1034–1055. MR2758432 https://doi.org/10.1214/09-AOAS245

[46] ROBBINS, H. and PITMAN, E. J. G. (1949). Application of the method of mixtures to quadratic forms in normal variates. *Ann. Math. Stat.* **20** 552–560. MR0032151 https://doi.org/10.1214/aoms/1177729947

[47] ROBERT, P. and ESCOUFIER, Y. (1976). A unifying tool for linear multivariate statistical methods: The RV-coefficient. *J. R. Stat. Soc. Ser. C. Appl. Stat.* **25** 257–265. MR0440801 https://doi.org/10.2307/2347233

[48] ROSIPAL, R. and KRÄMER, N. (2006). Overview and recent advances in partial least squares. In *Subspace, Latent Structure and Feature Selection. SLSFS* 2005 (C. Saunders, M. Grobelnik, S. Gunn and J. Shawe-Taylor, eds.). *Lecture Notes in Computer Science* **3940** 34–51. Springer, Berlin.

[49] ROSIPAL, R. and TREJO, L. J. (2001). Kernel partial least squares regression in reproducing kernel Hilbert space. *J. Mach. Learn. Res.* **2** 97–123.

[50] SCHÖLKOPF, B., SMOLA, A. and MÜLLER, K.-R. (1997). Kernel principal component analysis. In *Artificial Neural Networks – ICANN'97. ICANN* 1997 (W. Gerstner, A. Germond, M. Hasler and J. D. Nicoud, eds.). *Lecture Notes in Computer Science* **1327** 583–588. Springer, Berlin.

[51] 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

[52] SHEN, C., PRIEBE, C. E. and VOGELSTEIN, J. T. (2019). The exact equivalence of independence testing and two-sample testing. Preprint. arXiv:1910.08883.

[53] SHEN, C. and VOGELSTEIN, J. T. (2021). The exact equivalence of distance and kernel methods in hypothesis testing. *AStA Adv. Stat. Anal.* **105** 385–403. MR4304315 https://doi.org/10.1007/s10182-020-00378-1

[54] SOLARI, A., FINOS, L. and GOEMAN, J. J. (2014). Rotation-based multiple testing in the multivariate linear model. *Biometrics* **70** 954–961. MR3295756 https://doi.org/10.1111/biom.12238

[55] SOLARI, A., LE CESSIE, S. and GOEMAN, J. J. (2012). Testing goodness of fit in regression: A general approach for specified alternatives. *Stat. Med.* **31** 3656–3666. MR3041837 https://doi.org/10.1002/sim.5417

[56] SRIVASTAVA, M. S. and DU, M. (2008). A test for the mean vector with fewer observations than the dimension. *J. Multivariate Anal.* **99** 386–402. MR2396970 https://doi.org/10.1016/j.jmva.2006.11.002

[57] STEINWART, I. and SCOVEL, C. (2012). Mercer's theorem on general domains: On the interaction between measures, kernels, and RKHSs. *Constr. Approx.* **35** 363–417. MR2914365 https://doi.org/10.1007/s00365-012-9153-3

[58] SUN, H. (2005). Mercer theorem for RKHS on noncompact sets. *J. Complexity* **21** 337–349. MR2138444 https://doi.org/10.1016/j.jco.2004.09.002

[59] 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

[60] SZÉKELY, G. J. and RIZZO, M. L. (2013). The distance correlation t-test of independence in high dimension. *J. Multivariate Anal.* **117** 193–213. MR3053543 https://doi.org/10.1016/j.jmva.2013.02.012

[61] 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

[62] 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

[63] SZEKELY, G. J. and RIZZO, M. L. (2017). The energy of data. *Annu. Rev. Stat. Appl.* **4** 447–479.

[64] 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

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

[66] WANG, X., PAN, W., HU, W., TIAN, Y. and ZHANG, H. (2015). Conditional distance correlation. *J. Amer. Statist. Assoc.* **110** 1726–1734. MR3449068 https://doi.org/10.1080/01621459.2014.993081

[67] WOLD, S., SJÖSTRÖM, M. and ERIKSSON, L. (2001). PLS-regression: A basic tool of chemometrics. *Chemom. Intell. Lab. Syst.* **58** 109–130.

[68] WU, M. C., LEE, S., CAI, T., LI, Y., BOEHNKE, M. and LIN, X. (2011). Rare-variant association testing for sequencing data with the sequence kernel association test. *Am. J. Hum. Genet.* **89** 82–93.

[69] ZHANG, K., PETERS, J., JANZING, D. and SCHÖLKOPF, B. (2011). Kernel-based conditional independence test and application in causal discovery. In *Proceedings of the Twenty-Seventh Conference on Uncertainty in Artificial Intelligence. UAI'11* 804–813. AUAI Press, Arlington, VA, USA.

[70] ZHANG, Z., DAI, G., WANG, D. and JORDAN, M. I. (2010). Bayesian generalized kernel models. In *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics* 972–979.

[71] ZOU, H. and HASTIE, T. (2005). Regularization and variable selection via the elastic net. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **67** 301–320. MR2137327 https://doi.org/10.1111/j.1467-9868.2005.00503.x

Statistical Science
2022, Vol. 37, No. 4, 580–597
https://doi.org/10.1214/21-STS842

Methods to Compute Prediction Intervals: A Review and New Results

Qinglong Tian, Daniel J. Nordman and William Q. Meeker

Abstract. The purpose of this paper is to review both classic and modern methods for constructing prediction intervals. We focus, primarily, on model-based non-Bayesian methods for the prediction of a scalar random variable, but we also include Bayesian methods with objective prior distributions. Our review of non-Bayesian methods follows two lines: general methods based on (approximate) pivotal quantities and methods based on non-Bayesian predictive distributions. The connection between these two types of methods is described for distributions in the (log-)location-scale family. We also discuss extending the general prediction methods to data with complicated dependence structures as well as some nonparametric prediction methods (e.g., conformal prediction).

Key words and phrases: Bootstrap, calibration, coverage probability, prediction interval, predictive distribution.

REFERENCES

AITCHISON, J. and DUNSMORE, I. R. (1975). *Statistical Prediction Analysis*. Cambridge Univ. Press, Cambridge-Melbourne. MR0408097

AKTEKIN, T., POLSON, N. and SOYER, R. (2018). Sequential Bayesian analysis of multivariate count data. *Bayesian Anal.* **13** 385–409. MR3780428 https://doi.org/10.1214/17-BA1054

BAKER, G. A. (1935). The probability that the mean of a second sample will differ from the mean of a first sample by less than a certain multiple of the standard deviation of the first sample. *Ann. Math. Stat.* **6** 197–201.

BARNDORFF-NIELSEN, O. E. and COX, D. R. (1996). Prediction and asymptotics. *Bernoulli* **2** 319–340. MR1440272 https://doi.org/10.2307/3318417

BERAN, R. (1990). Calibrating prediction regions. *J. Amer. Statist. Assoc.* **85** 715–723. MR1138352

BERAN, R. and HALL, P. (1993). Interpolated nonparametric prediction intervals and confidence intervals. *J. Roy. Statist. Soc. Ser. B* **55** 643–652. MR1223933

BERRY, L. R. and WEST, M. (2020). Bayesian forecasting of many count-valued time series. *J. Bus. Econom. Statist.* **38** 872–887. MR4154894 https://doi.org/10.1080/07350015.2019.1604372

BESAG, J. (1974). Spatial interaction and the statistical analysis of lattice systems. *J. Roy. Statist. Soc. Ser. B* **36** 192–236. MR0373208

BJØRNSTAD, J. F. (1990). Predictive likelihood: A review. *Statist. Sci.* **5** 262–265.

BROCKWELL, P. J. and DAVIS, R. A. (2016). *Introduction to Time Series and Forecasting*, 3rd ed. *Springer Texts in Statistics*. Springer, Cham. MR3526245 https://doi.org/10.1007/978-3-319-29854-2

CASELLA, G. and BERGER, R. L. (1990). *Statistical Inference*. *The Wadsworth & Brooks/Cole Statistics/Probability Series*. Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove, CA. MR1051420

CASLETON, E., NORDMAN, D. and KAISER, M. (2017). A local structure model for network analysis. *Stat. Interface* **10** 355–367. MR3572115 https://doi.org/10.4310/SII.2017.v10.n2.a15

CHEN, P. and YE, Z.-S. (2017). Approximate statistical limits for a gamma distribution. *J. Qual. Technol.* **49** 64–77.

CHRISTENSEN, O. F. and WAAGEPETERSEN, R. (2002). Bayesian prediction of spatial count data using generalized linear mixed models. *Biometrics* **58** 280–286. MR1908167 https://doi.org/10.1111/j.0006-341X.2002.00280.x

CLARKE, B. S. and CLARKE, J. L. (2018). *Predictive Statistics*: *Analysis and Inference Beyond Models*. *Cambridge Series in Statistical and Probabilistic Mathematics* **46**. Cambridge Univ. Press, Cambridge. MR3791464 https://doi.org/10.1017/9781139236003

COX, D. R. (1975). Prediction intervals and empirical Bayes confidence intervals. In *Perspectives in Probability and Statistics* (*Papers in Honour of M. S. Bartlett on the Occasion of His* 65*th Birthday*) 47–55. MR0403046 https://doi.org/10.1017/s0021900200047550

CRESSIE, N. A. C. (2015). *Statistics for Spatial Data*, Revised ed. *Wiley Classics Library*. Wiley, New York. MR3559472

DE FINETTI, B. (1937). La prévision : Ses lois logiques, ses sources subjectives. *Ann. Inst. Henri Poincaré* **7** 1–68. MR1508036

DE FINETTI, B. (1975). *Theory of Probability*: *A Critical Introductory Treatment*. *Vol.* 2. *Wiley Series in Probability and Mathematical Statistics*. Wiley, London–Sydney. MR0440641

DE OLIVEIRA, V., KEDEM, B. and SHORT, D. A. (1997). Bayesian prediction of transformed Gaussian random fields. *J. Amer. Statist. Assoc.* **92** 1422–1433. MR1615252 https://doi.org/10.2307/2965412

Qinglong Tian is Ph.D. Candidate, Department of Statistics, Iowa State University, Ames, Iowa 50011, USA (e-mail: qltian@iastate.edu). Daniel J. Nordman is Professor, Department of Statistics, Iowa State University, Ames, Iowa 50011, USA (e-mail: dnordman@iastate.edu). William Q. Meeker is Distinguished Professor, Department of Statistics, Iowa State University, Ames, Iowa 50011, USA (e-mail: wqmeeker@iastate.edu).

DE OLIVEIRA, V. and KONE, B. (2015a). Prediction intervals for integrals of Gaussian random fields. *Comput. Statist. Data Anal.* **83** 37–51. MR3281794 https://doi.org/10.1016/j.csda.2014.09.013

DE OLIVEIRA, V. and KONE, B. (2015b). Prediction intervals for integrals of some types of non-Gaussian random fields: A semiparametric bootstrap approach. In *JSM Proceedings, Statistics and the Environment Section* 2588–2597.

DE OLIVEIRA, V. and RUI, C. (2009). On shortest prediction intervals in log-Gaussian random fields. *Comput. Statist. Data Anal.* **53** 4345–4357. MR2744330 https://doi.org/10.1016/j.csda.2009.05.030

DEMPSTER, A. P. (2008). The Dempster–Shafer calculus for statisticians. *Internat. J. Approx. Reason.* **48** 365–377. MR2419025 https://doi.org/10.1016/j.ijar.2007.03.004

DICICCIO, T. J., KUFFNER, T. A. and YOUNG, G. A. (2017). A simple analysis of the exact probability matching prior in the location-scale model. *Amer. Statist.* **71** 302–304. MR3750936 https://doi.org/10.1080/00031305.2016.1255662

DUNSMORE, I. R. (1976). A note on Faulkenberry's method of obtaining prediction intervals. *J. Amer. Statist. Assoc.* **71** 193–194.

EFRON, B. (2020). Prediction, estimation, and attribution. *J. Amer. Statist. Assoc.* **115** 636–655. MR4107663 https://doi.org/10.1080/01621459.2020.1762613

FAULKENBERRY, G. D. (1973). A method of obtaining prediction intervals. *J. Amer. Statist. Assoc.* **68** 433–435.

FISHER, R. A. (1935). The fiducial argument in statistical inference. *Ann. Eugenics* **6** 391–398.

FLIGNER, M. A. and WOLFE, D. A. (1976). Some applications of sample analogues to the probability integral transformation and a coverage property. *Amer. Statist.* **30** 78–85. MR0478462 https://doi.org/10.2307/2683799

FONSECA, G., GIUMMOLÈ, F. and VIDONI, P. (2014). Calibrating predictive distributions. *J. Stat. Comput. Simul.* **84** 373–383. MR3169333 https://doi.org/10.1080/00949655.2012.709518

FORTINI, S. and PETRONE, S. (2014). In Predictive Distribution (de Finetti's View). Wiley StatsRef: Statistics Reference Online 1–9.

FREY, J. (2013). Data-driven nonparametric prediction intervals. *J. Statist. Plann. Inference* **143** 1039–1048. MR3029231 https://doi.org/10.1016/j.jspi.2013.01.004

GEISSER, S. (1993). *Predictive Inference: An Introduction. Monographs on Statistics and Applied Probability* **55**. CRC Press, New York. MR1252174 https://doi.org/10.1007/978-1-4899-4467-2

GEORGE, E. I., LIANG, F. and XU, X. (2012). From minimax shrinkage estimation to minimax shrinkage prediction. *Statist. Sci.* **27** 82–94. MR2953497 https://doi.org/10.1214/11-STS383

GIUMMOLÈ, F. and VIDONI, P. (2010). Improved prediction limits for a general class of Gaussian models. *J. Time Series Anal.* **31** 483–493. MR2732602 https://doi.org/10.1111/j.1467-9892.2010.00680.x

GREGORY, K. B., LAHIRI, S. N. and NORDMAN, D. J. (2018). A smooth block bootstrap for quantile regression with time series. *Ann. Statist.* **46** 1138–1166. MR3797999 https://doi.org/10.1214/17-AOS1580

GUTTMAN, I. and TIAO, G. C. (1964). A Bayesian approach to some best population problems. *Ann. Math. Stat.* **35** 825–835. MR0161442 https://doi.org/10.1214/aoms/1177703582

HALL, P. and MAITI, T. (2006). On parametric bootstrap methods for small area prediction. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **68** 221–238. MR2188983 https://doi.org/10.1111/j.1467-9868.2006.00541.x

HALL, P., PENG, L. and TAJVIDI, N. (1999). On prediction intervals based on predictive likelihood or bootstrap methods. *Biometrika* **86** 871–880. MR1741983 https://doi.org/10.1093/biomet/86.4.871

HANDCOCK, M. S. and STEIN, M. L. (1993). A Bayesian analysis of kriging. *Technometrics* **35** 403–410.

HANNIG, J., IYER, H. and PATTERSON, P. (2006). Fiducial generalized confidence intervals. *J. Amer. Statist. Assoc.* **101** 254–269. MR2268043 https://doi.org/10.1198/016214505000000736

HANNIG, J., IYER, H., LAI, R. C. S. and LEE, T. C. M. (2016). Generalized fiducial inference: A review and new results. *J. Amer. Statist. Assoc.* **111** 1346–1361. MR3561954 https://doi.org/10.1080/01621459.2016.1165102

HARRIS, I. R. (1989). Predictive fit for natural exponential families. *Biometrika* **76** 675–684. MR1041412 https://doi.org/10.1093/biomet/76.4.675

HARVILLE, D. A. (2014). The need for more emphasis on prediction: A "nondenominational" model-based approach. *Amer. Statist.* **68** 71–83. MR3303836 https://doi.org/10.1080/00031305.2013.836987

HARVILLE, D. A. and CARRIQUIRY, A. L. (1992). Classical and Bayesian prediction as applied to an unbalanced mixed linear model. *Biometrics* **48** 987–1003. MR1212854 https://doi.org/10.2307/2532693

HINKLEY, D. (1979). Predictive likelihood. *Ann. Statist.* **7** 718–728. MR0532237

HULTING, F. L. and HARVILLE, D. A. (1991). Some Bayesian and non-Bayesian procedures for the analysis of comparative experiments and for small-area estimation: Computational aspects, frequentist properties, and relationships. *J. Amer. Statist. Assoc.* **86** 557–568. MR1147086

JENTSCH, C. and POLITIS, D. N. (2015). Covariance matrix estimation and linear process bootstrap for multivariate time series of possibly increasing dimension. *Ann. Statist.* **43** 1117–1140. MR3346699 https://doi.org/10.1214/14-AOS1301

KAISER, M. S. and CRESSIE, N. (2000). The construction of multivariate distributions from Markov random fields. *J. Multivariate Anal.* **73** 199–220. MR1763324 https://doi.org/10.1006/jmva.1999.1878

KAPLAN, A., KAISER, M. S., LAHIRI, S. N. and NORDMAN, D. J. (2020). Simulating Markov random fields with a conclique-based Gibbs sampler. *J. Comput. Graph. Statist.* **29** 286–296. MR4116042 https://doi.org/10.1080/10618600.2019.1668800

KOMAKI, F. (1996). On asymptotic properties of predictive distributions. *Biometrika* **83** 299–313. MR1439785 https://doi.org/10.1093/biomet/83.2.299

KREISS, J.-P. and LAHIRI, S. N. (2012). Bootstrap methods for time series. In *Handbook of Statistics* **30** 3–26. Elsevier, Amsterdam.

KREISS, J.-P., PAPARODITIS, E. and POLITIS, D. N. (2011). On the range of validity of the autoregressive sieve bootstrap. *Ann. Statist.* **39** 2103–2130. MR2893863 https://doi.org/10.1214/11-AOS900

KRISHNAMOORTHY, K. and LEE, M. (2010). Inference for functions of parameters in discrete distributions based on fiducial approach: Binomial and Poisson cases. *J. Statist. Plann. Inference* **140** 1182–1192. MR2581121 https://doi.org/10.1016/j.jspi.2009.11.004

KRISHNAMOORTHY, K. and MATHEW, T. (2009). *Statistical Tolerance Regions: Theory, Applications, and Computation. Wiley Series in Probability and Statistics*. Wiley, Hoboken, NJ. MR2500599 https://doi.org/10.1002/9780470473900

KRISHNAMOORTHY, K. and PENG, J. (2011). Improved closed-form prediction intervals for binomial and Poisson distributions. *J. Statist. Plann. Inference* **141** 1709–1718. MR2763201 https://doi.org/10.1016/j.jspi.2010.11.021

LAHIRI, S. N. (2006). Bootstrap methods: A review. In *Frontiers in Statistics* 231–255. Imp. Coll. Press, London. MR2326004 https://doi.org/10.1142/9781860948886_0011

LAWLESS, J. F. (1972). Conditional confidence interval procedures for the location and scale parameters of the Cauchy and logistic distributions. *Biometrika* **59** 377–386. MR0334380 https://doi.org/10.1093/biomet/59.2.377

LAWLESS, J. F. and FREDETTE, M. (2005). Frequentist prediction intervals and predictive distributions. *Biometrika* **92** 529–542. MR2202644 https://doi.org/10.1093/biomet/92.3.529

LEI, J. (2019). Fast exact conformalization of the lasso using piecewise linear homotopy. *Biometrika* **106** 749–764. MR4031197 https://doi.org/10.1093/biomet/asz046

LEI, J., G'SELL, M., RINALDO, A., TIBSHIRANI, R. J. and WASSERMAN, L. (2018). Distribution-free predictive inference for regression. *J. Amer. Statist. Assoc.* **113** 1094–1111. MR3862342 https://doi.org/10.1080/01621459.2017.1307116

MATHIASEN, P. E. (1979). Prediction functions. *Scand. J. Stat.* **6** 1–21. MR0529057

MCALINN, K. and WEST, M. (2019). Dynamic Bayesian predictive synthesis in time series forecasting. *J. Econometrics* **210** 155–169. MR3944768 https://doi.org/10.1016/j.jeconom.2018.11.010

MEEKER, W. Q., HAHN, G. J. and ESCOBAR, L. A. (2017). *Statistical Intervals: A Guide for Practitioners and Researchers*, Second Edition ed. Wiley, New York.

MUKHERJEE, G. and JOHNSTONE, I. M. (2015). Exact minimax estimation of the predictive density in sparse Gaussian models. *Ann. Statist.* **43** 937–961. MR3346693 https://doi.org/10.1214/14-AOS1251

NÁJERA, E. and O'REILLY, F. (2017). On fiducial generators. *Comm. Statist. Theory Methods* **46** 2232–2248. MR3576710 https://doi.org/10.1080/03610926.2015.1040505

NELSON, W. (1982). *Applied Life Data Analysis*. *Wiley Series in Probability and Mathematical Statistics*. Wiley, New York. MR0646615

PAN, L. and POLITIS, D. N. (2016). Bootstrap prediction intervals for linear, nonlinear and nonparametric autoregressions. *J. Statist. Plann. Inference* **177** 1–27. MR3511318 https://doi.org/10.1016/j.jspi.2014.10.003

PATEL, J. K. (1989). Prediction intervals—a review. *Comm. Statist. Theory Methods* **18** 2393–2465. MR1029127 https://doi.org/10.1080/03610928908830043

PEERS, H. W. (1965). On confidence points and Bayesian probability points in the case of several parameters. *J. Roy. Statist. Soc. Ser. B* **27** 9–16. MR0191029

POLITIS, D. N. (2003). The impact of bootstrap methods on time series analysis *Statist. Sci.* **18** 219–230. MR2026081 https://doi.org/10.1214/ss/1063994977

POLITIS, D. N. (2013). Model-free model-fitting and predictive distributions. *TEST* **22** 183–221. MR3062250 https://doi.org/10.1007/s11749-013-0317-7

PROSCHAN, F. (1953). Confidence and tolerance intervals for the normal distribution. *J. Amer. Statist. Assoc.* **48** 550–564. MR0057506

ROMANO, Y., PATTERSON, E. and CANDES, E. (2019). Conformalized quantile regression. In *Advances in Neural Information Processing Systems* (H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox and R. Garnett, eds.) **32**. Curran Associates, Red Hook.

SHAFER, G. and VOVK, V. (2008). A tutorial on conformal prediction. *J. Mach. Learn. Res.* **9** 371–421. MR2417240

SHEN, J., LIU, R. Y. and XIE, M. (2018). Prediction with confidence—a general framework for predictive inference. *J. Statist. Plann. Inference* **195** 126–140. MR3760843 https://doi.org/10.1016/j.jspi.2017.09.012

SHMUELI, G. (2010). To explain or to predict? *Statist. Sci.* **25** 289–310. MR2791669 https://doi.org/10.1214/10-STS330

SJÖSTEDT-DE LUNA, S. and YOUNG, A. (2003). The bootstrap and kriging prediction intervals. *Scand. J. Stat.* **30** 175–192. MR1965101 https://doi.org/10.1111/1467-9469.00325

STIGLER, S. M. (1986). Laplace's 1774 memoir on inverse probability. *Statist. Sci.* **1** 359–378. MR0858515

THATCHER, A. R. (1964). Relationships between Bayesian and confidence limits for predictions. *J. Roy. Statist. Soc. Ser. B* **26** 176–192 (discussion: 192–210). MR0175251

TIAN, Q., NORDMAN, D. J. and MEEKER, W. Q. (2022). Supplement to "Methods to compute prediction intervals: A review and new results." https://doi.org/10.1214/21-STS842SUPP

TIAN, Q., MENG, F., NORDMAN, D. and MEEKER, W. (2020). Predicting the number of future events. *J. Amer. Statist. Assoc.* https://doi.org/10.1080/01621459.2020.1850461

VIDONI, P. (1998). A note on modified estimative prediction limits and distributions. *Biometrika* **85** 949–953.

VIDONI, P. (2004). Improved prediction intervals for stochastic process models. *J. Time Series Anal.* **25** 137–154. MR2042115 https://doi.org/10.1111/j.1467-9892.2004.00341.x

VOVK, V., GAMMERMAN, A. and SHAFER, G. (2005). *Algorithmic Learning in a Random World*. Springer, New York. MR2161220

WANG, H. (2008). Coverage probability of prediction intervals for discrete random variables. *Comput. Statist. Data Anal.* **53** 17–26. MR2528589 https://doi.org/10.1016/j.csda.2008.07.017

WANG, H. (2010). Closed form prediction intervals applied for disease counts. *Amer. Statist.* **64** 250–256. MR2757171 https://doi.org/10.1198/tast.2010.09125

WANG, C. M., HANNIG, J. and IYER, H. K. (2012). Fiducial prediction intervals. *J. Statist. Plann. Inference* **142** 1980–1990. MR2903406 https://doi.org/10.1016/j.jspi.2012.02.021

WEST, M. and HARRISON, J. (1997). *Bayesian Forecasting and Dynamic Models*, 2nd ed. *Springer Series in Statistics*. Springer, New York. MR1482232

WILKS, S. S. (1941). Determination of sample sizes for setting tolerance limits. *Ann. Math. Stat.* **12** 91–96. MR0004451 https://doi.org/10.1214/aoms/1177731788

WILSON, E. B. (1927). Probable inference, the law of succession, and statistical inference. *J. Amer. Statist. Assoc.* **22** 209–212.

XIE, M. and SINGH, K. (2013). Confidence distribution, the frequentist distribution estimator of a parameter: A review. *Int. Stat. Rev.* **81** 3–39. MR3047496 https://doi.org/10.1111/insr.12000

ZHANG, H., ZIMMERMAN, J., NETTLETON, D. and NORDMAN, D. J. (2020). Random Forest Prediction Intervals. *Amer. Statist.* **74** 392–406. MR4168259 https://doi.org/10.1080/00031305.2019.1585288

Statistical Science
2022, Vol. 37, No. 4, 598–606
https://doi.org/10.1214/21-STS837

Approximate Confidence Intervals for a Binomial p—Once Again

Per Gösta Andersson

Abstract. The problem of constructing a reasonably simple yet well-behaved confidence interval for a binomial parameter p is old but still fascinating and surprisingly complex. During the last century, many alternatives to the poorly behaved standard Wald interval have been suggested. It seems though that the Wald interval is still much in use in spite of many efforts over the years through publications to point out its deficiencies. This paper constitutes yet another attempt to provide an alternative and it builds on a special case of a general technique for adjusted intervals primarily based on Wald type statistics. The main idea is to construct an approximate pivot with uncorrelated, or nearly uncorrelated, components. The resulting AN (Andersson–Nerman) interval, as well as a modification thereof, is compared with the well-renowned Wilson and AC (Agresti–Coull) intervals and the subsequent discussion will in itself hopefully shed some new light on this seemingly elementary interval estimation situation. Generally, an alternative to the Wald interval is to be judged not only by performance, its expression should also indicate *why* we will obtain a better behaved interval. It is argued that the well-behaved AN interval meets this requirement.

Key words and phrases: Correlation, coverage probability, score statistic, skewness, Wald statistic.

REFERENCES

AGRESTI, A. and CAFFO, B. (2000). Simple and effective confidence intervals for proportions and differences of proportions result from adding two successes and two failures. *Amer. Statist.* **54** 280–288. MR1814845 https://doi.org/10.2307/2685779

AGRESTI, A. and COULL, B. A. (1998). Approximate is better than "exact" for interval estimation of binomial proportions. *Amer. Statist.* **52** 119–126. MR1628435 https://doi.org/10.2307/2685469

ANDERSSON, P. G. (2009). A simple correlation adjustment procedure applied to confidence interval construction. *Amer. Statist.* **63** 258–262. MR2750351 https://doi.org/10.1198/tast.2009.08045

ANDERSSON, P. G. and NERMAN, O. (2000). A balanced adjusted confidence interval procedure applied to finite population sampling. In *Proc. of ICES II*, Buffalo, New York, USA (CD-ROM) Amer. Statist. Assoc., Alexandria, Virginia, USA.

BICKEL, P. J. and DOKSUM, K. A. (1977). *Mathematical Statistics: Basic Ideas and Selected Topics. Holden-Day Series in Probability and Statistics*. Holden-Day, San Francisco. MR0443141

BLYTH, C. R. (1986). Approximate binomial confidence limits. *J. Amer. Statist. Assoc.* **81** 843–855. MR0860519

BLYTH, C. R. and STILL, H. A. (1983). Binomial confidence intervals. *J. Amer. Statist. Assoc.* **78** 108–116. MR0696854

BROWN, L. D., CAI, T. T. and DASGUPTA, A. (2001). Interval estimation for a binomial proportion. *Statist. Sci.* **16** 101–133. With comments and a rejoinder by the authors. MR1861069 https://doi.org/10.1214/ss/1009213286

BROWN, L. D., CAI, T. T. and DASGUPTA, A. (2002). Confidence intervals for a binomial proportion and asymptotic expansions. *Ann. Statist.* **30** 160–201. MR1892660 https://doi.org/10.1214/aos/1015362189

CAI, T. T. (2005). One-sided confidence intervals in discrete distributions. *J. Statist. Plann. Inference* **131** 63–88. MR2136006 https://doi.org/10.1016/j.jspi.2004.01.005

CLOPPER, C. J. and PEARSON, E. S. (1934). The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika* **26** 404–413.

CROW, E. L. (1956). Confidence intervals for a proportion. *Biometrika* **43** 423–435. MR0093077 https://doi.org/10.1093/biomet/43.3-4.423

DASGUPTA, A. (2008). *Asymptotic Theory of Statistics and Probability. Springer Texts in Statistics*. Springer, New York. MR2664452

DEAN, N. and PAGANO, M. (2015). Evaluating confidence interval methods for binomial proportions in clustered surveys. *J. Surv. Statist. Meth.* **3** 484–503.

EASTERLING, R. (2015). There's nothing wrong with Clopper–Pearson binomial confidence limits. *Amer. Statist.* **69** 154–155.

GHOSH, B. K. (1979). A comparison of some approximate confidence intervals for the binomial parameter. *J. Amer. Statist. Assoc.* **74** 894–900. MR0556485

JIN, S., THULIN, M. and LARSSON, R. (2017). Approximate Bayesianity of frequentist confidence intervals for a binomial pro-

Per Gösta Andersson is Associate Professor, Department of Statistics, Stockholm University, Stockholm, Sweden (e-mail: per.gosta.andersson@stat.su.se).

portion. *Amer. Statist.* **71** 106–111. MR3668695 https://doi.org/10.1080/00031305.2016.1208630

JOHNSON, N. L., KEMP, A. W. and KOTZ, S. (2005). *Univariate Discrete Distributions*, 3rd ed. *Wiley Series in Probability and Statistics*. Wiley Interscience, Hoboken, NJ. MR2163227 https://doi.org/10.1002/0471715816

KOTT, P. S. (2017). A note on Wilson coverage intervals for proportions estimated from survey samples. *Surv. Methodol.* **43** 234–240.

KOTT, P. S., ANDERSSON, P. G. and NERMAN, O. (2001). Two-sided coverage intervals for small proportions based on survey data. In *Proc. FCSM Res. Conf.*, Arlington, VA, USA.

LAPLACE, P. S. (1812). *Théorie Analytique des Probabilités*. Courcier, Paris.

LOUIS, T. A. (1981). Confidence intervals for a binomial parameter after observing no successes. *Amer. Statist.* **35** 154.

NEWCOMBE, R. G. (1998). Two-sided confidence intervals for the single proportion: Comparison of seven methods. *Stat. Med.* **17** 857–872.

SANTNER, T. J. (1998). Teaching large-sample binomial confidence intervals. *Teach. Stat.* **20** 20–23.

SCHILLING, M. F. and DOI, J. A. (2014). A coverage probability approach to finding an optimal binomial confidence procedure. *Amer. Statist.* **68** 133–145. MR3246552 https://doi.org/10.1080/00031305.2014.899274

WILSON, E. B. (1927). Probable inference, the law of succession, and statistical inference. *J. Amer. Statist. Assoc.* **22** 209–212.

Statistical Science
2022, Vol. 37, No. 4, 607–624
https://doi.org/10.1214/22-STS849

A Conversation with David J. Aldous

Shankar Bhamidi

Abstract. David John Aldous was born in Exeter U.K. on July 13, 1952. He received a B.A. and Ph.D. in Mathematics in 1973 and 1977, respectively from Cambridge. After spending two years as a research fellow at St. John's College, Cambridge, he joined the Department of Statistics at the University of California, Berkeley in 1979 where he spent the rest of his academic career until retiring in 2018. He is known for seminal contributions on many topics within probability including weak convergence and tightness, exchangeability, Markov chain mixing times, Poisson clumping heuristic and limit theory for large discrete random structures including random trees, stochastic coagulation and fragmentation systems, models of complex networks and interacting particle systems on such structures. For his contributions to the field, he has received numerous honors and awards including the Rollo Davidson prize in 1980, the inaugural Loeve prize in Probability in 1993, and the Brouwer medal in 2021, and being elected as an IMS fellow in 1985, Fellow of the Royal Society in 1994, Fellow of the American Academy of Arts and Sciences in 2004, elected to the National Academy of Sciences (foreign associate) in 2010, ICM plenary speaker in 2010 and AMS fellow in 2012.

Key words and phrases: Exchangeability, Markov chain mixing times, scaling limits, local weak convergence, random graphs, network models.

REFERENCES

ALDOUS, D. J. (1977). Limit theorems for subsequences of arbitrarily-dependent sequences of random variables. *Z. Wahrsch. Verw. Gebiete* **40** 59–82. MR0455090 https://doi.org/10.1007/BF00535707

ALDOUS, D. (1978). Stopping times and tightness. *Ann. Probab.* **6** 335–340. MR0474446 https://doi.org/10.1214/aop/1176995579

ALDOUS, D. J. (1981a). Subspaces of L^1, via random measures. *Trans. Amer. Math. Soc.* **267** 445–463. MR0626483 https://doi.org/10.2307/1998664

ALDOUS, D. J. (1981b). Representations for partially exchangeable arrays of random variables. *J. Multivariate Anal.* **11** 581–598. MR0637937 https://doi.org/10.1016/0047-259X(81)90099-3

ALDOUS, D. (1983). Random walks on finite groups and rapidly mixing Markov chains. In *Seminar on Probability, XVII. Lecture Notes in Math.* **986** 243–297. Springer, Berlin. MR0770418 https://doi.org/10.1007/BFb0068322

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

ALDOUS, D. (1989). *Probability Approximations via the Poisson Clumping Heuristic. Applied Mathematical Sciences* **77**. Springer, New York. MR0969362 https://doi.org/10.1007/978-1-4757-6283-9

ALDOUS, D. J. (1990). The random walk construction of uniform spanning trees and uniform labelled trees. *SIAM J. Discrete Math.* **3** 450–465. MR1069105 https://doi.org/10.1137/0403039

ALDOUS, D. (1991a). The continuum random tree. I. *Ann. Probab.* **19** 1–28. MR1085326

ALDOUS, D. (1991b). The continuum random tree. II. An overview. In *Stochastic Analysis* (*Durham*, 1990). *London Mathematical Society Lecture Note Series* **167** 23–70. Cambridge Univ. Press, Cambridge. MR1166406 https://doi.org/10.1017/CBO9780511662980.003

ALDOUS, D. (1992). Asymptotics in the random assignment problem. *Probab. Theory Related Fields* **93** 507–534. MR1183889 https://doi.org/10.1007/BF01192719

ALDOUS, D. (1993). The continuum random tree. III. *Ann. Probab.* **21** 248–289. MR1207226

ALDOUS, D. (1997). Brownian excursions, critical random graphs and the multiplicative coalescent. *Ann. Probab.* **25** 812–854. MR1434128 https://doi.org/10.1214/aop/1024404421

ALDOUS, D. J. (1999). Deterministic and stochastic models for coalescence (aggregation and coagulation): A review of the mean-field theory for probabilists. *Bernoulli* **5** 3–48. MR1673235 https://doi.org/10.2307/3318611

ALDOUS, D. J. (2001). The $\zeta(2)$ limit in the random assignment problem. *Random Structures Algorithms* **18** 381–418. MR1839499 https://doi.org/10.1002/rsa.1015

ALDOUS, D. J. (2004). A tractable complex network model based on the stochastic mean-field model of distance. In *Complex Networks. Lecture Notes in Physics* **650** 51–87. Springer, Berlin. MR2108972 https://doi.org/10.1007/978-3-540-44485-5_3

ALDOUS, D. (2013a). Another conversation with Persi Diaconis. *Statist. Sci.* **28** 269–281. MR3112410 https://doi.org/10.1214/12-sts404

Shankar Bhamidi is Professor, Department of Statistics and Operations Research, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA (e-mail: bhamidi@email.unc.edu).

ALDOUS, D. J. (2013b). Using prediction market data to illustrate undergraduate probability. *Amer. Math. Monthly* **120** 583–593. MR3096465 https://doi.org/10.4169/amer.math.monthly.120.07.583

ALDOUS, D. (2016a). The incipient giant component in bond percolation on general finite weighted graphs. *Electron. Commun. Probab.* **21** Paper No. 68, 9 pp. MR3564215 https://doi.org/10.1214/16-ECP21

ALDOUS, D. J. (2016b). Weak concentration for first passage percolation times on graphs and general increasing set-valued processes. *ALEA Lat. Am. J. Probab. Math. Stat.* **13** 925–940. MR3550985

ALDOUS, D. (2017). Waves in a spatial queue. *Stoch. Syst.* **7** 197–236. MR3663341 https://doi.org/10.1287/15-SSY208

ALDOUS, D. J. (2021). A prediction tournament paradox. *Amer. Statist.* **75** 243–248. MR4298580 https://doi.org/10.1080/00031305.2019.1604430

ALDOUS, D. and BARTHELEMY, M. (2019). Optimal geometry of transportation networks. *Phys. Rev. E* **99** 052303. https://doi.org/10.1103/PhysRevE.99.052303

ALDOUS, D. J. and BHAMIDI, S. (2010). Edge flows in the complete random-lengths network. *Random Structures Algorithms* **37** 271–311. MR2724663 https://doi.org/10.1002/rsa.20306

ALDOUS, D. and DIACONIS, P. (1986). Shuffling cards and stopping times. *Amer. Math. Monthly* **93** 333–348. MR0841111 https://doi.org/10.2307/2323590

ALDOUS, D. and DIACONIS, P. (1987). Strong uniform times and finite random walks. *Adv. in Appl. Math.* **8** 69–97. MR0876954 https://doi.org/10.1016/0196-8858(87)90006-6

ALDOUS, D. and DIACONIS, P. (1995). Hammersley's interacting particle process and longest increasing subsequences. *Probab. Theory Related Fields* **103** 199–213. MR1355056 https://doi.org/10.1007/BF01204214

ALDOUS, D. and DIACONIS, P. (1999). Longest increasing subsequences: From patience sorting to the Baik–Deift–Johansson theorem. *Bull. Amer. Math. Soc.* (*N.S.*) **36** 413–432. MR1694204 https://doi.org/10.1090/S0273-0979-99-00796-X

ALDOUS, D. and FILL, J. (2002). Reversible Markov chains and random walks on graphs.

ALDOUS, D. and GANESAN, K. (2013). True scale-invariant random spatial networks. *Proc. Natl. Acad. Sci. USA* **110** 8782–8785. MR3082274 https://doi.org/10.1073/pnas.1304329110

ALDOUS, D. J. and KENDALL, W. S. (2008). Short-length routes in low-cost networks via Poisson line patterns. *Adv. in Appl. Probab.* **40** 1–21. MR2411811 https://doi.org/10.1239/aap/1208358883

ALDOUS, D. and LIMIC, V. (1998). The entrance boundary of the multiplicative coalescent. *Electron. J. Probab.* **3** no. 3, 59 pp. MR1491528 https://doi.org/10.1214/EJP.v3-25

ALDOUS, D. and LYONS, R. (2007). Processes on unimodular random networks. *Electron. J. Probab.* **12** 1454–1508. MR2354165 https://doi.org/10.1214/EJP.v12-463

ALDOUS, D., MIERMONT, G. and PITMAN, J. (2005). Weak convergence of random p-mappings and the exploration process of inhomogeneous continuum random trees. *Probab. Theory Related Fields* **133** 1–17. MR2197134 https://doi.org/10.1007/s00440-004-0407-2

ALDOUS, D. and PITMAN, J. (1998). The standard additive coalescent. *Ann. Probab.* **26** 1703–1726. MR1675063 https://doi.org/10.1214/aop/1022855879

ALDOUS, D. and POPOVIC, L. (2005). A critical branching process model for biodiversity. *Adv. in Appl. Probab.* **37** 1094–1115. MR2193998 https://doi.org/10.1239/aap/1134587755

ALDOUS, D. J. and SHUN, J. (2010). Connected spatial networks over random points and a route-length statistic. *Statist. Sci.* **25** 275–288. MR2791668 https://doi.org/10.1214/10-STS335

ALDOUS, D. and STEELE, J. M. (2004). The objective method: Probabilistic combinatorial optimization and local weak convergence. In *Probability on Discrete Structures*. *Encyclopaedia Math. Sci.* **110** 1–72. Springer, Berlin. MR2023650 https://doi.org/10.1007/978-3-662-09444-0_1

AUSTIN, T. (2012). Exchangeable random arrays. In *Notes for IAS Workshop*.

BAIK, J., DEIFT, P. and JOHANSSON, K. (1999). On the distribution of the length of the longest increasing subsequence of random permutations. *J. Amer. Math. Soc.* **12** 1119–1178. MR1682248 https://doi.org/10.1090/S0894-0347-99-00307-0

BAILEY, N. T. J. (1957). *The Mathematical Theory of Epidemics*. Hafner Publishing Co., New York. MR0095085

BARBOUR, A. D., HOLST, L. and JANSON, S. (1992). *Poisson Approximation*. *Oxford Studies in Probability* **2**. The Clarendon Press, Oxford Univ. Press, New York. MR1163825

BARTHÉLEMY, M. (2011). Spatial networks. *Phys. Rep.* **499** 1–101. MR2770962 https://doi.org/10.1016/j.physrep.2010.11.002

BARTHELEMY, M. (2018). *Morphogenesis of Spatial Networks*. *Lecture Notes in Morphogenesis*. Springer, Cham. MR3752147 https://doi.org/10.1007/978-3-319-20565-6

BAYER, D. and DIACONIS, P. (1992). Trailing the dovetail shuffle to its lair. *Ann. Appl. Probab.* **2** 294–313. MR1161056

BENJAMINI, I. and SCHRAMM, O. (2011). Recurrence of distributional limits of finite planar graphs. In *Selected Works of Oded Schramm*. *Volumes* 1, 2. *Sel. Works Probab. Stat.* 533–545. Springer, New York. MR2883381 https://doi.org/10.1007/978-1-4419-9675-6_15

BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. *Wiley Series in Probability and Statistics*: *Probability and Statistics*. Wiley, New York. MR1700749 https://doi.org/10.1002/9780470316962

BREIMAN, L. (1968). *Probability*. Addison-Wesley Co., Reading, MA. MR0229267

CHUNG, K. L. (1968). *A Course in Probability Theory*. Harcourt, Brace & World, Inc., New York. MR0229268

DAWID, A. P. (1981). Some matrix-variate distribution theory: Notational considerations and a Bayesian application. *Biometrika* **68** 265–274. MR0614963 https://doi.org/10.1093/biomet/68.1.265

DIACONIS, P. (1988). *Group Representations in Probability and Statistics*. *Institute of Mathematical Statistics Lecture Notes—Monograph Series* **11**. IMS, Hayward, CA. MR0964069

DIACONIS, P. and JANSON, S. (2008). Graph limits and exchangeable random graphs. *Rend. Mat. Appl.* (7) **28** 33–61. MR2463439

DIACONIS, P. and SHAHSHAHANI, M. (1981). Generating a random permutation with random transpositions. *Z. Wahrsch. Verw. Gebiete* **57** 159–179. MR626813 https://doi.org/10.1007/BF00535487

DRAKE, R. L. (1972). A General Mathematics Survey of the Coagulation Equation. In *Topics in Current Aerosol Research* Pergamon Press, Elmsford.

FREEDMAN, D. (2012). *Markov Chains*. Holden-Day, San Francisco, CA. MR0292176

HAMMERSLEY, J. M. (1972). A few seedlings of research. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability* (*Univ. California*, *Berkeley*, *Calif.*, 1970/1971), *Vol. I*: *Theory of Statistics* 345–394. Univ. California Press, Berkeley, CA. MR0405665

JANSON, S., KNUTH, D. E., ŁUCZAK, T. and PITTEL, B. (1993). The birth of the giant component. *Random Structures Algorithms* **4** 231–358. MR1220220 https://doi.org/10.1002/rsa.3240040303

KEMENY, J. G. and SNELL, J. L. (1960). *Finite Markov Chains*. *The University Series in Undergraduate Mathematics*. D. Van Nostrand Co., Inc., Princeton, NJ. MR0115196

KREPS, D. M. (1981). Arbitrage and equilibrium in economies with infinitely many commodities. *J. Math. Econom.* **8** 15–35. MR0611252 https://doi.org/10.1016/0304-4068(81)90010-0

KUCHARSKI, A. (2020). *The Rules of Contagion*: *Why Things Spread—And Why They Stop*. Hachette UK.

LE GALL, J.-F. (1993). The uniform random tree in a Brownian excursion. *Probab. Theory Related Fields* **96** 369–383. MR1231930 https://doi.org/10.1007/BF01292678

LEVIN, D. A. and PERES, Y. (2017). *Markov Chains and Mixing Times*. Amer. Math. Soc., Providence, RI. MR3726904 https://doi.org/10.1090/mbk/107

LOVÁSZ, L. and SZEGEDY, B. (2006). Limits of dense graph sequences. *J. Combin. Theory Ser. B* **96** 933–957. MR2274085 https://doi.org/10.1016/j.jctb.2006.05.002

MÉZARD, M. and PARISI, G. (1987). On the solution of the random link matching problems. *J. Phys.* **48** 1451–1459.

PITMAN, J. (2006). *Combinatorial Stochastic Processes*: *Ecole d'Eté de Probabilités de Saint-Flour XXXII—2002*. *Lecture Notes in Math*. **1875**. Springer, Berlin. MR2245368

ROOT, D. H. (1969). The existence of certain stopping times on Brownian motion. *Ann. Math. Stat.* **40** 715–718. MR0238394 https://doi.org/10.1214/aoms/1177697749

SMOLUCHOWSKI, M. V. (1916). Drei vortrage uber diffusion, brownsche bewegung und koagulation von kolloidteilchen. *Z. Phys.* **17** 557–585.

WALI, K. C. (1991). *Chandra*: *A Biography of S. Chandrasekhar*. Univ. Chicago Press, Chicago, IL. MR1086635 https://doi.org/10.1007/978-1-4899-3472-7_7

Statistical Science
2022, Vol. 37, No. 4, 625–627
https://doi.org/10.1214/22-STS862

Comments on *Confidence as Likelihood* by Pawitan and Lee in *Statistical Science*, November 2021

Michael Lavine and Jan F. Bjørnstad

Abstract. Pawitan and Lee (*Statist. Sci.* **36** (2021) 509–517) attempt to show a correspondence between confidence and likelihood, specifically, that "confidence is in fact an extended likelihood" (*Statist. Sci.* **36** (2021) 509–517, abstract). The word "extended" means that the likelihood function can accommodate unobserved random variables such as random effects and future values; see (*J. Amer. Statist. Assoc.* **91** (1996) 791–806) for details. Here, we argue that the extended likelihood presented by (*Statist. Sci.* **36** (2021) 509–517) is not the correct extended likelihood and does not justify interpreting confidence as likelihood.

Key words and phrases: Likelihood, confidence.

REFERENCES

BJØRNSTAD, J. F. (1996). On the generalization of the likelihood function and the likelihood principle. *J. Amer. Statist. Assoc.* **91** 791–806. MR1395746 https://doi.org/10.2307/2291674

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

PAWITAN, Y. and LEE, Y. (2021). Confidence as likelihood. *Statist. Sci.* **36** 509–517. MR4323049 https://doi.org/10.1214/20-sts811

Michael Lavine is Professor Emeritus, Department of Mathematics and Statistics, University of Massachusetts Amherst, Amherst, Massachusetts 01003, USA (e-mail: lavine@math.umass.edu). Jan F. Bjørnstad is Professor Emeritus, University of Oslo and Statistics Norway (e-mail: Jan.Bjornstad@ssb.no).

Statistical Science
2022, Vol. 37, No. 4, 628–629
https://doi.org/10.1214/22-STS869

Rejoinder: Confidence as Likelihood

Yudi Pawitan and Youngjo Lee

REFERENCES

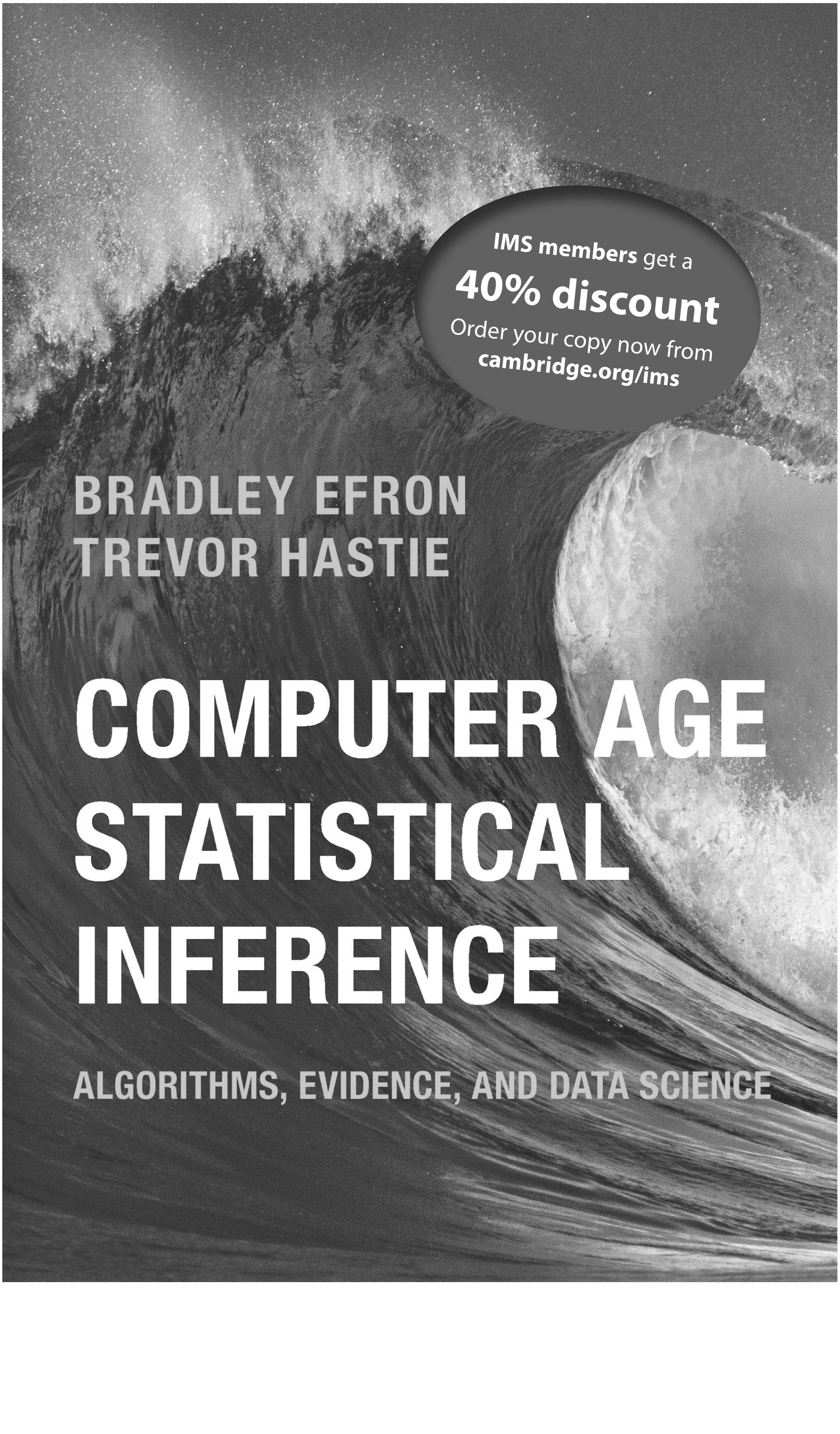
LEE, Y., NELDER, J. A. and PAWITAN, Y. (2017). *Generalized Linear Models with Random Effects*: *Unified analysis via H-likelihood*, 2nd ed. *Monographs on Statistics and Applied Probability* **153**. CRC Press, Boca Raton, FL. MR3726914

Yudi Pawitan is Professor, Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden (e-mail: yudi.pawitan@ki.se). Youngjo Lee is Professor Emeritus, Department Statistics, Seoul National University, Seoul, Korea (e-mail: youngjo@snu.ac.kr).

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 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