STATISTICAL SCIENCE

Volume 35, Number 2 May 2020

Statistical Science **[ISSN 0883-4237 (print); ISSN 2168-8745 (online)]**, Volume 35, Number 2, May 2020. Published quarterly by the Institute of Mathematical Statistics, 3163 Somerset Drive, Cleveland, OH 44122, 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, 9650 Rockville Pike—Suite L2310, Bethesda, MD 20814-3998, USA.

Statistical Science **Volume 35, Number 2 (159–333) May 2020**

Volume 35 **Number 2** **May 2020**

Statistical Science
2020, Vol. 35, No. 2, 159–170
https://doi.org/10.1214/19-STS696

Laplace's Theories of Cognitive Illusions, Heuristics and Biases

Joshua B. Miller and Andrew Gelman

Abstract. In his book from the early 1800s, *Essai Philosophique sur les Probabilités*, the mathematician Pierre-Simon de Laplace anticipated many ideas developed within the past 50 years in cognitive psychology and behavioral economics, explaining human tendencies to deviate from norms of rationality in the presence of probability and uncertainty. A look at Laplace's theories and reasoning is striking, both in how modern they seem, how much progress he made without the benefit of systematic experimentation, and the novelty of a few of his unexplored conjectures. We argue that this work points to these theories being more fundamental and less contingent on recent experimental findings than we might have thought.

REFERENCES

ASHRAF, N., CAMERER, C. and LOEWENSTEIN, G. (2005). Adam Smith, behavioral economist. *J. Econ. Perspect.* **19** 131–145.

AYTON, P. and FISCHER, I. (2004). The hot hand fallacy and the gambler's fallacy: Two faces of subjective randomness? *Memory & Cognition* **21** 1369–1378.

BAR-HILLEL, M. (1980). The base-rate fallacy in probability judgments. *Acta Psychol.* **44** 211–233.

BAR-HILLEL, M. and WAGENAAR, W. A. (1991). The perception of randomness. *Adv. in Appl. Math.* **12** 428–454.

BARON, J. (2008). *Thinking and Deciding*. Cambridge Univ. Press, Cambridge.

BENJAMIN, D. J. (forthcoming). Errors in probabilistic reasoning and judgment biases. In *Handbook of Behavioral Economics* (D. Bernheim, S. DellaVigna and D. Laibson, eds.) Elsevier, New York.

BENJAMIN, D. J., RABIN, M. and RAYMOND, C. (2016). A model of nonbelief in the law of large numbers. *J. Eur. Econ. Assoc.* **14** 515–544.

BOWERS, J. S. and DAVIS, C. J. (2012). Bayesian just-so stories in psychology and neuroscience. *Psychol. Bull.* **138** 389–414.

CAMERER, C. F., LOWENSTEIN, G. and RABIN, M. (2003). *Advances in Behavioral Economics*. Princeton Univ. Press, Princeton, NJ.

CLOTFELTER, C. T. and COOK, P. J. (1993). The "gambler's fallacy" in lottery play. *Manage. Sci.* **39** 1521–1525.

COX, R. T. (1946). Probability, frequency and reasonable expectation. *Am. J. Phys.* **14** 1–13. MR0015688 https://doi.org/10.1119/1.1990764

CROSON, R. and SUNDALI, J. (2005). The gambler's fallacy and the hot hand: Empirical data from casinos. *J. Risk Uncertain.* **30** 195–209.

D'ALEMBERT, J. (1761). *Opuscules Mathématiques*. David, Paris.

DASTON, L. (1988). *Classical Probability in the Enlightenment*. Princeton Univ. Press, Princeton, NJ. MR0988886

DEMPSTER, A. P. (1968). A generalization of Bayesian inference. *J. Roy. Statist. Soc. Ser. B* **30** 205–247. MR0238428

EDWARDS, W. (1968). *Conservatism in Human Information Processing*. Wiley, New York.

ELSTER, J. (1989). *Nuts and Bolts for the Social Sciences*. Cambridge Univ. Press, Cambridge.

EPLEY, N. and GILOVICH, T. (2006). The anchoring-and-adjustment heuristic: Why the adjustments are insufficient. *Psychol. Sci.* **17** 311–318.

EREV, I., ERT, E., ROTH, A. E., HARUVY, E., HERZOG, S. M., HAU, R., HERTWIG, R., STEWART, T., WEST, R. et al. (2010). A choice prediction competition: Choices from experience and from description. *J. Behav. Decis. Mak.* **23** 15–47.

FISCHHOFF, B., SLOVIC, P., LICHTENSTEIN, S., READ, S. and COMBS, B. (1978). How safe is safe enough? A psychometric study of attitudes towards technological risks and benefits. *Policy Sci.* **9** 127–152.

GELMAN, A. (2011). Josh Tenenbaum presents ...a model of folk physics! Available at https://andrewgelman.com/2011/11/06/josh-tenenbaum-presents-a-model-of-folk-physics/.

GIGERENZER, G. (2004). Fast and frugal heuristics: The tools of bounded rationality. In *Blackwell Handbook of Judgment and Decision Making* (D. J. Koehler and N. Harvey, eds.) 62–88. Blackwell, London.

GIGERENZER, G. (2005). I think, therefore I err. *Soc. Res.* **72** 195–208.

GIGERENZER, G. and BRIGHTON, H. (2009). Homo heuristicus: Why biased minds make better inferences. *Top. Cogn. Sci.* **1** 107–143.

GIGERENZER, G. and GOLDSTEIN, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychol. Rev.* **103** 650.

GIGERENZER, G. and HOFFRAGE, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. *Psychol. Rev.* **102** 684–704.

GIGERENZER, G. and SELTEN, R. (2001). *Bounded Rationality*: *The Adaptive Toolbox*. MIT Press, Cambridge, MA.

GILOVICH, T., GRIFFIN, D. and KAHNEMAN, D. (2002). *Heuristics and Biases*: *The Psychology of Intuitive Judgment*. Cambridge Univ. Press, Cambridge.

Joshua B. Miller is an Associate Professor in the Department of Economics at the University of Melbourne, Australia (e-mail: joshua.benjamin.miller@gmail.com). Andrew Gelman is a Professor in the Department of Statistics and Department of Political Science at Columbia University, New York, New York, USA (e-mail: gelman@stat.columbia.edu).

GILOVICH, T., VALLONE, R. and TVERSKY, A. (1985). The hot hand in basketball: On the misperception of random sequences. *Cogn. Psychol.* **17** 295–314.

GORROOCHURN, P. (2012). *Classic Problems of Probability*. Wiley, Hoboken, NJ. MR2976816 https://doi.org/10.1002/9781118314340

HATFIELD, G. (1994). Psychology as a natural science in the eighteenth century. *Revue de Synthèse* **115** 375–391.

HERTWIG, R. and EREV, I. (2009). The description-experience gap in risky choice. *Trends Cogn. Sci.* **13** 517–523.

HERTWIG, R., GIGERENZER, G. and HOFFRAGE, U. (1997). The reiteration effect in hindsight bias. *Psychol. Rev.* **104** 194.

HERTWIG, R., PACHUR, T. and KURZENHÄUSER, S. (2005). Judgments of risk frequencies: Tests of possible cognitive mechanisms. *J. Exper. Psychol., Learn., Mem., Cogn.* **31** 621.

HERTWIG, R., BARRON, G., WEBER, E. U. and EREV, I. (2004). Decisions from experience and the effect of rare events in risky choice. *Psychol. Sci.* **15** 534–539.

IMAS, A. (2016). The realization effect: Risk-taking after realized versus paper losses. *Am. Econ. Rev.* **106** 2086–2109.

JAMES, W. (1890). *The Principles of Psychology*. Henry Holt and Company, New York.

JAYNES, E. T. (1957). Information theory and statistical mechanics. *Phys. Rev.* **106** 620–630. MR0087305

KAHNEMAN, D. (1991). Judgment and decision making: A personal view. *Psychol. Sci.* **2** 142–145.

KAHNEMAN, D. (2011). *Thinking, Fast and Slow*. Farrar, Straus and Giroux, New York.

KAHNEMAN, D. and KLEIN, G. (2009). Conditions for intuitive expertise: A failure to disagree. *Am. Psychol.* **64** 515–526.

KAHNEMAN, D., SLOVIC, P. and TVERSKY, A. (1982). *Judgment Under Uncertainty: Heuristics and Biases*. Cambridge Univ. Press, Cambridge.

KAHNEMAN, D. and TVERSKY, A. (1972a). On prediction and judgment. *ORI Research Monograph* **12**. (10) (a).

KAHNEMAN, D. and TVERSKY, A. (1972b). Subjective probability: A judgement of representativeness. *Cogn. Psychol.* **3** 430–454.

KAHNEMAN, D. and TVERSKY, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica* **47** 263–292.

KAHNEMAN, D. and TVERSKY, A. (1982). On the study of statistical intuitions. *Cognition* **11** 123–141.

KAMENICA, E. (2012). Behavioral economics and psychology of incentives. *Ann. Rev. Econ.* **4** 427–452.

KEYNES, J. M. (1921). *A Treatise on Probability*. Macmillan and Company, Limited, London.

LAPLACE, P.-S. (1825). *Essai Philosophique sur les Probabilités, fifth edition*. Springer, New York. Translated by A. I. Dale (1995).

LOEWENSTEIN, G. F., WEBER, E. U., HSEE, C. K. and WELCH, N. (2001). Risk as feelings. *Psychol. Bull.* **127** 267–286.

MARCUS, G. F. and DAVIS, E. (2013). How robust are probabilistic models of higher-level cognition? *Psychol. Sci.* **24** 2351–2360.

MEEHL, P. E. (1954). *Clinical Versus Statistical Predictions: A Theoretical Analysis and Revision of the Literature*. Univ. Minnesota Press, Minneapolis, MN.

MEEHL, P. E. and ROSEN, A. (1955). Antecedent probability and the efficiency of psychometric signs, patterns, or cutting scores. *Psychol. Bull.* **52** 194–216.

MILLER, J. B. and MANIADIS, Z. (2013). The weight of personal experience: An experimental measurement. Working paper.

MILLER, J. B. and SANJURJO, A. (2018). A Visible (Hot) Hand? Expert Players Bet on the Hot Hand and Win. OSF Preprints. October 30 https://osf.io/sd32u/.

MILLER, J. B. and SANJURJO, A. (2018). Surprised by the hot hand fallacy? A truth in the law of small numbers. *Econometrica* **86** 2019–2047. MR3901245 https://doi.org/10.3982/ECTA14943

MORRISON, R. S. and ORDESHOOK, P. C. (1975). Rational choice, light guessing and the gambler's fallacy. *Public Choice* **22** 79–89.

NICKERSON, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Rev. Gen. Psychol.* **2** 175.

NICKERSON, R. S. (2002). The production and perception of randomness. *Psychol. Rev.* **109** 350–357.

OAKSFORD, M. and CHATER, N. (2009). Precis of Bayesian rationality: The probabilistic approach to human reasoning. *Behav. Brain Sci.* **32** 69–84.

ODEAN, T. (1998). Are investors reluctant to realize their losses? *J. Finance* **53** 1775–1798.

OSKARSSON, A. T., BOVEN, L. V., MCCLELLAND, G. H. and HASTIE, R. (2009). What's next? Judging sequences of binary events. *Psychol. Bull.* **135** 262–385.

PEARSON, K. (1978). *The History of Statistics in the 17th and 18th Centuries Against the Changing Background of Intellectual, Scientific and Religious Thought. Lectures Given at University College London*, 1921–1933. Macmillan Co., New York. MR0612481

RABIN, M. (2002). Inference by believers in the law of small numbers. *Q. J. Econ.* **117** 775–816.

RAMSEY, F. P. (1926). Truth and probability. In *The Foundations of Mathematics and Other Logical Essays* (R. B. Braithwaite, ed.) 156–198. Harcourt, Brace and Company, New York.

SHAFER, G. (1976). *A Mathematical Theory of Evidence*. Princeton Univ. Press, Princeton, NJ. MR0464340

SHEFRIN, H. and STATMAN, M. (1985). The disposition to sell winners too early and ride losers too long: Theory and evidence. *J. Finance* **40** 777–790.

SIMON, H. A. (1956). Rational choice and the structure of the environment. *Psychol. Rev.* **63** 129–138.

SIMON, H. A. (1957). *Models of Man, Social and Rational*. Wiley, New York. MR0113732

SIMONSOHN, U., KARLSSON, N., LOEWENSTEIN, G. and ARIELY, D. (2008). The tree of experience in the forest of information: Overweighing experienced relative to observed information. *Games Econom. Behav.* **62** 263–286. MR2384867 https://doi.org/10.1016/j.geb.2007.03.010

STIGLER, S. (2003). Casanova, "Bonaparte," and the Loterie de France. *Journal de la Société Française de Statistique* **144** 5–34.

STIGLER, S. (2005). P. S. Laplace, Théorie Analytique des Probabilités, First Edition (1812); Essai Philosophique sur les Probabilitiés, first edition (1814),. In *Landmark Writings in Western Mathematics*, 1640–1940 (I. Grattan-Guinness, ed.) 329–340. Elsevier, Amsterdam.

STIGLER, S. M. (2012). Laplace's unpublished manuscript on associationist psychology. *J. Électron. Hist. Probab. Stat.* **8** 6. MR3025117

SUNDALI, J. and CROSON, R. (2006). Biases in casino betting: The hot hand and the gambler's fallacy. *Judgm. Decis. Mak.* **1** 1–12.

TENENBAUM, J. B., KEMP, C., GRIFFITHS, T. L. and GOODMAN, N. D. (2011). How to grow a mind: Statistics, structure, and abstraction. *Science* **331** 1279–1285. MR2807660 https://doi.org/10.1126/science.1192788

TERRELL, D. (1994). A test of the gambler's fallacy: Evidence from pari-mutuel games. *J. Risk Uncertain.* **8** 309–317.

THALER, R. (1980). Toward a positive theory of consumer choice. *J. Econ. Behav. Organ.* **1** 39–60.

TODD, P. M. and GIGERENZER, G. (2000). Précis of simple heuristics that make us smart. *Behav. Brain Sci.* **23**.

TVERSKY, A. and KAHNEMAN, D. (1973). Availability: A heuristic for judging frequency and probability. *Cogn. Psychol.* **5** 207–232.

TVERSKY, A. and KAHNEMAN, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science* **185** 1124–1131.

VON NEUMANN, J. and MORGENSTERN, O. (1944). *Theory of Games and Economic Behavior*. Princeton Univ. Press, Princeton, NJ. MR0011937

WILKE, A. and BARRETT, H. C. (2009). The hot hand phenomenon as a cognitive adaptation to clumped resources. *Evol. Hum. Behav.* **30** 161–169.

Statistical Science
2020, Vol. 35, No. 2, 171–172
https://doi.org/10.1214/19-STS750

Comment: Laplace and Cognitive Illusions

Daniel Kahneman and Maya Bar-Hillel

Abstract. Reports in the 1970s of cognitive illusions in judgments of uncertainty had been anticipated by Laplace 150 years earlier. We discuss Miller and Gelman's remark that Laplace's anticipation of the main ideas of the heuristics and biases approach "gives us a new perspective on these ideas as more universal and less contingent on particular developments [that came much] later."

Key words and phrases: Cognitive illusions, heuristics and biases, judgment under uncertainty.

REFERENCES

GOULD, S. J. (1991). *Bully for Brontosaurus*: *Reflections in Natural History*. W. W. Norton, New York.

LAPLACE, P.-S. (1995). *Philosophical Essay on Probabilities*. *Sources in the History of Mathematics and Physical Sciences* **13**. Springer, New York. Translated from the fifth (1825) French edition, and with notes and a preface by Andrew I. Dale. MR1325241 https://doi.org/10.1007/978-1-4612-4184-3

MILLER, J. and GELMAN, S. (2020). Laplace's theories of cognitive illusions, heuristics, and biases. *Statist. Sci.* **35** 159–170.

TVERSKY, A. and KAHNEMAN, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychol. Rev.* **90** 293–315.

Daniel Kahneman is Professor, Princeton University, 70 East 10th Street, PH-D, New York, New York 10003, USA (e-mail: Kahneman@princeton.edu). Maya Bar-Hillel is Professor, Hebrew University, Federmann Center for the Study of Rationality, Safra Campus, The Hebrew University, Jerusalem 91904, Israel (e-mail: maya@huji.ac.il).

Statistical Science
2020, Vol. 35, No. 2, 173–174
https://doi.org/10.1214/19-STS751

Comment: Illusions, Then and Now

Glenn Shafer

Abstract. The situation of mathematical statistics today resembles the situation of Laplacean probability in the mid-19th century. Strong claims are made about how probability theory should be used. Yet it is widely misused and increasingly not used in the analysis of data. Perhaps we need a philosophy of probability that makes more modest and realistic claims.

Key words and phrases: Laplace, Cournot, cognitive illusions, norms, foundations of probability.

REFERENCES

[1] BERTRAND, J. (1889). *Calcul des probabilités*. Gauthier-Villars, Paris. Second edition 1907.

[2] BRU, B., BRU, M.-F. and BIENAYMÉ, O. (1997). La statistique critiquée par le calcul des probabilités: Deux manuscrits inédits d'Irenée Jules Bienaymé. *Rev. Histoire Math.* **3** 137–239. MR1620388

[3] BRU, M.-F. and BRU, B. (2018). *Les jeux de l'infini et du hasard*. Presses Universitaires de Franche-Comté, Besançon. Two volumes. MR3822089

[4] COURNOT, A. A. (1843). *Exposition de la théorie des chances et des probabilités*. Hachette, Paris.

[5] DE CONDORCET, M. (1785). *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*. l'Imprimerie royale, Paris.

[6] HACKING, I. (1990). *The Taming of Chance*. Cambridge Univ. Press, New York.

[7] JOZEAU, M.-F. (1997). Géodésie au XIXème siècle: De l'hégémonie française à l'hégémonie allemande. Regards belges. Ph.D. thesis, Université Denis Diderot Paris VII, Paris.

[8] LACROIX, S. F. (1816). *Traité élémentaire du calcul des probabilités*. Courcier, Paris. Second edition 1822.

Glenn Shafer is University Professor, Rutgers Business School, 1 Washington Park, Newark, New Jersey 07102, USA (e-mail: gshafer@business.rutgers.edu).

Statistical Science
2020, Vol. 35, No. 2, 175–177
https://doi.org/10.1214/20-STS779

Rejoinder: Laplace's theories of cognitive illusions, heuristics and biases

Joshua B. Miller and Andrew Gelman

Abstract. We appreciate the thoughtful comments from Glenn Shafer and from Daniel Kahneman and Maya Bar-Hillel and respond to each in turn.

Key words and phrases: Cognitive illusions, history of statistics, probability, psychology.

REFERENCES

BRU, M.-F. and BRU, B. (2018). *Les Jeux de L'infini et du Hasard. Sciences*: *Concepts et Problèmes*. [*Sciences*: *Concepts and Problems*]. Presses Universitaires de Franche-Comté, Besançon. Vol. 1. Les probabilités dénombrables à la portée de tous. [Vol. 1. Countable probabilities accessible to all], Vol. 2. Les probabilités indénombrables à la portée de tous. Annexes et appendices. [Vol. 2. Uncountable probabilities accessible to all. Additional texts and appendices]. MR3822089

GILOVICH, T., VALLONE, R. and TVERSKY, A. (1985). The hot hand in basketball: On the misperception of random sequences. *Cognitive Psychology* **17** 295–314.

GORROOCHURN, P. (2012). *Classic Problems of Probability*. Wiley, Hoboken, NJ. MR2976816 https://doi.org/10.1002/9781118314340

LAPLACE, P.-S. (1773, published 1776). Recherches, 1°, sur l'intégration des équations différentielles aux différences finies, et sur leur usage dans la théorie des hasards. 2°, sur Ie principe de la gravitation universelle, et sur les inégalités séculaires des planètes qui en dépendent. *Mémoires de Mathématique et de Physique, Présentés à L'Académie Royale des Sciences Par Divers Savans et Lûs dans Ses Assemblées*.

LAPLACE, P.-S. (1995). *Philosophical Essay on Probabilities*. *Sources in the History of Mathematics and Physical Sciences* **13**. Springer, New York. Translated from the fifth French edition (1825), and with notes and a preface by Andrew I. Dale. MR1325241 https://doi.org/10.1007/978-1-4612-4184-3

MILLER, J. B. and SANJURJO, A. (2018). Surprised by the hot hand fallacy? A truth in the law of small numbers. *Econometrica* **86** 2019–2047. MR3901245 https://doi.org/10.3982/ECTA14943

STIGLER, S. M. (2012). Laplace's unpublished manuscript on associationist psychology. *J. Électron. Hist. Probab. Stat.* **8** 6. MR3025117

TVERSKY, A. and KAHNEMAN, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science* **185** 1124–1131.

TVERSKY, A. and KAHNEMAN, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review* **90** 293–315.

Joshua B. Miller is Associate Professor, Department of Economics, University of Melbourne, Australia. Andrew Gelman is Professor, Department of Statistics and Department of Political Science, Columbia University, New York, New York 10027, USA (e-mail: gelman@stat.columbia.edu).

Statistical Science
2020, Vol. 35, No. 2, 178–201
https://doi.org/10.1214/19-STS716

Fano's Inequality for Random Variables

Sébastien Gerchinovitz, Pierre Ménard and Gilles Stoltz

Abstract. We extend Fano's inequality, which controls the average probability of events in terms of the average of some f-divergences, to work with arbitrary events (not necessarily forming a partition) and even with arbitrary $[0, 1]$-valued random variables, possibly in continuously infinite number. We provide two applications of these extensions, in which the consideration of random variables is particularly handy: we offer new and elegant proofs for existing lower bounds, on Bayesian posterior concentration (minimax or distribution-dependent) rates and on the regret in nonstochastic sequential learning.

Key words and phrases: Multiple-hypotheses testing, lower bounds, information theory, Bayesian posterior concentration.

REFERENCES

AERON, S., SALIGRAMA, V. and ZHAO, M. (2010). Information theoretic bounds for compressed sensing. *IEEE Trans. Inform. Theory* **56** 5111–5130. MR2808668 https://doi.org/10.1109/TIT.2010.2059891

ALI, S. M. and SILVEY, S. D. (1966). A general class of coefficients of divergence of one distribution from another. *J. Roy. Statist. Soc. Ser. B* **28** 131–142. MR0196777

BEREND, D. and KONTOROVICH, A. (2013). On the concentration of the missing mass. *Electron. Commun. Probab.* **18** no. 3, 7. MR3011530 https://doi.org/10.1214/ECP.v18-2359

BIRGÉ, L. (2005). A new lower bound for multiple hypothesis testing. *IEEE Trans. Inform. Theory* **51** 1611–1615. MR2241522 https://doi.org/10.1109/TIT.2005.844101

BOUCHERON, S., LUGOSI, G. and MASSART, P. (2013). *Concentration Inequalities*: *A Nonasymptotic Theory of Independence*. Oxford Univ. Press, Oxford. MR3185193 https://doi.org/10.1093/acprof:oso/9780199535255.001.0001

BRETAGNOLLE, J. and HUBER, C. (1978). Estimation des densités: Risque minimax. In *Séminaire de Probabilités*, *XII* (*Univ. Strasbourg*, *Strasbourg*, 1976/1977). *Lecture Notes in Math.* **649** 342–363. Springer, Berlin. MR0520011

BRETAGNOLLE, J. and HUBER, C. (1979). Estimation des densités: Risque minimax. *Z. Wahrsch. Verw. Gebiete* **47** 119–137. MR0523165 https://doi.org/10.1007/BF00535278

CERF, R. and PETIT, P. (2011). A short proof of Cramér's theorem in $\mathbb{R}$. *Amer. Math. Monthly* **118** 925–931. MR2869520 https://doi.org/10.4169/amer.math.monthly.118.10.925

CESA-BIANCHI, N. and LUGOSI, G. (2006). *Prediction*, *Learning*, *and Games*. Cambridge Univ. Press, Cambridge. MR2409394 https://doi.org/10.1017/CBO9780511546921

CESA-BIANCHI, N., LUGOSI, G. and STOLTZ, G. (2005). Minimizing regret with label efficient prediction. *IEEE Trans. Inform. Theory* **51** 2152–2162. MR2235288 https://doi.org/10.1109/TIT.2005.847729

CESA-BIANCHI, N., FREUND, Y., HAUSSLER, D., HELMBOLD, D. P., SCHAPIRE, R. E. and WARMUTH, M. K. (1997). How to use expert advice. *J. ACM* **44** 427–485. MR1470152 https://doi.org/10.1145/258128.258179

CHEN, X., GUNTUBOYINA, A. and ZHANG, Y. (2016). On Bayes risk lower bounds. *J. Mach. Learn. Res.* **17** Paper No. 218, 58. MR3595153

CHERNOFF, H. (1952). A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *Ann. Math. Stat.* **23** 493–507. MR0057518 https://doi.org/10.1214/aoms/1177729330

COVER, T. M. and THOMAS, J. A. (2006). *Elements of Information Theory*, 2nd ed. Wiley, Hoboken, NJ. MR2239987

CRAMÉR, H. (1938). Sur un nouveau théorème limite de la théorie des probabilités. *Actualités Scientifiques et Industrielles* **736** 5–23.

CSISZÁR, I. (1963). Eine informationstheoretische Ungleichung und ihre Anwendung auf den Beweis der Ergodizität von Markoffschen Ketten. *Magy. Tud. Akad. Mat. Kut. Intéz. Közl.* **8** 85–108. MR0164374

CSISZÁR, I. and KÖRNER, J. (1981). *Information Theory*: *Coding Theorems for Discrete Memoryless Systems*, 3rd ed. Akadémiai Kiadó, Budapest. MR1492180

DUCHI, J. and WAINWRIGHT, M. J. (2013). Distance-based and continuum Fano inequalities with applications to statistical estimation. Available at arXiv:1311.2669.

DURRETT, R. (2010). *Probability*: *Theory and Examples*, 4th ed. *Cambridge Series in Statistical and Probabilistic Mathematics* **31**. Cambridge Univ. Press, Cambridge. MR2722836 https://doi.org/10.1017/CBO9780511779398

FERGUSON, T. S. (1967). *Mathematical Statistics*: *A Decision Theoretic Approach*. *Probability and Mathematical Statistics*, *Vol.* 1. Academic Press, New York. MR0215390

GARIVIER, A., MÉNARD, P. and STOLTZ, G. (2019). Explore first, exploit next: The true shape of regret in bandit problems. *Math.*

Sébastien Gerchinovitz is Assistant Professor, Institut de Mathématiques de Toulouse—Université Paul Sabatier, Toulouse, France (e-mail: sebastien.gerchinovitz@math.univ-toulouse.fr). Pierre Ménard is Ph.D. Student, Institut de Mathématiques de Toulouse—Université Paul Sabatier, Toulouse, France (e-mail: pierre.menard@math.univ-toulouse.fr). Gilles Stoltz is Senior research fellow, Laboratoire de Mathématiques d'Orsay—Université Paris-Sud, CNRS, Université Paris-Saclay, Orsay, France & Affiliate Professor, GREGHEC—HEC Paris, CNRS, Jouy-en-Josas, France (e-mail: gilles.stoltz@math.u-psud.fr).

Oper. Res. **44** 377–399. MR3959077 https://doi.org/10.1287/moor.2017.0928

GERCHINOVITZ, S., MÉNARD, P. and STOLTZ, G. (2018). Fano's inequality for random variables. Available at arXiv:1702.05985.

GHOSAL, S., GHOSH, J. K. and VAN DER VAART, A. W. (2000). Convergence rates of posterior distributions. *Ann. Statist*. **28** 500–531. MR1790007 https://doi.org/10.1214/aos/1016218228

GUNTUBOYINA, A. (2011). Lower bounds for the minimax risk using f-divergences, and applications. *IEEE Trans. Inform. Theory* **57** 2386–2399. MR2809097 https://doi.org/10.1109/TIT.2011.2110791

GUSHCHIN, A. A. (2003). Fano's lemma and analogous inequalities for minimax risk. *Theory Probab. Math. Statist*. **67** 26–37.

HAN, T. S. and VERDÚ, S. (1994). Generalizing the Fano inequality. *IEEE Trans. Inform. Theory* **40** 1247–1251. MR1301430 https://doi.org/10.1109/18.335943

HARREMOËS, P. and VAJDA, I. (2011). On pairs of f-divergences and their joint range. *IEEE Trans. Inform. Theory* **57** 3230–3235. MR2817015 https://doi.org/10.1109/TIT.2011.2137353

HAYASHI, M. (2017). *Quantum Information Theory*, 2nd ed. *Graduate Texts in Physics*. Springer, Berlin. MR3558531 https://doi.org/10.1007/978-3-662-49725-8

HOFFMANN, M., ROUSSEAU, J. and SCHMIDT-HIEBER, J. (2015). On adaptive posterior concentration rates. *Ann. Statist*. **43** 2259–2295. MR3396985 https://doi.org/10.1214/15-AOS1341

IBRAGIMOV, I. A. and HAS'MINSKIĬ, R. Z. (1981). *Statistical Estimation*: *Asymptotic Theory*. *Applications of Mathematics* **16**. Springer, New York. MR0620321

KAUFMANN, E., CAPPÉ, O. and GARIVIER, A. (2016). On the complexity of best-arm identification in multi-armed bandit models. *J. Mach. Learn. Res*. **17** Paper No. 1, 42. MR3482921

KEARNS, M. and SAUL, L. (1998). Large deviation methods for approximate probabilistic inference. In *Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence* (*UAI*'98) 311–319.

KWON, J. and PERCHET, V. (2016). Gains and losses are fundamentally different in regret minimization: The sparse case. *J. Mach. Learn. Res*. **17** Paper No. 227, 32. MR3595163

LE CAM, L. (1986). *Asymptotic Methods in Statistical Decision Theory*. *Springer Series in Statistics*. Springer, New York. MR0856411 https://doi.org/10.1007/978-1-4612-4946-7

LE CAM, L. and YANG, G. L. (2000). *Asymptotics in Statistics*: *Some Basic Concepts*, 2nd ed. *Springer Series in Statistics*. Springer, New York. MR1784901 https://doi.org/10.1007/978-1-4612-1166-2

MASSART, P. (2007). *Concentration Inequalities and Model Selection*. *Lecture Notes in Math*. **1896**. Springer, Berlin. MR2319879

ORDENTLICH, E. and WEINBERGER, M. J. (2005). A distribution dependent refinement of Pinsker's inequality. *IEEE Trans. Inform. Theory* **51** 1836–1840. MR2235683 https://doi.org/10.1109/TIT.2005.846407

PARDO, L. (2006). *Statistical Inference Based on Divergence Measures*. *Statistics*: *Textbooks and Monographs* **185**. CRC Press/CRC, Boca Raton, FL. MR2183173

ROCKAFELLAR, R. T. (1970). *Convex Analysis*. *Princeton Mathematical Series*, *No*. 28. Princeton Univ. Press, Princeton, NJ. MR0274683

STOLTZ, G. (2007). An introduction to the prediction of individual sequences: (1) oracle inequalities; (2) prediction with partial monitoring. Statistics seminar of Univ. Paris VI and Paris VII, Chevaleret, November 12 and 26, 2007; written version of the pair of seminar talks available upon request.

TSYBAKOV, A. B. (2009). *Introduction to Nonparametric Estimation*. *Springer Series in Statistics*. Springer, New York. MR2724359 https://doi.org/10.1007/b13794

YU, B. (1997). Assouad, Fano, and Le Cam. In *Festschrift for Lucien Le Cam* (D. Pollard, E. Torgersen and G. L. Yang, eds.) 423–435. Springer, New York. MR1462963

ZHANG, T. (2006). Information-theoretic upper and lower bounds for statistical estimation. *IEEE Trans. Inform. Theory* **52** 1307–1321. MR2241190 https://doi.org/10.1109/TIT.2005.864439

Statistical Science
2020, Vol. 35, No. 2, 202–217
https://doi.org/10.1214/19-STS719

Equitability, Interval Estimation, and Statistical Power

Yakir A. Reshef, David N. Reshef, Pardis C. Sabeti and Michael Mitzenmacher

Abstract. Emerging high-dimensional data sets often contain many nontrivial relationships, and, at modern sample sizes, screening these using an independence test can sometimes yield too many relationships to be a useful exploratory approach. We propose a framework to address this limitation centered around a property of measures of dependence called *equitability*. Given some measure of relationship strength, an equitable measure of dependence is one that assigns similar scores to equally strong relationships of different types. We formalize equitability within a semiparametric inferential framework in terms of interval estimates of relationship strength, and we then use the correspondence of these interval estimates to hypothesis tests to show that equitability is equivalent under moderate assumptions to requiring that a measure of dependence yield well-powered tests not only for distinguishing nontrivial relationships from trivial ones but also for distinguishing stronger relationships from weaker ones. We then show that equitability, to the extent it is achieved, implies that a statistic will be well powered to detect all relationships of a certain minimal strength, across different relationship types in a family. Thus, equitability is a strengthening of power against independence that enables exploration of data sets with a small number of strong, interesting relationships and a large number of weaker, less interesting ones.

Key words and phrases: Equitability, measure of dependence, statistical power, independence test, semiparametric inference.

REFERENCES

ACHARYA, J., DASKALAKIS, C. and KAMATH, G. (2015). Optimal testing for properties of distributions. In *Advances in Neural Information Processing Systems* 3591–3599.

ARIAS-CASTRO, E., PELLETIER, B. and SALIGRAMA, V. (2018). Remember the curse of dimensionality: The case of goodness-of-fit testing in arbitrary dimension. *J. Nonparametr. Stat.* **30** 448–471. MR3794401 https://doi.org/10.1080/10485252.2018.1435875

BARAUD, Y. (2002). Non-asymptotic minimax rates of testing in signal detection. *Bernoulli* **8** 577–606. MR1935648

BREIMAN, L. and FRIEDMAN, J. H. (1985). Estimating optimal transformations for multiple regression and correlation. *J. Amer. Statist. Assoc.* **80** 580–619. MR0803258

CASELLA, G. and BERGER, R. L. (2002). *Statistical Inference*, *Vol. 2. The Wadsworth & Brooks/Cole Statistics/Probability Series*. Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove, CA. MR1051420

COVER, T. M. and THOMAS, J. A. (2006). *Elements of Information Theory*, 2nd ed. Wiley Interscience, Hoboken, NJ. MR2239987

CSISZÁR, I. (2008). Axiomatic characterizations of information measures. *Entropy* **10** 261–273.

DING, A. A., DY, J. G., LI, Y. and CHANG, Y. (2017). A robust-equitable measure for feature ranking and selection. *J. Mach. Learn. Res.* **18** Paper No. 71, 46. MR3714234

EMILSSON, V., THORLEIFSSON, G., ZHANG, B., LEONARDSON, A. S., ZINK, F., ZHU, J., CARLSON, S., HELGASON, A., WALTERS, G. B. et al. (2008). Genetics of gene expression and its effect on disease. *Nature* **452** 423–428.

FAUST, K. and RAES, J. (2012). Microbial interactions: From networks to models. *Nat. Rev., Microbiol.* **10** 538–550. https://doi.org/10.1038/nrmicro2832

FROMONT, M., LERASLE, M. and REYNAUD-BOURET, P. (2016). Family-wise separation rates for multiple testing. *Ann. Statist.* **44** 2533–2563. MR3576553 https://doi.org/10.1214/15-AOS1418

GRETTON, A., BOUSQUET, O., SMOLA, A. and SCHÖLKOPF, B. (2005a). Measuring statistical dependence with Hilbert–Schmidt norms. In *Algorithmic Learning Theory. Lecture Notes in*

Yakir A. Reshef is Ph.D. candidate, School of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts 02138, USA (e-mail: yakir@seas.harvard.edu). David N. Reshef is Ph.D. candidate, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA (e-mail: dnreshef@mit.edu). Pardis C. Sabeti is Professor, Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, Massachusetts 02138, USA (e-mail: pardis@broadinstitute.org). Michael Mitzenmacher is Professor, School of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts 02138, USA (e-mail: michaelm@eecs.harvard.edu).

Computer Science **3734** 63–77. Springer, Berlin. MR2255909 https://doi.org/10.1007/11564089_7

GRETTON, A., HERBRICH, R., SMOLA, A., BOUSQUET, O. and SCHÖLKOPF, B. (2005b). Kernel methods for measuring independence. *J. Mach. Learn. Res.* **6** 2075–2129. MR2249882

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

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

HELLER, R., HELLER, Y., KAUFMAN, S., BRILL, B. and GORFINE, M. (2016). Consistent distribution-free K-sample and independence tests for univariate random variables. *J. Mach. Learn. Res.* **17** Paper No. 29, 54. MR3491123

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

HUO, X. and SZÉKELY, G. J. (2016). Fast computing for distance covariance. *Technometrics* **58** 435–447. MR3556612 https://doi.org/10.1080/00401706.2015.1054435

INGSTER, Y. I. (1989). Asymptotic minimax testing of independence hypothesis. *J. Sov. Math.* **44** 466–476.

INGSTER, Y. I. (1987). Asymptotically minimax testing of nonparametric hypotheses. In *Probability Theory and Mathematical Statistics*, *Vol. I* (*Vilnius*, 1985) 553–574. VNU Sci. Press, Utrecht. MR0901514

INGSTER, Y. I. and SUSLINA, I. A. (2003). *Nonparametric Goodness-of-Fit Testing Under Gaussian Models*. *Lecture Notes in Statistics* **169**. Springer, New York. MR1991446 https://doi.org/10.1007/978-0-387-21580-8

JIANG, B., YE, C. and LIU, J. S. (2015). Nonparametric K-sample tests via dynamic slicing. *J. Amer. Statist. Assoc.* **110** 642–653. MR3367254 https://doi.org/10.1080/01621459.2014.920257

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

KRASKOV, A., STÖGBAUER, H. and GRASSBERGER, P. (2004). Estimating mutual information. *Phys. Rev. E* (3) **69** 066138, 16. MR2096503 https://doi.org/10.1103/PhysRevE.69.066138

LEPSKI, O. V. and SPOKOINY, V. G. (1999). Minimax nonparametric hypothesis testing: The case of an inhomogeneous alternative. *Bernoulli* **5** 333–358. MR1681702 https://doi.org/10.2307/3318439

LINFOOT, E. H. (1957). An informational measure of correlation. *Inf. Control* **1** 85–89. MR0092706

LOPEZ-PAZ, D., HENNIG, P. and SCHÖLKOPF, B. (2013). The randomized dependence coefficient. In *Advances in Neural Information Processing Systems* 1–9.

LOPEZ-PAZ, D., MUANDET, K., SCHÖLKOPF, B. and TOLSTIKHIN, I. (2015). Towards a learning theory of causation. In *International Conference on Machine Learning* (*ICML*).

MÓRI, T. F. and SZÉKELY, G. J. (2019). Four simple axioms of dependence measures. *Metrika* **82** 1–16. MR3897521 https://doi.org/10.1007/s00184-018-0670-3

MURRELL, B., MURRELL, D. and MURRELL, H. (2014). R2-equitability is satisfiable. *Proc. Natl. Acad. Sci. USA* **111** E2160–E2160.

MURRELL, B., MURRELL, D. and MURRELL, H. (2016). Discovering general multidimensional associations. *PLoS ONE* **11** e0151551. https://doi.org/10.1371/journal.pone.0151551

PANINSKI, L. (2008). A coincidence-based test for uniformity given very sparsely sampled discrete data. *IEEE Trans. Inform. Theory* **54** 4750–4755. MR2591136 https://doi.org/10.1109/TIT.2008.928987

WANG, Y. X. R., WATERMAN, M. S. and HUANG, H. (2014). Gene coexpression measures in large heterogeneous samples using count statistics. *Proc. Natl. Acad. Sci. USA* **111** 16371–16376.

REIMHERR, M. and NICOLAE, D. L. (2013). On quantifying dependence: A framework for developing interpretable measures. *Statist. Sci.* **28** 116–130. MR3075341 https://doi.org/10.1214/12-STS405

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

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

RESHEF, D. N., RESHEF, Y. A., MITZENMACHER, M. and SABETI, P. C. (2014). Cleaning up the record on the maximal information coefficient and equitability. *Proc. Natl. Acad. Sci. USA* **111** E3362–E3363.

RESHEF, Y. A., RESHEF, D. N., FINUCANE, H. K., SABETI, P. C. and MITZENMACHER, M. (2016). Measuring dependence powerfully and equitably. *J. Mach. Learn. Res.* **17** Paper No. 212, 63. MR3595146

RESHEF, D. N., RESHEF, Y. A., SABETI, P. C. and MITZENMACHER, M. (2018). An empirical study of the maximal and total information coefficients and leading measures of dependence. *Ann. Appl. Stat.* **12** 123–155. MR3773388 https://doi.org/10.1214/17-AOAS1093

ROMANO, S., VINH, N. X., VERSPOOR, K. and BAILEY, J. (2018). The randomized information coefficient: Assessing dependencies in noisy data. *Mach. Learn.* **107** 509–549. MR3761295 https://doi.org/10.1007/s10994-017-5664-2

SCHWEIZER, B. and WOLFF, E. F. (1981). On nonparametric measures of dependence for random variables. *Ann. Statist.* **9** 879–885. MR0619291

SIMON, N. and TIBSHIRANI, R. (2012). Comment on "Detecting novel associations in large data sets." Unpublished.

SPEED, T. (2011). A correlation for the 21st century. *Science* **334** 1502–1503.

STOREY, J. D. and TIBSHIRANI, R. (2003). Statistical significance for genomewide studies. *Proc. Natl. Acad. Sci. USA* **100** 9440–9445. MR1994856 https://doi.org/10.1073/pnas.1530509100

SUGIYAMA, M. and BORGWARDT, K. M. (2013). Measuring statistical dependence via the mutual information dimension. In *The International Joint Conferences on Artificial Intelligence* (*IJCAI*) 1692–1698. AAAI Press, Menlo Park, CA.

SUN, N. and ZHAO, H. (2014). Putting things in order. *Proc. Natl. Acad. Sci. USA* **111** 16236–16237.

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

SZÉKELY, G. J., 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

TURK-BROWNE, N. B. (2013). Functional interactions as big data in the human brain. *Science* **342** 580–584.

WANG, X., JIANG, B. and LIU, J. S. (2017). Generalized R-squared for detecting dependence. *Biometrika* **104** 129–139. MR3626486 https://doi.org/10.1093/biomet/asw071

YATRACOS, Y. G. (1985). On the existence of uniformly consistent estimates. *Proc. Amer. Math. Soc.* **94** 479–486. MR0787899 https://doi.org/10.2307/2045240

YODÉ, A. F. (2011). Adaptive minimax test of independence. *Math. Methods Statist.* **20** 246–268. MR2908761 https://doi.org/10.3103/S1066530711030069

ZHANG, K. (2016). Bet on independence. Preprint. Available at arXiv:1610.05246.

Statistical Science
2020, Vol. 35, No. 2, 218–233
https://doi.org/10.1214/19-STS727

LGM Split Sampler: An Efficient MCMC Sampling Scheme for Latent Gaussian Models

Óli Páll Geirsson, Birgir Hrafnkelsson, Daniel Simpson and Helgi Sigurdarson

Abstract. A general and flexible class of latent Gaussian models is proposed in this paper. The latent Gaussian model is adapted to the generalized additive model for location, scale and shape (GAMLSS), that is, the data density function of each data point can depend on more than a single linear predictor of the latent parameters. We refer to this framework as extended latent Gaussian models. The most commonly applied latent Gaussian models (LGMs) are such that a linear predictor is proposed only for the location parameter. Extended LGMs allow proposing linear predictors also for the scale parameter and potentially other parameters. We propose a novel computationally efficient Markov chain Monte Carlo sampling scheme for the extended LGMs which we refer to as the LGM split sampler. It is a two block Gibbs sampling scheme designed to exploit the model structure of the extended LGMs. An extended LGM is constructed for a simulated dataset and the LGM split sampler is implemented for posterior simulations. The results demonstrate the flexibility of the extended LGM framework and the efficiency of the LGM split sampler.

Key words and phrases: Bayesian hierarchical models, Gibbs sampling, latent Gaussian models, Markov chain Monte Carlo, posterior simulation.

REFERENCES

CHILÈS, J.-P. and DELFINER, P. (2012). *Geostatistics*: *Modeling Spatial Uncertainty*, 2nd ed. *Wiley Series in Probability and Statistics*. Wiley, Hoboken, NJ. MR2850475 https://doi.org/10.1002/9781118136188

COOLEY, D., NYCHKA, D. and NAVEAU, P. (2007). Bayesian spatial modeling of extreme precipitation return levels. *J. Amer. Statist. Assoc.* **102** 824–840. MR2411647 https://doi.org/10.1198/016214506000000780

CRESSIE, N. A. C. (1993). *Statistics for Spatial Data*, 2nd ed. *Wiley Series in Probability and Mathematical Statistics*: *Applied Probability and Statistics*. Wiley, New York. MR1239641 https://doi.org/10.1002/9781119115151

CROCHET, P. (2012). Estimating the flood frequency distribution for ungauged catchments using an index flood procedure. Application to ten catchments in Northern Iceland. Technical Report No. VÍ 2012-005, The Icelandic Meteorological Office.

CROCHET, P., JÓHANNESSON, T., JÓNSSON, T., SIGURÐSSON, O., BJÖRNSSON, H., PÁLSSON, F. and BARSTAD, I. (2007). Estimating the spatial distribution of precipitation in Iceland using a linear model of orographic precipitation. *J. Hydrometeorol.* **8** 1285–1306.

CUNNANE, C. and NASH, J. E. (1971). Bayesian estimation of frequency of hydrological events. *Proc. Warsaw Symp. Math. Models Hydrol.* **1** 47–55.

DAVÍÐSSON, Ó. B. (2015). Bayesian flood frequency analysis using monthly maxima. Master's thesis, Univ. Iceland.

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

DIGGLE, P. J., TAWN, J. A. and MOYEED, R. A. (1998). Model-based geostatistics. *J. R. Stat. Soc. Ser. C. Appl. Stat.* **47** 299–350. MR1626544 https://doi.org/10.1111/1467-9876.00113

FAHRMEIR, L. and TUTZ, G. (1994). *Multivariate Statistical Modelling Based on Generalized Linear Models*. *Springer Series in Statistics*. Springer, New York. MR1284203 https://doi.org/10.1007/978-1-4899-0010-4

FERKINGSTAD, E., FRIGESSI, A., RUE, H., THORLEIFSSON, G. and KONG, A. (2008). Unsupervised empirical Bayesian multi-

Óli Páll Geirsson is Chief Data Officer, City of Reykjavik, Borgartún 12-14, 105 Reykjavik, Iceland (e-mail: olipalli@gmail.com). Birgir Hrafnkelsson is Professor of Statistics, Department of Mathematics, Faculty of Physical Sciences, University of Iceland, Dunhagi 5, 107 Reykjavik, Iceland (e-mail: birgirhr@hi.is). Daniel Simpson is Assistant Professor, Department of Statistical Sciences, University of Toronto, 100 St. George Street, Toronto, Ontario M5S 3G3, Canada (e-mail: simpson@utstat.toronto.edu). Helgi Sigurdarson is Project Manager, Air Navigation Services, Isavia, Reykjavik Airport, 102 Reykjavik, Iceland (e-mail: Helgi.Sigurdarson@isavia.is).

ple testing with external covariates. *Ann. Appl. Stat.* **2** 714–735. MR2524353 https://doi.org/10.1214/08-AOAS158

FILIPPONE, M., ZHONG, M. and GIROLAMI, M. (2013). A comparative evaluation of stochastic-based inference methods for Gaussian process models. *Mach. Learn.* **93** 93–114. MR3093120 https://doi.org/10.1007/s10994-013-5388-x

GEIRSSON, Ó. P., HRAFNKELSSON, B. and SIMPSON, D. (2015). Computationally efficient spatial modeling of annual maximum 24-h precipitation on a fine grid. *Environmetrics* **26** 339–353. MR3366968 https://doi.org/10.1002/env.2343

GEIRSSON, Ó. P., HRAFNKELSSON, B., SIMPSON, D. and SIGURDARSON, H. (2020). Supplement to "LGM split sampler: An efficient MCMC sampling scheme for latent Gaussian models." https://doi.org/10.1214/19-STS727SUPP.

GIROLAMI, M. and CALDERHEAD, B. (2011). Riemann manifold Langevin and Hamiltonian Monte Carlo methods. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **73** 123–214. MR2814492 https://doi.org/10.1111/j.1467-9868.2010.00765.x

GREHYS (1996). Presentation and review of some methods for regional flood frequency analysis. *J. Hydrol.* **186** 63–84.

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

HRAFNKELSSON, B., MORRIS, J. S. and BALADANDAYUTHAPANI, V. (2012). Spatial modeling of annual minimum and maximum temperatures in Iceland. *Meteorol. Atmos. Phys.* **116** 43–61.

KLEIN, N., KNEIB, T., LANG, S. et al. (2013). Bayesian structured additive distributional regression. Technical Report No. 2013-23, Department of Public Finance, Univ. Innsbruck.

KLEIN, N., KNEIB, T., KLASEN, S. and LANG, S. (2015). Bayesian structured additive distributional regression for multivariate responses. *J. R. Stat. Soc. Ser. C. Appl. Stat.* **64** 569–591. MR3367789 https://doi.org/10.1111/rssc.12090

KNEIB, T. (2013). Beyond mean regression. *Stat. Model.* **13** 275–303. MR3179527 https://doi.org/10.1177/1471082X13494159

KNORR-HELD, L. and RUE, H. (2002). On block updating in Markov random field models for disease mapping. *Scand. J. Stat.* **29** 597–614. MR1988414 https://doi.org/10.1111/1467-9469.00308

LAWSON, A. B. (2013). *Bayesian Disease Mapping*: *Hierarchical Modeling in Spatial Epidemiology*. *Interdisciplinary Statistics*. CRC Press, Boca Raton, FL. MR2484272

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

MARTINO, S., AAS, K., LINDQVIST, O., NEEF, L. R. and RUE, H. (2011). Estimating stochastic volatility models using integrated nested Laplace approximations. *Eur. J. Finance* **17** 487–503.

MARTINS, T. G., SIMPSON, D., LINDGREN, F. and RUE, H. (2013a). Bayesian computing with INLA: New features. *Comput. Statist. Data Anal.* **67** 68–83. MR3079584 https://doi.org/10.1016/j.csda.2013.04.014

MARTINS, T. G., SIMPSON, D., ILLIAN, J. B., RUE, H. and GEIRSSON, Ó. P. (2013b). Discussion of 'Beyond mean regression' [MR3179527]. *Stat. Model.* **13** 355–361. MR3179533 https://doi.org/10.1177/1471082X13494527

NEAL, R. M. (1993). Probabilistic inference using Markov chain Monte Carlo methods. Technical Report No. CRG-TR-93-1, Department of Computer Science, Univ. Toronto.

PETTITT, A. N., WEIR, I. S. and HART, A. G. (2002). A conditional autoregressive Gaussian process for irregularly spaced multivariate data with application to modelling large sets of binary data. *Stat. Comput.* **12** 353–367. MR1951708 https://doi.org/10.1023/A:1020792130229

RIGBY, R. A. and STASINOPOULOS, D. M. (2005). Generalized additive models for location, scale and shape. *J. R. Stat. Soc. Ser. C. Appl. Stat.* **54** 507–554. MR2137253 https://doi.org/10.1111/j.1467-9876.2005.00510.x

ROBERTS, G. O., GELMAN, A. and GILKS, W. R. (1997). Weak convergence and optimal scaling of random walk Metropolis algorithms. *Ann. Appl. Probab.* **7** 110–120. MR1428751 https://doi.org/10.1214/aoap/1034625254

ROBERTS, G. O. and ROSENTHAL, J. S. (1998). Optimal scaling of discrete approximations to Langevin diffusions. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **60** 255–268. MR1625691 https://doi.org/10.1111/1467-9868.00123

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

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

RUE, H., RIEBLER, A., SØRBYE, .S. H., ILLIAN, J. B., SIMPSON, D. P. and LINDGREN, F. K. (2017). Bayesian computing with INLA: A review. *Ann. Rev. Stat. Appl.* **4** 395–421.

SCHAEFLI, B., TALAMBA, D. B. and MUSY, A. (2007). Quantifying hydrological modeling errors through a mixture of normal distributions. *J. Hydrol.* **332** 303–315.

ZHANG, H. (2004). Inconsistent estimation and asymptotically equal interpolations in model-based geostatistics. *J. Amer. Statist. Assoc.* **99** 250–261. MR2054303 https://doi.org/10.1198/016214504000000241

Statistical Science
2020, Vol. 35, No. 2, 234–253
https://doi.org/10.1214/19-STS731

Checking for Prior-Data Conflict Using Prior-to-Posterior Divergences

David J. Nott, Xueou Wang, Michael Evans and Berthold-Georg Englert

Abstract. When using complex Bayesian models to combine information, checking consistency of the information contributed by different components of the model for inference is good statistical practice. Here a new method is developed for detecting prior-data conflicts in Bayesian models based on comparing the observed value of a prior-to-posterior divergence to its distribution under the prior predictive distribution for the data. The divergence measure used in our model check is a measure of how much beliefs have changed from prior to posterior, and can be thought of as a measure of the overall size of a relative belief function. It is shown that the proposed method is intuitive, has desirable properties, can be extended to hierarchical settings, and is related asymptotically to Jeffreys' and reference prior distributions. In the case where calculations are difficult, the use of variational approximations as a way of relieving the computational burden is suggested. The methods are compared in a number of examples with an alternative but closely related approach in the literature based on the prior predictive distribution of a minimal sufficient statistic.

Key words and phrases: Bayesian inference, model checking, prior data-conflict, variational Bayes, Bayesian inference.

REFERENCES

AL LABADI, L. and EVANS, M. (2017). Optimal robustness results for relative belief inferences and the relationship to prior-data conflict. *Bayesian Anal.* **12** 705–728. MR3655873 https://doi.org/10.1214/16-BA1024

ALBERT, J. (2009). *Bayesian Computation with R*, 2nd ed. *Use R!* Springer, Dordrecht. MR2839312 https://doi.org/10.1007/978-0-387-92298-0

ANDRADE, J. A. A. and O'HAGAN, A. (2006). Bayesian robustness modeling using regularly varying distributions. *Bayesian Anal.* **1** 169–188. MR2227369 https://doi.org/10.1214/06-BA106

ATTIAS, H. (1999). Inferring parameters and structure of latent variable models by variational Bayes. In *Proceedings of the* 15*th Conference on Uncertainty in Artificial Intelligence* (K. Laskey and H. Prade, eds.) 21–30. Morgan Kaufmann, San Francisco, CA.

BASKURT, Z. and EVANS, M. (2013). Hypothesis assessment and inequalities for Bayes factors and relative belief ratios. *Bayesian Anal.* **8** 569–590. MR3102226 https://doi.org/10.1214/13-BA824

BAYARRI, M. J. and BERGER, J. O. (2000). p values for composite null models. *J. Amer. Statist. Assoc.* **95** 1127–1142, 1157–1170. MR1804239 https://doi.org/10.2307/2669749

BAYARRI, M. J. and CASTELLANOS, M. E. (2007). Bayesian checking of the second levels of hierarchical models. *Statist. Sci.* **22** 322–343. MR2416808 https://doi.org/10.1214/07-STS235

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

BOUSQUET, N. (2008). Diagnostics of prior-data agreement in applied Bayesian analysis. *J. Appl. Stat.* **35** 1011–1029. MR2522125 https://doi.org/10.1080/02664760802192981

BOX, G. E. P. (1980). Sampling and Bayes' inference in scientific modelling and robustness (with discussion). *J. Roy. Statist. Soc. Ser. A* **143** 383–430. MR0603745 https://doi.org/10.2307/2982063

CAROTA, C., PARMIGIANI, G. and POLSON, N. G. (1996). Diagnostic measures for model criticism. *J. Amer. Statist. Assoc.* **91** 753–762. MR1395742 https://doi.org/10.2307/2291670

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

David J. Nott is Associate Professor and the corresponding author, Department of Statistics and Applied Probability, Singapore 117546, and Operations Research and Analytics Cluster, National University of Singapore, Singapore 119077 (e-mail: standj@nus.edu.sg). Xueou Wang is Ph.D Student, Department of Statistics and Applied Probability, National University of Singapore, Singapore 117546 (e-mail: a0095911@u.nus.edu). Michael Evans is Professor, Department of Statistical Sciences, University of Toronto, Toronto, Ontario, M5S 3G3, Canada (e-mail: mevans@utstat.utoronto.ca). Berthold-Georg Englert is Professor and Principal Investigator, Center for Quantum Technologies and Department of Physics, National University of Singapore, 117542 and MajuLab, CNRS-UNS-NUS-NTU International Joint Research Unit, UMI 3654, Singapore (e-mail: cqtebg@nus.edu.sg).

COX, D. R. and HINKLEY, D. V. (1974). *Theoretical Statistics*. CRC Press, London. MR0370837

DAHL, F. A., GÅSEMYR, J. and NATVIG, B. (2007). A robust conflict measure of inconsistencies in Bayesian hierarchical models. *Scand. J. Stat.* **34** 816–828. MR2396940 https://doi.org/10.1111/j.1467-9469.2007.00560.x

DEY, D. K., GELFAND, A. E., SWARTZ, T. B. and VLACHOS, P. K. (1998). A simulation-intensive approach for checking hierarchical models. *TEST* **7** 325–346.

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

EVANS, M. and JANG, G. H. (2010). Invariant P-values for model checking. *Ann. Statist.* **38** 512–525. MR2589329 https://doi.org/10.1214/09-AOS727

EVANS, M. and JANG, G. H. (2011a). A limit result for the prior predictive applied to checking for prior-data conflict. *Statist. Probab. Lett.* **81** 1034–1038. MR2803740 https://doi.org/10.1016/j.spl.2011.02.025

EVANS, M. and JANG, G. H. (2011b). Weak informativity and the information in one prior relative to another. *Statist. Sci.* **26** 423–439. MR2917964 https://doi.org/10.1214/11-STS357

EVANS, M. and MOSHONOV, H. (2006). Checking for prior-data conflict. *Bayesian Anal.* **1** 893–914. MR2282210 https://doi.org/10.1016/j.spl.2011.02.025

GÅSEMYR, J. and NATVIG, B. (2009). Extensions of a conflict measure of inconsistencies in Bayesian hierarchical models. *Scand. J. Stat.* **36** 822–838. MR2573310 https://doi.org/10.1111/j.1467-9469.2009.00659.x

GELMAN, A. (2006). Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). *Bayesian Anal.* **1** 515–533. MR2221284 https://doi.org/10.1214/06-BA117A

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

GELMAN, A. and SHALIZI, C. R. (2013). Philosophy and the practice of Bayesian statistics. *Br. J. Math. Stat. Psychol.* **66** 8–38. MR3044854 https://doi.org/10.1111/j.2044-8317.2011.02037.x

GERSHMAN, S., HOFFMAN, M. D. and BLEI, D. M. (2012). Nonparametric variational inference. In *Proceedings of the* 29*th International Conference on Machine Learning*, *ICML* 2012.

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

GHOSH, J. K., DELAMPADY, M. and SAMANTA, T. (2006). *An Introduction to Bayesian Analysis*: *Theory and Methods*. *Springer Texts in Statistics*. Springer, New York. MR2247439

GIL, M., ALAJAJI, F. and LINDER, T. (2013). Rényi divergence measures for commonly used univariate continuous distributions. *Inform. Sci.* **249** 124–131. MR3105467 https://doi.org/10.1016/j.ins.2013.06.018

GUTTMAN, I. (1967). The use of the concept of a future observation in goodness-of-fit problems. *J. Roy. Statist. Soc. Ser. B* **29** 83–100. MR0216699

HERSHEY, J. R. and OLSEN, P. A. (2007). Approximating the Kullback Leibler divergence between Gaussian mixture models. In 2007 *IEEE International Conference on Acoustics*, *Speech and Signal Processing—ICASSP* '07 **4** IV-317–IV-320.

HJORT, N. L., DAHL, F. A. and STEINBAKK, G. H. (2006). Post-processing posterior predictive p-values. *J. Amer. Statist. Assoc.* **101** 1157–1174. MR2324154 https://doi.org/10.1198/016214505000001393

JAYNES, E. T. (1976). Confidence intervals vs. Bayesian intervals (1976). In *Foundations of Probability Theory*, *Statistical Inference*, *and Statistical Theories of Science*, *Vol. II* 175–257. Reidel, Dordrecht.

JOHNSON, R. A. (1970). Asymptotic expansions associated with posterior distributions. *Ann. Math. Stat.* **41** 851–864. MR0263198 https://doi.org/10.1214/aoms/1177696963

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.

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

LI, X., SHANG, J., NG, H. K. and ENGLERT, B.-G. (2016). Optimal error intervals for properties of the quantum state. *Phys. Rev. A* **94** 062112.

LIESE, F. and VAJDA, I. (1987). *Convex Statistical Distances*. *Teubner-Texte zur Mathematik* [*Teubner Texts in Mathematics*] **95**. BSB B. G. Teubner Verlagsgesellschaft, Leipzig. With German, French and Russian summaries. MR0926905

MARSHALL, E. C. and SPIEGELHALTER, D. J. (2007). Identifying outliers in Bayesian hierarchical models: A simulation-based approach. *Bayesian Anal.* **2** 409–444. MR2312289 https://doi.org/10.1214/07-BA218

O'HAGAN, A. (2003). HSSS model criticism. In *Highly Structured Stochastic Systems*. *Oxford Statist. Sci. Ser.* **27** 423–453. Oxford Univ. Press, Oxford. MR2082418

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

REIMHERR, M., MENG, X.-L. and NICOLAE, D. L. (2014). Being an informed Bayesian: Assessing prior informativeness and prior likelihood conflict. Available at arXiv:1406.5958.

RÉNYI, A. (1961). On measures of entropy and information. In *Proc. 4th Berkeley Sympos. Math. Statist. and Prob.*, *Vol. I* 547–561. Univ. California Press, Berkeley, CA. MR0132570

ROBINS, J. M., VAN DER VAART, A. and VENTURA, V. (2000). Asymptotic distribution of p values in composite null models. *J. Amer. Statist. Assoc.* **95** 1143–1167, 1171–1172. MR1804240 https://doi.org/10.2307/2669750

ROHDE, D. and WAND, M. P. (2016). Semiparametric mean field variational Bayes: General principles and numerical issues. *J. Mach. Learn. Res.* **17** 172. MR3567440

RUBIN, D. B. (1984). Bayesianly justifiable and relevant frequency calculations for the applied statistician. *Ann. Statist.* **12** 1151–1172. MR0760681 https://doi.org/10.1214/aos/1176346785

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

SCHEEL, I., GREEN, P. J. and ROUGIER, J. C. (2011). A graphical diagnostic for identifying influential model choices in Bayesian hierarchical models. *Scand. J. Stat.* **38** 529–550. MR2833845 https://doi.org/10.1111/j.1467-9469.2010.00717.x

SPIEGELHALTER, D. J., AYLIN, P., BEST, N. G., EVANS, S. J. W. and MURRAY, G. D. (2002). Commissioned analysis of surgical performance using routine data: Lessons from the Bristol inquiry. *J. Roy. Statist. Soc. Ser. A* **165** 191–231. MR1869173 https://doi.org/10.1111/1467-985X.02021

TAN, L. S. L. and NOTT, D. J. (2014). A stochastic variational framework for fitting and diagnosing generalized linear mixed models. *Bayesian Anal*. **9** 963–1004. MR3293964 https://doi.org/10.1214/14-BA885

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

Statistical Science
2020, Vol. 35, No. 2, 254–271
https://doi.org/10.1214/19-STS732

A Generalized Approach to Power Analysis for Local Average Treatment Effects

Kirk Bansak

Abstract. This study introduces a new approach to power analysis in the context of estimating a local average treatment effect (LATE), where the study subjects exhibit noncompliance with treatment assignment. As a result of distributional complications in the LATE context, compared to the simple ATE context, there is currently no standard method of power analysis for the LATE. Moreover, existing methods and commonly used substitutes—which include instrumental variable (IV), intent-to-treat (ITT) and scaled ATE power analyses—require specifying generally unknown variance terms and/or rely upon strong and unrealistic assumptions, thus providing unreliable guidance on the power of tests of the LATE. This study develops a new approach that uses standardized effect sizes to place bounds on the power for the most commonly used estimator of the LATE, the Wald IV estimator, whereby variance terms and distributional parameters need not be specified nor assumed. Instead, in addition to the effect size, sample size and error tolerance parameters, the only other parameter that must be specified by the researcher is the compliance rate. Additional conditions can also be introduced to further narrow the bounds on the power calculation. The result is a generalized approach to power analysis in the LATE context that is simple to implement.

Key words and phrases: Experimental design, local average treatment effects, noncompliance, principal stratification, statistical power.

REFERENCES

ABADIE, A., ANGRIST, J. and IMBENS, G. (2002). Instrumental variables estimates of the effect of subsidized training on the quantiles of trainee earnings. *Econometrica* **70** 91–117. MR1926256 https://doi.org/10.1111/1468-0262.00270

ANGRIST, J. D. (1990). Lifetime earnings and the Vietnam era draft lottery: Evidence from Social Security administrative records. *Am. Econ. Rev.* **80** 313–336.

ANGRIST, J. D. and IMBENS, G. W. (1995). Two-stage least squares estimation of average causal effects in models with variable treatment intensity. *J. Amer. Statist. Assoc.* **90** 431–442. MR1340501

ANGRIST, J. D., IMBENS, G. W. and RUBIN, D. B. (1996). Identification of causal effects using instrumental variables. *J. Amer. Statist. Assoc.* **91** 444–455.

ANGRIST, J. D. and PISCHKE, J.-S. (2009). *Mostly Harmless Econometrics*: *An Empiricist's Companion*. Princeton Univ. Press, Princeton, NJ.

BAIOCCHI, M., CHENG, J. and SMALL, D. S. (2014). Instrumental variable methods for causal inference. *Stat. Med.* **33** 2297–2340. MR3257582 https://doi.org/10.1002/sim.6128

BANSAK, K. (2020). Supplement to "A Generalized Approach to Power Analysis for Local Average Treatment Effects." https://doi.org/10.1214/19-STS732SUPP.

BLOOM, H. S. (1995). Minimum detectable effects: A simple way to report the statistical power of experimental designs. *Evaluation Review* **19** 547–556.

BLOOM, H. S. (2006). The core analytics of randomized experiments for social research. Technical report, MDRC.

BLOOM, H. S., ORR, L. L., BELL, S. H., CAVE, G., DOOLITTLE, F., LIN, W. and BOS, J. M. (1997). The benefits and costs of JTPA Title II-A programs: Key findings from the National Job Training Partnership Act Study. *J. Hum. Resour.* **32** 549–576.

BRION, M.-J. A., SHAKHBAZOV, K. and VISSCHER, P. M. (2013). Calculating statistical power in Mendelian randomization studies. *Int. J. Epidemiol.* **42** 1497–1501.

COHEN, J. (1962). The statistical power of abnormal-social psychological research. *J. Abnorm. Soc. Psychol.* **65** 145–153.

COHEN, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Lawrence Erlbaum Associates, Hillsdale, NJ.

DUFLO, E., GLENNERSTER, R. and KREMER, M. (2007). Using randomization in development economics research: A toolkit. Technical report, Centre for Economic Policy Research, London.

FREEMAN, G., COWLING, B. J. and SCHOOLING, C. M. (2013). Power and sample size calculations for Mendelian randomization studies using one genetic instrument. *Int. J. Epidemiol.* **42** 1157–1163.

Kirk Bansak is Assistant Professor, Department of Political Science, University of California San Diego, Social Sciences Building 301, La Jolla, California 92093-0521, USA (e-mail: kbansak@ucsd.edu).

GERBER, A. S. and GREEN, D. P. (2000). The effects of canvassing, telephone calls, and direct mail on voter turnout: A field experiment. *Am. Polit. Sci. Rev.* **94** 653–663.

GERBER, A. S. and GREEN, D. P. (2012). *Field Experiments*: *Design*, *Analysis*, *and Interpretation*. W. W. Norton & Company, New York.

HAHN, J., TODD, P. and VAN DER KLAAUW, W. (2001). Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica* **69** 201–209.

HAINMUELLER, J., HANGARTNER, D. and PIETRANTUONO, G. (2016). Naturalization fosters the long-term political integration of immigrants. *Proc. Natl. Acad. Sci. USA* **112** 12651–12656.

HIRANO, K., IMBENS, G. W., RUBIN, D. B. and ZHOU, X.-H. (2000). Assessing the effect of an influenza vaccine in an encouragement design. *Biostatistics* **1** 69–88.

IMBENS, G. W. (2014a). Instrumental variables: An econometrician's perspective. *Statist. Sci.* **29** 323–358. MR3264545 https://doi.org/10.1214/14-STS480

IMBENS, G. (2014b). Rejoinder: "Instrumental variables: An econometrician's perspective" [MR3264546; MR3264547; MR3264548; MR3264549; MR3264545]. *Statist. Sci.* **29** 375–379. MR3264550 https://doi.org/10.1214/14-STS496

IMBENS, G. W. and ANGRIST, J. D. (1994). Identification and estimation of local average treatment effects. *Econometrica* **62** 467–475.

IMBENS, G. W. and RUBIN, D. B. (1997). Bayesian inference for causal effects in randomized experiments with noncompliance. *Ann. Statist.* **25** 305–327. MR1429927 https://doi.org/10.1214/aos/1034276631

IMBENS, G. W. and RUBIN, D. B. (2015). *Causal Inference for Statistics*, *Social*, *and Biomedical Sciences*: *An Introduction*. Cambridge Univ. Press, New York. MR3309951 https://doi.org/10.1017/CBO9781139025751

JO, B. (2002). Statistical power in randomized intervention studies with noncompliance. *Psychol. Methods* **7** 178–193.

KITAGAWA, T. (2014). Instrumental variables before and LATEr [discussion of MR3264545]. *Statist. Sci.* **29** 359–362. MR3264546 https://doi.org/10.1214/14-STS494

LIPSEY, M. W. (1990). *Design Sensitivity*: *Statistical Power for Experimental Research*. Sage Publications, Newbury Park, CA.

NEYMAN, J. (1923). Sur les applications de la théorie des probabilités aux experiences agricoles: Essai des principes. *Roczniki Nauk Rolniczych* **10** 1–51.

PIERCE, B. L., AHSAN, H. and VANDERWEELE, T. J. (2011). Power and instrument strength requirements for Mendelian randomization studies using multiple genetic variants. *Int. J. Epidemiol.* **40** 740–752.

RUBIN, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *J. Educ. Psychol.* **66** 688.

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

RUBIN, D. B. (1980). Comment: Randomization analysis of experimental data: The Fisher randomization test. *J. Amer. Statist. Assoc.* **75** 591–593.

RUBIN, D. B. (1990). Comment on J. Neyman and causal inference in experiments and observational studies: "On the application of probability theory to agricultural experiments. Essay on principles. Section 9" [Ann. Agric. Sci. **10** (1923), 1–51]. *Statist. Sci.* **5** 472–480. MR1092987

SWANSON, S. A. and HERNÁN, M. A. (2014). Think globally, act globally: An epidemiologist's perspective on instrumental variable estimation [discussion of MR3264545]. *Statist. Sci.* **29** 371–374. MR3264549 https://doi.org/10.1214/14-STS491

TSANG, R., COLLEY, L. and LYND, L. D. (2009). Inadequate statistical power to detect clinically significant differences in adverse event rates in randomized controlled trials. *J. Clin. Epidemiol.* **62** 609–616.

TVERSKY, A. and KAHNEMAN, D. (1971). The belief in the law of small numbers. *Psychol. Bull.* **76** 105–110.

WANG, X., JIANG, Y., ZHANG, N. R. and SMALL, D. S. (2018). Sensitivity analysis and power for instrumental variable studies. *Biometrics* **74** 1150–1160. MR3908133

Statistical Science
2020, Vol. 35, No. 2, 272–279
https://doi.org/10.1214/19-STS737

On the Probability That Two Random Integers Are Coprime

Jing Lei and Joseph B. Kadane

Abstract. We show that there is a nonempty class of finitely additive probabilities on $\mathbb{N}^2$ such that for each member of the class, each set with limiting relative frequency p has probability p. Hence, in that context the probability that two random integers are coprime is $6/\pi^2$. We also show that two other interpretations of "random integer," namely residue classes and shift invariance, support any number in $[0, 6/\pi^2]$ for that probability. Finally, we specify a countably additive probability space that also supports $6/\pi^2$.

Key words and phrases: Number theory, coprime, uniform distribution, finitely additive probability.

REFERENCES

ABRAMOVICH, S. and NIKITIN, Y. Y. (2017). On the probability of co-primality of two natural numbers chosen at random: From Euler identity to Haar measure on the ring of Adeles. *Bernoulli News* **24** 7–13.

ABRAMS, A. and PARIS, M. (1992). The probability that $(a, b) = 1$. *College Math. J.* **23** 47.

BHASKARA RAO, K. P. S. and BHASKARA RAO, M. (1983). *Theory of Charges*: *A Study of Finitely Additive Measures*. *Pure and Applied Mathematics* **109**. Academic Press, New York. MR0751777

BILLINGSLEY, P. (1995). *Probability and Measure*, 3rd ed. *Wiley Series in Probability and Mathematical Statistics*. Wiley, New York. MR1324786

COLLINS, G. E. and JOHNSON, J. R. (1988). The probability of relative primality of Gaussian integers. In *International Symposium on Symbolic and Algebraic Computation* 252–258. Springer. MR1034718

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

HARDY, G. H. and WRIGHT, E. M. (2008). *An Introduction to the Theory of Numbers*, 6th ed. Oxford Univ. Press, Oxford. MR2445243

KADANE, J. B. and O'HAGAN, A. (1995). Using finitely additive probability: Uniform distributions on the natural numbers. *J. Amer. Statist. Assoc.* **90** 626–631. MR1340516

KOLMOGOROFF, A. (1933). *Grundbegriffe der Wahrscheinlichkeitsrechnung*. Springer, Berlin.

KOLMOGOROV, A. N. and BHARUCHA-REID, A. T. (2018). *Foundations of the Theory of Probability*, 2nd English ed. Chelsea, New York.

KYBURG, H. E. and SMOKLER, H. E. (1980). Studies in subjective probability.

SCHIROKAUER, O. and KADANE, J. B. (2007). Uniform distributions on the natural numbers. *J. Theoret. Probab.* **20** 429–441. MR2337135 https://doi.org/10.1007/s10959-007-0066-1

Jing Lei is Associate Professor of Statistics and Data Science, 132 Baker Hall, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, USA (e-mail: jinglei@andrew.cmu.edu). Joseph B. Kadane is Leonard J. Savage University Professor of Statistics and Social Sciences, Emeritus, 132 Baker Hall, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, USA (e-mail: kadane@stat.cmu.edu).

Statistical Science
2020, Vol. 35, No. 2, 280–307
https://doi.org/10.1214/19-STS742

Comparative Study of Differentially Private Data Synthesis Methods

Claire McKay Bowen and Fang Liu

Abstract. When sharing data among researchers or releasing data for public use, there is a risk of exposing sensitive information of individuals in the data set. Data synthesis is a statistical disclosure limitation technique for releasing synthetic data sets with pseudo individual records. Traditional data synthesis techniques often rely on strong assumptions of a data intruder's behaviors and background knowledge to assess disclosure risk. Differential privacy (DP) formulates a theoretical approach for a strong and robust privacy guarantee in data release without having to model intruders' behaviors. Efforts have been made aiming to incorporate the DP concept in the data synthesis process. In this paper, we examine current DIfferentially Private Data Synthesis (DIPS) techniques for releasing individual-level surrogate data for the original data, compare the techniques conceptually and evaluate the statistical utility and inferential properties of the synthetic data via each DIPS technique through extensive simulation studies. Our work sheds light on the practical feasibility and utility of the various DIPS approaches, and suggests future research directions for DIPS.

Key words and phrases: Differential privacy, DIPS, sufficient statistics, parametric DIPS, nonparametric DIPS, statistical disclosure limitation.

REFERENCES

ABADI, M., CHU, A., GOODFELLOW, I., MCMAHAN, H. B., MIRONOV, I., TALWAR, K. and ZHANG, L. (2016). Deep learning with differential privacy. In *Proceedings of the* 2016 *ACM SIGSAC Conference on Computer and Communications Security* 308–318. ACM, New York.

ABOWD, J. M. and SCHMUTTE, I. M. (2015). Revisiting the economics of privacy: Population statistics and confidentiality protection as public goods. Cornell University ILR School. Available at https://digitalcommons.ilr.cornell.edu/ldi/22/.

ABOWD, J. M., SCHNEIDER, M. J. and VILHUBER, L. (2013). Differential privacy applications to Bayesian and linear mixed model estimation. *J. Priv. Confid.* **5** 4.

ABOWD, J. M. and VILHUBER, L. (2008). How protective are synthetic data? In *Privacy in Statistical Databases* 239–246. Springer, Berlin.

ABOWD, J., ALVISI, L., DWORK, C., KANNAN, S., MACHANAVAJJHALA, A. and REITER, J. (2017). Privacy-preserving data analysis for the Federal Statistical Agencies. Preprint. Available at arXiv:1701.00752.

ACS, G., CASTELLUCCIA, C. and CHEN, R. (2012). Differentially private histogram publishing through lossy compression. In 2012 *IEEE* 12*th International Conference on Data Mining* 1–10. IEEE, Washington, DC.

BARAK, B., CHAUDHURI, K., DWORK, C., KALE, S., MCSHERRY, F. and TALWAR, K. (2007). Privacy, accuracy, and consistency too: A holistic solution to contingency table release. In *Proceedings of the Twenty-Sixth ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems* 273–282. ACM, New York.

BARRIENTOS, A. F., BOLTON, A., BALMAT, T., REITER, J. P., DE FIGUEIREDO, J. M., MACHANAVAJJHALA, A., CHEN, Y., KNEIFEL, C. and DELONG, M. (2018). Providing access to confidential research data through synthesis and verification: An application to data on employees of the U.S. federal government. *Ann. Appl. Stat.* **12** 1124–1156. MR3834297 https://doi.org/10.1214/18-AOAS1194

BARRIENTOS, A. F., REITER, J. P., MACHANAVAJJHALA, A. and CHEN, Y. (2019). Differentially private significance tests for regression coefficients. *J. Comput. Graph. Statist.* **28** 440–453. MR3974892 https://doi.org/10.1080/10618600.2018.1538881

BLUM, A., LIGETT, K. and ROTH, A. (2013). A learning theory approach to noninteractive database privacy. *J. ACM* **60** Art. 12. MR3060810 https://doi.org/10.1145/2450142.2450148

BOWEN, C. M. K. and LIU, F. (2018). Differentially private release and analysis of youth voter registration data via statistical election to partition sequentially. Preprint. Available at arXiv:1803.06763.

BOWEN, C. M. and LIU, F. (2020). Supplement to "Comparative study of differentially private data synthesis methods." https://doi.org/10.1214/19-STS742SUPP.

CHAREST, A. S. (2010). How can we analyze differentially private synthetic datasets. *J. Priv. Confid.* **2** 3.

Claire McKay Bowen is Lead Data Scientist, Urban Institute, 500 L'Enfant Plaza SW, Washington, DC 20024, USA (e-mail: cbowen@urban.org). Fang Liu is Associate Professor, Department of Applied and Computational Mathematics and Statistics, University of Notre Dame, 201B Crowley Hall, Notre Dame, Indiana 46556, USA (e-mail: fang.liu.131@nd.edu).

CHAREST, A.-S. and HOU, Y. (2017). On the meaning and limits of empirical differential privacy. *J. Priv. Confid.* **7** 3.

CHAUDHURI, K. and MONTELEONI, C. (2009). Privacy-preserving logistic regression. In *Advances in Neural Information Processing Systems* 289–296.

CHAUDHURI, K., MONTELEONI, C. and SARWATE, A. D. (2011). Differentially private empirical risk minimization. *J. Mach. Learn. Res.* **12** 1069–1109. MR2786918

CHAUDHURI, K., SARWATE, A. and SINHA, K. (2012). Near-optimal differentially private principal components. In *Advances in Neural Information Processing Systems* 989–997.

CHEN, R., FUNG, B. C., MOHAMMED, N., DESAI, B. C. and WANG, K. (2013). Privacy-preserving trajectory data publishing by local suppression. *Inform. Sci.* **231** 83–97.

DING, B., WINSLETT, M., HAN, J. and LI, Z. (2011). Differentially private data cubes: Optimizing noise sources and consistency. In *Proceedings of the ACM SIGMOD International Conference on Management of Data* (*SIGMOD* 2011) 217–228. ACM, New York.

DRECHSLER, J. (2011). *Synthetic Datasets for Statistical Disclosure Control*: *Theory and Implementation*. *Lecture Notes in Statistics* **201**. Springer, New York. MR2809912 https://doi.org/10.1007/978-1-4614-0326-5

DUCHI, J. C., JORDAN, M. I. and WAINWRIGHT, M. J. (2013). Local privacy and statistical minimax rates. In 2013 *IEEE* 54*th Annual Symposium on Foundations of Computer Science—FOCS* 2013 429–438. IEEE Computer Soc., Los Alamitos, CA. MR3246246 https://doi.org/10.1109/FOCS.2013.53

DWORK, C. (2008). Differential privacy: A survey of results. In *Theory and Applications of Models of Computation*. *Lecture Notes in Computer Science* **4978** 1–19. Springer, Berlin. MR2472670 https://doi.org/10.1007/978-3-540-79228-4_1

DWORK, C. (2011). Differential privacy. In *Encyclopedia of Cryptography and Security* 338–340. Springer, Berlin.

DWORK, C. and ROTH, A. (2014). The algorithmic foundations of differential privacy. *Found. Trends Theor. Comput. Sci.* **9** 211–407. MR3254020 https://doi.org/10.1561/0400000042

DWORK, C. and ROTHBLUM, G. N. (2016). Concentrated differential priacy. Preprint. Available at arXiv:1603.01887.

DWORK, C., MCSHERRY, F., NISSIM, K. and SMITH, A. (2006a). Calibrating noise to sensitivity in private data analysis. In *Theory of Cryptography*. *Lecture Notes in Computer Science* **3876** 265–284. Springer, Berlin. MR2241676 https://doi.org/10.1007/11681878_14

DWORK, C., KENTHAPADI, K., MCSHERRY, F., MIRONOV, I. and NAOR, M. (2006b). Our data, ourselves: Privacy via distributed noise generation. In *Advances in Cryptology—EUROCRYPT* 2006. *Lecture Notes in Computer Science* **4004** 486–503. Springer, Berlin. MR2423560 https://doi.org/10.1007/11761679_29

ERLINGSSON, Ú., PIHUR, V. and KOROLOVA, A. (2014). Rappor: Randomized aggregatable privacy-preserving ordinal response. In *Proceedings of the* 2014 *ACM SIGSAC Conference on Computer and Communications Security* 1054–1067. ACM, New York.

FANTI, G., PIHUR, V. and ERLINGSSON, Ú. (2016). Building a RAPPOR with the unknown: Privacy-preserving learning of associations and data dictionaries. *Proc. Priv. Enhanc. Technol.* **2016** 41–61.

FRIEDMAN, A., BERKOVSKY, S. and KAAFAR, M. A. (2016). A differential privacy framework for matrix factorization recommender systems. *User Model. User-Adapt. Interact.* **26** 425–458.

GABOARDI, M., HONAKER, J., KING, G., NISSIM, K., ULLMAN, J. and VADHAN, S. (2016). PSI (Ψ): A private data sharing interface. Preprint. Available at arXiv:1609.04340.

GARDNER, J., XIONG, L., XIAO, Y., GAO, J., POST, A. R., JIANG, X. and OHNO-MACHADO, L. (2013). SHARE: System design and case studies for statistical health information release. *J. Am. Med. Inform. Assoc.* **20** 109–116.

GIL, D., GIRELA, J. L., DE JUAN, J., GOMEZ-TORRES, M. J. and JOHNSSON, M. (2012). Predicting seminal quality with artificial intelligence methods. *Expert Syst. Appl.* **39** 12564–12573.

GÖTZ, M., MACHANAVAJJHALA, A., WANG, G., XIAO, X. and GEHRKE, J. (2012). Publishing search logs—a comparative study of privacy guarantees. *IEEE Trans. Knowl. Data Eng.* **24** 520–532.

HALL, R., RINALDO, A. and WASSERMAN, L. (2013). Differential privacy for functions and functional data. *J. Mach. Learn. Res.* **14** 703–727. MR3033345

HARDT, M., LIGETT, K. and MCSHERRY, F. (2012). A simple and practical algorithm for differentially private data release. In *Advances in Neural Information Processing Systems* 2339–2347.

HARDT, M. and ROTHBLUM, G. N. (2010). A multiplicative weights mechanism for privacy-preserving data analysis. In 2010 *IEEE* 51*st Annual Symposium on Foundations of Computer Science—FOCS* 2010 61–70. IEEE Computer Soc., Los Alamitos, CA. MR3024776

HARDT, M. and TALWAR, K. (2010). On the geometry of differential privacy. In *STOC'*10*—Proceedings of the* 2010 *ACM International Symposium on Theory of Computing* 705–714. ACM, New York. MR2743320

HAY, M., RASTOGI, V., MIKLAU, G. and SUCIU, D. (2010). Boosting the accuracy of differentially private histograms through consistency. *Proc. VLDB Endow.* **3** 1021–1032.

HAY, M., MACHANAVAJJHALA, A., MIKLAU, G., CHEN, Y. and ZHANG, D. (2016). Principled evaluation of differentially private algorithms using dpbench. In *Proceedings of the* 2016 *International Conference on Management of Data* 139–154. ACM, New York.

HE, X., CORMODE, G., MACHANAVAJJHALA, A., PROCOPIUC, C. M. and SRIVASTAVA, D. (2015). DPT: Differentially private trajectory synthesis using hierarchical reference systems. *Proc. VLDB Endow.* **8** 1154–1165.

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

HOMER, N., SZELINGER, S., REDMANN, M., DUGGAN, D., TEMBE, W., MUEHLING, J., PEARSON, J. V., STEPHAN, D. A., NELSON, S. F. et al. (2008). Resolving individuals contributing trace amounts of DNA to highly complex mixtures using high-density SNP genotyping microarrays. *PLoS Genet.* **4** e1000167.

HUNDEPOOL, A., DOMINGO-FERRER, J., FRANCONI, L., GIESSING, S., NORDHOLT, E. S., SPICER, K. and DE WOLF, P.-P. (2012). *Statistical Disclosure Control*. *Wiley Series in Survey Methodology*. Wiley, Chichester. MR3026260 https://doi.org/10.1002/9781118348239

KARWA, V., KRIVITSKY, P. N. and SLAVKOVIĆ, A. B. (2017). Sharing social network data: Differentially private estimation of exponential family random-graph models. *J. R. Stat. Soc. Ser. C. Appl. Stat.* **66** 481–500. MR3632338 https://doi.org/10.1111/rssc.12185

KERMAN, J. (2011). Neutral noninformative and informative conjugate beta and gamma prior distributions. *Electron. J. Stat.* **5** 1450–1470. MR2851686 https://doi.org/10.1214/11-EJS648

KIFER, D., SMITH, A. and THAKURTA, A. (2012). Private convex empirical risk minimization and high-dimensional regression. *J. Mach. Learn. Res. Workshop Conf. Proc.* **23** 1–40.

KINNEY, S. K., REITER, J. P., REZNEK, A. P., MIRANDA, J., JARMIN, R. S. and ABOWD, J. M. (2011). Towards unrestricted public use business microdata: The synthetic longitudinal business database. *Int. Stat. Rev.* **79** 362–384.

KOTSOGIANNIS, I., MACHANAVAJJHALA, A., HAY, M. and MIKLAU, G. (2017). Pythia: Data dependent differentially private algorithm selection. In *Proceedings of the* 2017 *ACM International Conference on Management of Data* 1323–1337. ACM, New York.

KOWALCZYK, L., MALKIN, T., ULLMAN, J. and WICHS, D. (2017). Hardness of non-interactive differential privacy from one-way functions. *Semantic Scholar*. Available at https://pdfs.semanticscholar.org/5a40/c47078efdd3d6b6d15f323d6c3bc0d709ea8.pdf.

LEE, J. and CLIFTON, C. (2011). How much is enough? Choosing for differential privacy. In *Information Security*: *Proceedings of the* 14*th*, *International Conference*, *ISC* 2011 (X. Lai, J. Zhou and H. Li, eds.) 325–340. ACM, New York.

LEI, J., CHAREST, A.-S., SLAVKOVIC, A., SMITH, A. and FIENBERG, S. (2018). Differentially private model selection with penalized and constrained likelihood. *J. Roy. Statist. Soc. Ser. A* **181** 609–633. MR3807500 https://doi.org/10.1111/rssa.12324

LI, H., XIONG, L. and JIANG, X. (2014a). Differentially private synthesization of multi-dimensional data using copula functions. In *Advances in Database Technology*: *Proceedings*. *International Conference on Extending Database Technology* **2014** 475. NIH Public Access.

LI, H., XIONG, L. and JIANG, X. (2014b). Differentially private synthesization of multi-dimensional data using copula functions. In *Proc*. 17*th International Conference on Extending Database Technology* (*EDBT*) 475–486.

LI, N., LYU, M., SU, D. and YANG, W. (2016). Differential privacy: From theory to practice. *Synth. Lect. Inf. Secur. Priv. Trust* **8** 1–138.

LITTLE, R. J. A. (1993). Statistical analysis of masked data. *J. Off. Stat*. **9** 407.

LITTLE, R. J. A., LIU, F. and RAGHUNATHAN, T. E. (2004). Statistical disclosure techniques based on multiple imputation. In *Applied Bayesian Modeling and Causal Inference from Incomplete-Data Perspectives Wiley Ser. Probab. Stat*. 141–152. Wiley, Chichester. MR2138251 https://doi.org/10.1002/0470090456.ch13

LIU, F. (2016). Model-based differentially private data synthesis. Preprint. Available at arXiv:1606.08052.

LIU, F. (2019a). Generalized Gaussian mechanism for differential privacy. *IEEE Trans. Knowl. Data Eng*. **31** 747–756.

LIU, F. (2019b). Statistical properties of sanitized results from differentially private Laplace mechanisms with noninformative bounding. *Trans. Data Priv.* **12** 169–195.

LIU, F. and LITTLE, R. J. A. (2003). SMIKe vs. data swapping and PRAM for statistical disclosure limitation in microdata: A simulation study. In *Proceedings of* 2003 *American Statistical Association Joint Statistical Meeting*.

MACHANAVAJJHALA, A., KIFER, D., ABOWD, J., GEHRKE, J. and VILHUBER, L. (2008). Privacy: Theory meets practice on the map. In *IEEE ICDE IEEE* 24*th International Conference* 277–286.

MACLEOD, A. J. and HENDERSON, G. R. (1984). Bounds for the sample standard deviation. *Teach. Stat*. **6** 72–76.

MANRIQUE-VALLIER, D. and REITER, J. P. (2012). Estimating identification disclosure risk using mixed membership models. *J. Amer. Statist. Assoc*. **107** 1385–1394. MR3036402 https://doi.org/10.1080/01621459.2012.710508

MCCLURE, D. and REITER, J. P. (2012). Differential privacy and statistical disclosure risk measures: An investigation with binary synthetic data. *Trans. Data Priv.* **5** 535–552. MR3018910

MCSHERRY, F. (2009). Privacy integrated queries: An extensible platform for privacy-preserving data analysis. In *Proceedings of the* 2009 *ACM SIGMOD International Conference on Management of Data* 19–30. ACM, New York.

MCSHERRY, F. and TALWAR, K. (2007). Mechanism design via differential privacy. In *Foundations of Computer Science*, 2007. *FOCS'*07. 48*th Annual IEEE Symposium* 94–103. IEEE, Washington, DC.

MOHAMMED, N., CHEN, R., FUNG, B. and YU, P. S. (2011). Differentially private data release for data mining. In *Proceedings of the* 17*th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 493–501. ACM, New York.

NARAYANAN, A. and SHMATIKOV, V. (2008). Robust de anonymization of large sparse datasets. In *IEEE Symposium on Security and Privacy* 111–125.

NISSIM, K., RASKHODNIKOVA, S. and SMITH, A. (2007). Smooth sensitivity and sampling in private data analysis. In *STOC'*07*—Proceedings of the* 39*th Annual ACM Symposium on Theory of Computing* 75–84. ACM, New York. MR2402430 https://doi.org/10.1145/1250790.1250803

NISSIM, K. and STEMMER, U. (2015). On the generalization properties of differential privacy. *CoRR*, Abs/1504.05800.

ORR, A. (2017). Google's differential privacy may be better than Apple's. https://www.macobserver.com/analysis/google-apple-differential-privacy/.

PROSERPIO, D., GOLDBERG, S. and MCSHERRY, F. (2012). A workflow for differentially-private graph synthesis. In *Proceedings of the* 2012 *ACM Workshop on Workshop on Online Social Networks* 13–18. ACM, New York.

QARDAJI, W., YANG, W. and LI, N. (2013). Understanding hierarchical methods for differentially private histograms. *Proc. VLDB Endow.* **6** 1954–1965.

RAAB, G. M., NOWOK, B. and DIBBEN, C. (2017). Practical data synthesis for large samples. *J. Priv. Confid*. **7** 67–97.

RAGHUNATHAN, T. E., REITER, J. P. and RUBIN, D. B. (2003). Multiple imputation for statistical disclosure limitation. *J. Off. Stat*. **19** 1–16.

RAGHUNATHAN, T. E., SOLENBERGER, P. W. and HOEWYK, J. V. (2017). IVEware: Imputation and Variance Estimation software for SRMI. Available at http://www.isr.umich.edu/src/smp/ive (accessed December 20, 2018).

RAGHUNATHAN, T. E., LEPKOWSKI, J. M., VANHOEWYK, J. and SOLENBERGER, P. (2001). A multivariate technique for multiply imputing missing values using a sequence of regression models. *Surv. Methodol*. **27** 85–95.

REITER, J. P. (2002). Satisfying disclosure restrictions with synthetic data sets. *J. Off. Stat*. **18** 531–543.

REITER, J. P. (2003). Inference for partially synthetic, public use microdata sets. *Surv. Methodol*. **29** 181–188.

REITER, J. P. (2005). Estimating risks of identification disclosure in microdata. *J. Amer. Statist. Assoc*. **100** 1103–1112. MR2236926 https://doi.org/10.1198/016214505000000619

REITER, J. P. (2009). Using multiple imputation to integrate and disseminate confidential microdata. *Int. Stat. Rev.* **77** 179–195.

REITER, J. P. and KINNEY, S. K. (2012). Inferentially valid, partially synthetic data: Generating from posterior predictive distributions not necessary. *J. Off. Stat*. **28** 583–590.

ROTH, A. and ROUGHGARDEN, T. (2010). Interactive privacy via the median mechanism. In *STOC'*10*—Proceedings of the* 2010 *ACM International Symposium on Theory of Computing* 765–774. ACM, New York. MR2743326

RUBIN, D. B. (1993). Discussion statistical disclosure limitation. *J. Off. Stat*. **9** 461–468.

RUBINSTEIN, B. I., BARTLETT, P. L., HUANG, L. and TAFT, N. (2009). Learning in a large function space: Privacy-preserving mechanisms for SVM learning. Preprint. Available at arXiv:0911.5708.

SCOTT, D. W. (2015). *Multivariate Density Estimation*: *Theory*, *Practice*, *and Visualization*, 2nd ed. *Wiley Series in Probability and Statistics*. Wiley, Hoboken, NJ. MR3329609

SHEFFET, O. (2017). Differentially private ordinary least squares. In *Proceedings of the* 34*th International Conference on Machine Learning—Vol*. 70 3105–3114. JMLR.org.

SMITH, A. and THAKURTA, A. (2013). Differentially private model selection via stability arguments and the robustness of the Lasso. *J. Mach. Learn. Res. Workshop Conf. Proc*. **30** 1–31.

STEINKE, T. and ULLMAN, J. (2017). Between pure and approximate differential privacy. *J. Priv. Confid*. **7** 2.

SWEENEY, L. (2013). Matching known patients to health records in Washington state data. *Social Science Research Network*. id=2289850.

TANG, J., KOROLOVA, A., BAI, X., WANG, X. and WANG, X. (2017). Privacy loss in Apple's implementation of differential privacy on macOS 10.12. Preprint. Available at arXiv:1709.02753.

VADHAN, S. (2017). The complexity of differential privacy. In *Tutorials on the Foundations of Cryptography*. *Inf. Secur. Cryptography* 347–450. Springer, Cham. MR3837668

WANG, Y.-X., FIENBERG, S. and SMOLA, A. (2015). Privacy for free: Posterior sampling and stochastic gradient Monte Carlo. In *International Conference on Machine Learning* 2493–2502.

WANG, Q., ZHANG, Y., LU, X., WANG, Z., QIN, Z. and REN, K. (2016). RescueDP: Real-time spatio-temporal crowd-sourced data publishing with differential privacy. In *Computer Communications, IEEE INFOCOM* 2016*—the* 35*th Annual IEEE International Conference on* 1–9. IEEE, Washington, DC.

WASSERMAN, L. and ZHOU, S. (2010). A statistical framework for differential privacy. *J. Amer. Statist. Assoc*. **105** 375–389. MR2656057 https://doi.org/10.1198/jasa.2009.tm08651

XIAO, Y., GARDNER, J. and XIONG, L. (2012). Dpcube: Releasing differentially private data cubes for health information. In 2012 *IEEE* 28*th International Conference on Data Engineering* 1305–1308. IEEE, Washington, DC.

XIAO, Y. and XIONG, L. (2015). Protecting locations with differential privacy under temporal correlations. In *Proceedings of the* 22*nd ACM SIGSAC Conference on Computer and Communications Security* 1298–1309. ACM, New York.

XU, J., ZHANG, Z., XIAO, X., YANG, Y., YU, G. and WINSLETT, M. (2013). Differentially private histogram publication. *VLDB J*. **22** 797–822.

YU, F., FIENBERG, S. E., SLAVKOVIC, A. B. and UHLER, C. (2014). Scalable privacy-preserving data sharing methodology for genome-wide association studies. *J. Biomed. Inform*. **50** 133–141.

ZHANG, J., ZHANG, Z., XIAO, X., YANG, Y. and WINSLETT, M. (2012). Functional mechanism: Regression analysis under differential privacy. *Proc. VLDB Endow.* **5** 1364–1375.

ZHANG, J., XIAO, X., YANG, Y., ZHANG, Z. and WINSLETT, M. (2013). PrivGene: Differentially private model fitting using genetic algorithms. In *Proceedings of the* 2013 *ACM SIGMOD International Conference on Management of Data* 665–676. ACM, New York.

ZHANG, J., CORMODE, G., PROCOPIUC, C. M., SRIVASTAVA, D. and XIAO, X. (2017). PrivBayes: Private data release via Bayesian networks. *ACM Trans. Database Syst*. **42** Art. 25. MR3730676 https://doi.org/10.1145/3134428

Statistical Science
2020, Vol. 35, No. 2, 308–320
https://doi.org/10.1214/19-STS734

A Conversation with Grace Wahba

Douglas Nychka, Ping Ma and Douglas Bates

Abstract. Grace Wahba (née Goldsmith, born August 3, 1934), I. J. Schoenberg-Hilldale Professor of Statistics at the University of Wisconsin-Madison (Emerita), is a pioneer in methods for smoothing noisy data. Her research combines theoretical analysis, computation and methodology motivated by innovative scientific applications. Best known for the development of generalized cross-validation (GCV), the connection between splines and Bayesian posterior estimates, and "Wahba's problem," she has developed methods with applications in demographic studies, machine learning, DNA microarrays, risk modeling, medical imaging and climate prediction.

Grace grew up in the Washington, DC area and New Jersey, and graduated from Montclair High School. She was educated at Cornell (B.A. 1956), University of Maryland, College Park (M.A. 1962) and Stanford (Ph.D. 1966), and worked in industry for several years before receiving her doctorate in 1966 and settling in Madison in 1967. Although holding several visiting appointments, she has made Madison her home for over 50 years. She is the author of *Spline Models for Observational Data* which has garnered more than 8000 citations. Grace is treasured as an academic advisor and has mentored 39 Ph.D. students that have resulted in more than 330 academic descendants. She was elected to the United States National Academy of Sciences in 2000 and received an honorary degree of Doctor of Science from the University of Chicago in 2007. Wahba is a Fellow of several academic societies including the American Academy of Arts and Sciences, the American Association for the Advancement of Science, the American Statistical Association and the Institute of Mathematical Statistics. Over the years, she has received a selection of notable awards in the statistics community: R. A. Fisher Lectureship (2014), Gottfried E. Noether Senior Researcher Award (2009), Committee of Presidents of Statistical Societies Elizabeth Scott Award (1996) and the first Emanuel and Carol Parzen Prize for Statistical Innovation (1994).

Key words and phrases: Spline, reproducing kernel Hilbert spaces, RKHS, Wahba's problem, cross-validation, generalized cross-validation, support vector machines.

REFERENCES

[1] Box, G. E. P., Hunter, W. G. and Hunter, J. S. (1978). *Statistics for Experimenters*: *An Introduction to Design*, *Data Analysis*, *and Model Building*. *Wiley Series in Probability and Mathematical Statistics*. Wiley, New York. MR0483116

[2] Craven, P. and Wahba, G. (1978/79). Smoothing noisy data with spline functions. Estimating the correct degree of smoothing by the method of generalized cross-validation. *Numer. Math.* **31** 377–403. MR0516581 https://doi.org/10.1007/BF01404567

[3] Golub, G. H., Heath, M. and Wahba, G. (1979). Generalized cross-validation as a method for choosing a good ridge parameter. *Technometrics* **21** 215–223. MR0533250 https://doi.org/10.2307/1268518

[4] Gong, J., Wahba, G., Johnson, D. and Tribbia, J. (1998). Adaptive tuning of numerical weather prediction models: Simultaneous estimation of weighting, smoothing and physical parameters. *Mon. Weather Rev.* **126** 210–231.

[5] Kimeldorf, G. and Wahba, G. (1971). Some results on Tchebycheffian spline functions. *J. Math. Anal. Appl.* **33** 82–95. MR0290013 https://doi.org/10.1016/0022-247X(71)90184-3

[6] Kimeldorf, G. S. and Wahba, G. (1970). A correspondence between Bayesian estimation on stochastic processes and smoothing by splines. *Ann. Math. Stat.* **41** 495–502. MR0254999

Douglas Nychka is Professor, Department of Applied Mathematics and Statistics, Colorado School of Mines, 1500 Illinois St., Golden, Colorado 80401, USA (e-mail: nychka@mines.edu). Ping Ma is Professor, Department of Statistics, 310 Herty Drive, University of Georgia, Athens, Georgia 30602, USA (e-mail: pingma@uga.edu). Douglas Bates is Emeritus Professor, Department of Statistics, University of Wisconsin, 1300 University Ave, Madison, Wisconsin 53706-1685, USA (e-mail: bates@stat.wisc.edu).

https://doi.org/10.1214/aoms/1177697089

[7] KIMELDORF, G. S. and WAHBA, G. (1970). Spline functions and stochastic processes. *Sankhya*, *Ser. A* **32** 173–180. MR0303594

[8] LEE, Y., LIN, Y. and WAHBA, G. (2004). Multicategory support vector machines: Theory and application to the classification of microarray data and satellite radiance data. *J. Amer. Statist. Assoc.* **99** 67–81. MR2054287 https://doi.org/10.1198/016214504000000098

[9] LIN, Y., LEE, Y. and WAHBA, G. (2002). Support vector machines for classification in nonstandard situations. *Mach. Learn.* **46** 191–202.

[10] LIN, Y., WAHBA, G., ZHANG, H. and LEE, Y. (2002). Statistical properties and adaptive tuning of support vector machines. *Mach. Learn.* **48** 115–136.

[11] NYCHKA, D., WAHBA, G., GOLDFARB, S. and PUGH, T. (1984). Cross-validated spline methods for the estimation of three-dimensional tumor size distributions from observations on two-dimensional cross sections. *J. Amer. Statist. Assoc.* **79** 832–846. MR0770276

[12] O'SULLIVAN, F. and WAHBA, G. (1985). A cross validated Bayesian retrieval algorithm for non-linear remote sensing. *J. Comput. Phys.* **59** 441–455.

[13] PARZEN, E. (1960). *Modern Probability Theory and Its Applications. A Wiley Publication in Mathematical Statistics*. Wiley, New York. MR0112166

[14] PARZEN, E. (1962). *Stochastic Processes. Holden-Day Series in Probability and Statistics*. Holden-Day, San Francisco, CA. MR0139192

[15] VAN DER WAERDEN, B. L. (1953). *Algebra. Vol* 1. Ungar, New York.

[16] WAHBA, G. (1965). Problem 65-1, a least squares estimate of satellite attitude: Problem proposal. *SIAM Rev.* **7** 409.

[17] WAHBA, G. (1966). Problem 65-1, a least squares estimate of satellite attitude. *SIAM Rev.* **8** 384–386.

[18] WAHBA, G. (1977). Practical approximate solutions to linear operator equations when the data are noisy. *SIAM J. Numer. Anal.* **14** 651–667. MR0471299 https://doi.org/10.1137/0714044

[19] WAHBA, G. (1987). Three topics in ill-posed problems. In *Inverse and Ill-Posed Problems* (*Sankt Wolfgang*, 1986). *Notes Rep. Math. Sci. Engrg.* **4** 37–51. Academic Press, Boston, MA. MR1020307 https://doi.org/10.1016/B978-0-12-239040-1.50008-9

[20] WAHBA, G., LIN, Y., LEE, Y. and ZHANG, H. (2003). Optimal properties and adaptive tuning of standard and nonstandard support vector machines. In *Nonlinear Estimation and Classification* (*Berkeley*, *CA*, 2001). *Lect. Notes Stat.* **171** 129–147. Springer, New York. MR2005787 https://doi.org/10.1007/978-0-387-21579-2_8

[21] WAHBA, G. and WENDELBERGER, J. (1980). Some new mathematical methods for variational objective analysis using splines and cross-validation. *Mon. Weather Rev.* **108** 1122–1145.

[22] WAHBA, G. and WOLD, S. (1975). A completely automatic French curve: Fitting spline functions by cross validation. *Commun. Stat.* **4** 1–17. MR0655432 https://doi.org/10.1080/03610917508548493

Statistical Science
2020, Vol. 35, No. 2, 321–333
https://doi.org/10.1214/19-STS764

A Conversation with Francisco J. Samaniego

George G. Roussas and Debasis Bhattacharya

Abstract. In a wide-ranging interview, Professor George G. Roussas of the University of California, Davis, and Professor Debasis Bhatacharya of Visva-Bharati University, India, and a frequent visitor to UC Davis, engage Professor Francisco J. Samaniego in a discussion about his personal background, his education and the highlights of his professional career as a teacher and researcher.

Key words and phrases: Education, teaching and research, reliability theory, comparative statistical inference.

REFERENCES

[1] ROUSSAS, G. G. and BHATTACHARYA, D. (2020). Supplement to "A Conversation with Francisco J. Samaniego." https://doi.org/10.1214/19-STS764SUPP

George G. Roussas is a Distinguished Professor Emeritus of Statistics, University of California, Davis, Davis, California 95616, USA (e-mail: ggroussas@ucdavis.edu). Debasis Bhattacharya is a Professor of Statistics, Visva-Bharati University, Santiniketan, India (e-mail: debasis_us@yahoo.com).

INSTITUTE OF MATHEMATICAL STATISTICS

(Organized September 12, 1935)

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

IMS OFFICERS

President: Susan Murphy, Department of Statistics, Harvard University, Cambridge, Massachusetts 02138-2901, USA

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

Past President: Xiao-Li Meng, Department of Statistics, Harvard University, Cambridge, Massachusetts 02138-2901, USA

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

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

Program Secretary: Ming Yuan, Department of Statistics, Columbia University, New York, NY 10027-5927, USA

IMS EDITORS

The Annals of Statistics. *Editors*: Richard J. Samworth, Statistical Laboratory, Centre for Mathematical Sciences, University of Cambridge, Cambridge, CB3 0WB, UK. Ming Yuan, Department of Statistics, Columbia University, New York, NY 10027, USA

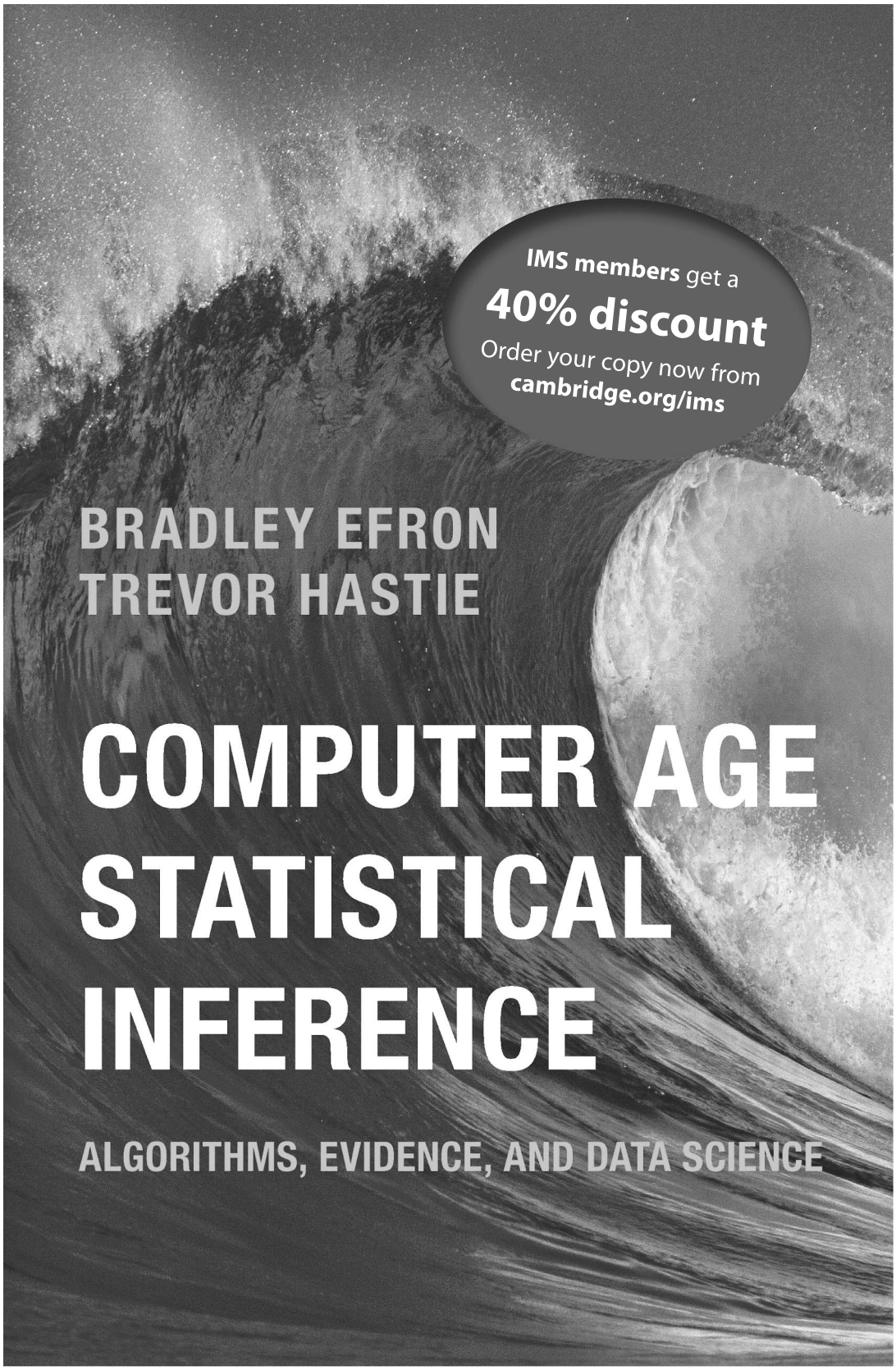
The Annals of Applied Statistics. *Editor-in-Chief*: Karen Kafadar, Department of Statistics, University of Virginia, Heidelberg Institute for Theoretical Studies, Charlottesville, VA 22904-4135, USA

The Annals of Probability. *Editor*: Amir Dembo, Department of Statistics and Department of Mathematics, Stanford University, Stanford, California 94305, USA

The Annals of Applied Probability. *Editors*: François Delarue, Laboratoire J. A. Dieudonné, Université de Nice Sophia-Antipolis, France-06108 Nice Cedex 2. Peter Friz, Institut für Mathematik, Technische Universität Berlin, 10623 Berlin, Germany and Weierstrass-Institut für Angewandte Analysis und Stochastik, 10117 Berlin, Germany

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

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


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