Contents

Proceedings of XIII Brazilian School of Bayesian Statistics

Preface 667

S. L. SCOTT
Comparing consensus Monte Carlo strategies for distributed Bayesian computation 668

D. DRAPER and A. TERENIN
Comment: A brief survey of the current state of play for Bayesian computation in data science at big-data scale 686

A. GELMAN and A. VEHTARI
Comment: Consensus Monte Carlo using expectation propagation 692

S. L. SCOTT
Rejoinder 697

J. NAKAJIMA and M. WEST
Dynamics & sparsity in latent threshold factor models: A study in multivariate EEG signal processing 701

F. B. GONÇALVES, K. ŁATUSZYŃSKI and G. O. ROBERTS
Barker's algorithm for Bayesian inference with intractable likelihoods 732

K. S. CONCEIÇÃO, A. K. SUZUKI, M. G. ANDRADE and F. LOUZADA
A Bayesian approach for a zero modified Poisson model to predict match outcomes applied to the 2012–13 La Liga season 746

L. COSTA, T. NICHOLS and J. Q. SMITH
Studying the effective brain connectivity using multiregression dynamic models 765

F. FERRAZ DO NASCIMENTO and M. BOURGUIGNON PEREIRA
A Bayesian approach to extended models for exceedance 801

BJPS is an IMS supported journal

ISSN 0103-0752

Brazilian Statistical Association

Brazilian Journal of Probability and Statistics

Volume 31 • Number 4 • November 2017

Brazilian Journal of Probability and Statistics

Volume 31 • Number 4 • November 2017

Proceedings of XIII Brazilian School of Bayesian Statistics

www.imstat.org/bjps

ISSN 0103-0752 (Print) ISSN 2317-6199 (Online), Volume 31, Number 4, November 2017. Published quarterly by the Brazilian Statistical Association.

POSTMASTER:

Send address changes to Brazilian Journal of Probability and Statistics, Institute of Mathematical Statistics, Dues and Subscriptions Office, 9650 Rockville Pike, Suite L 2310, Bethesda, Maryland 20814-3998, USA.

Brazilian Statistical Association members should send address changes to Rua do Matão, 1010 sala 250A, 05508-090 São Paulo/SP Brazil (address of the BSA office).

Printed in the United States of America

Partial financial support:
CNPq and CAPES (Brazil).

Brazilian Journal of Probability and Statistics
2017, Vol. 31, No. 4, 667
https://doi.org/10.1214/17-BJPS314PRE

Preface

Brazilian Journal of Probability and Statistics
2017, Vol. 31, No. 4, 668–685
https://doi.org/10.1214/17-BJPS365

Comparing consensus Monte Carlo strategies for distributed Bayesian computation[1]

Steven L. Scott

Google

Abstract. Consensus Monte Carlo is an algorithm for conducting Monte Carlo based Bayesian inference on large data sets distributed across many worker machines in a data center. The algorithm operates by running a separate Monte Carlo algorithm on each worker machine, which only sees a portion of the full data set. The worker-level posterior samples are then combined to form a Monte Carlo approximation to the full posterior distribution based on the complete data set. We compare several methods of carrying out the combination, including a new method based on approximating worker-level simulations using a mixture of multivariate Gaussian distributions. We find that resampling and kernel density based methods break down after 10 or sometimes fewer dimensions, while the new mixture-based approach works well, but the necessary mixture models take too long to fit.

References

Ahn, S., Shahbaba, B., Welling, M., et al. (2014). Distributed stochastic gradient MCMC. In *ICML* 1044–1052.

Bardenet, R., Doucet, A. and Holmes, C. (2014). Towards scaling up Markov chain Monte Carlo: An adaptive subsampling approach. In *Proceedings of the 31st International Conference on Machine Learning (ICML-14)* (T. Jebara and E. P. Xing, eds.) 405–413. JMLR Workshop and Conference Proceedings.

Blei, D. M., Jordan, M. I., et al. (2006). Variational inference for Dirichlet process mixtures. *Bayesian Analysis* **1**, 121–144. MR2227367

Chang, F., Dean, J., Ghemawat, S., Hsieh, W. C., Wallach, D. A., Burrows, M., Chandra, T., Fikes, A. and Gruber, R. E. (2008). Bigtable: A distributed storage system for structured data. *ACM Transactions on Computer Systems* **26**, 4.

Chen, H., Seita, D., Pan, X. and Canny, J. (2016). An efficient minibatch acceptance test for Metropolis–Hastings. arXiv preprint, available at arXiv:1610.06848v1.

Dean, J. and Ghemawat, S. (2008). MapReduce: Simplified data processing on large clusters. *Communications of the ACM* **51**, 107–113.

Doucet, A., De Frietas, N. and Gordon, N. (2001). *Sequential Monte Carlo in Practice*. Springer. MR1847783

Lee, A., Yao, C., Giles, M. B., Doucet, A. and Holmes, C. C. (2010). On the utility of graphics cards to perform massively parallel simulation of advanced Monte Carlo methods. *Journal of Computational and Graphical Statistics* **19**, 769–789.

Maclaurin, D. and Adams, R. P. (2014). Firefly Monte Carlo: Exact MCMC with subsets of data. Available at arXiv:1403.5693.

Key words and phrases. Big data, cloud computing, Bayesian modeling.

McGrory, C. A. and Titterington, D. M. (2007). Variational approximations in Bayesian model selection for finite mixture distributions. *Computational Statistics & Data Analysis* **51**, 5352–5367. MR2370876

McLachlan, G. J. and Peel, D. (2000). *Finite Mixture Models*. New York: John Wiley & Sons. MR1789474

McLachlan, G. J., Peel, D. and Bean, R. W. (2003). Modelling high-dimensional data by mixtures of factor analyzers. *Computational Statistics & Data Analysis* **41**, 379–388. MR1973720

Miroshnikov, A. and Conlon, E. (2014). parallelMCMCcombine: Methods for combining independent subset Markov chain Monte Carlo (MCMC) posterior samples to estimate a posterior density given the full data set. R package version 1.0. Available at http://CRAN.R-project.org/package=parallelMCMCcombine.

Neal, R. M. (2000). Markov chain sampling methods for Dirichlet process mixture models. *Journal of Computational and Graphical Statistics* **9**, 249–265. MR1823804

Neiswanger, W., Wang, C. and Xing, E. (2013). Asymptotically exact, embarrassingly parallel MCMC. arXiv preprint, available at arXiv:1311.4780.

Quiroz, M., Villani, M. and Kohn, R. (2016). Exact subsampling MCMC. Available at arXiv:1603.08232.

Rousseau, J. and Mengersen, K. (2011). Asymptotic behaviour of the posterior distribution in overfitted mixture models. *Journal of the Royal Statistical Society, Series B, Statistical Methodology* **73**, 689–710. MR2867454

Scott, D. W. and Sain, S. R. (2005). Multidimensional density estimation. In *Handbook of Statistics* **24**, 229–261. Elsevier.

Scott, S. L. (2010). A modern Bayesian look at the multi-armed bandit. *Applied Stochastic Models in Business and Industry* **26**, 639–658 (with discussion). MR2752378

Scott, S. L. (2015). Multi-armed bandit experiments in the online service economy. *Applied Stochastic Models in Business and Industry* **31**, 37–45. MR3326375

Scott, S. L., Blocker, A. W., Bonassi, F. V., Chipman, H. A., George, E. I. and McCulloch, R. E. (2016). Bayes and big data: The consensus Monte Carlo algorithm. *International Journal of Management Science and Engineering Management* **11**, 78–88.

Srivastava, S., Li, C. and Dunson, D. B. (2015). Scalable Bayes via barycenter in Wasserstein space. arXiv preprint, available at arXiv:1508.05880.

Suchard, M. A., Wang, Q., Chan, C., Frelinger, J., Cron, A. and West, M. (2010). Understanding GPU programming for statistical computation: Studies in massively parallel massive mixtures. *Journal of Computational and Graphical Statistics* **19**, 419–438. MR2758309

Tadesse, M. G., Sha, N. and Vannucci, M. (2005). Bayesian variable selection in clustering high-dimensional data. *Journal of the American Statistical Association* **100**, 602–617. MR2160563

Wang, X. and Dunson, D. B. (2013). Parallel MCMC via Weierstrass sampler. arXiv preprint, available at arXiv:1312.4605.

Wang, X., Fangjian, G., Heller, K. A. and Dunson, D. B. (2015). Parallelizing MCMC with random partition trees. Researchgate.net, DOI:10.13140/RG.2.1.2921.4883.

Brazilian Journal of Probability and Statistics
2017, Vol. 31, No. 4, 686–691
https://doi.org/10.1214/17-BJPS365B
Main article: https://doi.org/10.1214/17-BJPS365

Comment: A brief survey of the current state of play for Bayesian computation in data science at big-data scale

David Draper[a] **and Alexander Terenin**[b]

[a]*University of California, Santa Cruz*
[b]*Imperial College London*

References

Barkalov, K., Gergel, V. and Lebedev, I. (2016). Solving global optimization problems on a GPU cluster. *AIP Conference Proceedings* **1738**.

Beam, A. L., Ghosh, S. K. and Doyle, J. (2016). Fast Hamiltonian Monte Carlo using GPU computing. *Journal of Computational and Graphical Statistics* **25**, 536–548. MR3499693

Betancourt, M. (2017). A conceptual introduction to Hamiltonian Monte Carlo. Available at arXiv:1701.02434.

Bierkens, J., Fearnhead, P. and Roberts, G. (2016). The zig-zag process and super-efficient sampling for Bayesian analysis of big data. Available at arXiv:1607.03188.

Blei, D. (2012). Probabilistic topic models. *Communications of the ACM* **55**, 77–84.

Bouchard-Côté, A., Vollmer, S. J. and Doucet, A. (2015). The bouncy particle sampler: A non-reversible rejection-free Markov chain Monte Carlo method. Available at arXiv:1510.02451.

Carvalho, C. M., Polson, N. G. and Scott, J. G. (2010). The horseshoe estimator for sparse signals. *Biometrika* **97**, 465–480. MR2650751

Krizhevsky, A., Sutskever, I. and Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems*, 1097–1105.

Lee, A., Yau, C., Giles, M. B., Doucet, A. and Holmes, C. C. (2010). On the utility of graphics cards to perform massively parallel simulation of advanced Monte Carlo methods. *Journal of Computational and Graphical Statistics* **19**, 769–789.

Lunn, D. J., Thomas, A., Best, N. and Spiegelhalter, D. (2000). WinBUGS—A Bayesian modelling framework: Concepts, structure, and extensibility. *Statistics and Computing* **10**, 325–337.

Newman, D., Asuncion, A., Smyth, P. and Welling, M. (2009). Distributed algorithms for topic models. *Journal of Machine Learning Research* **10**, 1801–1828. MR2540777

Plummer, M. (2003). JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. In *Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003)*.

Pollock, M., Fearnhead, P., Johansen, A. M. and Roberts, G. O. (2016). The scalable Langevin exact algorithm: Bayesian inference for big data. Available at arXiv:1609.03436.

Recht, B., Re, C., Wright, S. and Niu, F. (2011). Hogwild: A lock-free approach to parallelizing stochastic gradient descent. In *Advances in Neural Information Processing Systems*, 693–701.

Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural Networks* **61**, 85–117.

Srivastava, S., Li, C. and Dunson, D. B. (2015). Scalable Bayes via barycenter in Wasserstein space. Available at arXiv:1508.05880v2.

Stan Development Team (2016). Stan modeling language: Users' guide and reference manual, version 2.14.0. Available at http://mc-stan.org.

Terenin, A., Dong, S. and Draper, D. (2016). GPU-accelerated Gibbs sampling: A case study of the horseshoe probit model. Available at arXiv:1608.04329v2.

Terenin, A., Magnusson, M., Jonsson, L. and Draper, D. (2017). Pólya urn latent Dirichlet allocation: A sparse massively parallel sampler. Available at arXiv:1704.03581v1.

Terenin, A., Simpson, D. and Draper, D. (2017). Asynchronous Gibbs sampling. Available at arXiv:1509.08999v3.

Tran, D., Toulis, P. and Airoldi, E. (2016). Stochastic gradient descent methods for estimation with large data sets. Available at https://cran.r-project.org/web/packages/sgd/vignettes/sgd-jss.pdf.

Welling, M. and Teh, Y. W. (2011). Bayesian learning via stochastic gradient Langevin dynamics. In *Proceedings of the 28th International Conference on Machine Learning (ICML-11)*, 681–688. MR3157685

Brazilian Journal of Probability and Statistics
2017, Vol. 31, No. 4, 692–696
https://doi.org/10.1214/17-BJPS365A
Main article: https://doi.org/BJPS365

Comment: Consensus Monte Carlo using expectation propagation

Andrew Gelman[a] **and Aki Vehtari**[b]

[a]*Columbia University*
[b]*Aalto University*

References

Ahn, S., Korattikara, A. and Welling, M. (2012). Bayesian posterior sampling via stochastic gradient Fisher scoring. In *Proceedings of the 29th International Conference on Machine Learning*.

Gelman, A. (2016). Explanations for that shocking 2% shift. Statistical modeling, causal inference, and social science blog, 9 Nov. Available at http://andrewgelman.com/2016/11/09/explanations-shocking-2-shift/.

Gelman, A., Vehtari, A., Jylanki, P., Robert, C., Chopin, N. and Cunningham, J. P. (2014). Expectation propagation as a way of life. Available at arXiv:1412.4869.

Gershman, S., Hoffman, M. and Blei, D. (2012). Nonparametric variational inference. In *Proceedings of the 29th International Conference on Machine Learning*.

Heskes, T., Opper, M., Wiegerinck, W., Winther, O. and Zoeter, O. (2005). Approximate inference techniques with expectation constraints. *Journal of Statistical Mechanics: Theory and Experiment* P11015.

Hoffman, M., Blei, D. M., Wang, C. and Paisley, J. (2013). Stochastic variational inference. *Journal of Machine Learning Research* **14**, 1303–1347. MR3081926

Huang, Z. and Gelman, A. (2005). Sampling for Bayesian computation with large datasets. Technical report, Department of Statistics, Columbia University.

Minka, T. (2001). Expectation propagation for approximate Bayesian inference. In *Proceedings of the Seventeenth Conference on Uncertainty in Artificial Intelligence* (J. Breese and D. Koller, eds.) 362–369.

Neiswanger, W., Wang, C. and Xing, E. (2013). Asymptotically exact, embarrassingly parallel MCMC. Available at arXiv:1311.4780.

Scott, S. L. (2017). Comparing consensus Monte Carlo strategies for distributed Bayesian computation. *Brazilian Journal of Probability and Statistics*. To appear.

Scott, S. L., Blocker, A. W., Bonassi, F. V., Chipman, H. A., George, E. I. and McCulloch, R. E. (2013). Bayes and big data: The consensus Monte Carlo algorithm. In *Bayes 250*. Available at http://research.google.com/pubs/pub41849.html.

Tresp, V. (2000). A Bayesian committee machine. *Neural Computation* **12**, 2719–2741.

Wang, C. and Blei, D. M. (2013). Variational inference in nonconjugate models. *Journal of Machine Learning Research* **14**, 899–925. MR3063617

Wang, C., Chen, M. H., Schifano, E., Wu, J. and Yan, J. (2015a). Statistical methods and computing for big data. Available at arXiv:1502.07989.

Wang, W., Rothschild, D., Goel, S. and Gelman, A. (2015b). Forecasting elections with non-representative polls. *International Journal of Forecasting* **31**, 980–991.

Wang, X. and Dunson, D. B. (2013). Parallelizing MCMC via Weierstrass sampler. arXiv preprint, available at arXiv:1312.4605.

Brazilian Journal of Probability and Statistics
2017, Vol. 31, No. 4, 697–700
https://doi.org/10.1214/17-BJPS368
Main article: https://doi.org/10.1214/17-BJPS365

Rejoinder

Steven L. Scott

Google

References

Bumbuca, F., Misra, S. and Rossi, P. E. (2017). Distributed Markov chain Monte Carlo for Bayesian hierarchical models. Technical report, available at https://ssrn.com/abstract=2964646.

Huang, Z. and Gelman, A. (2005). Sampling for Bayesian computation with large datasets. Technical report, Columbia University Department of Statistics.

Scott, S. L., Blocker, A. W., Bonassi, F. V., Chipman, H. A., George, E. I. and McCulloch, R. E. (2016). Bayes and big data: The consensus Monte Carlo algorithm. *International Journal of Management Science and Engineering Management* **11**, 78–88.

Wikipedia (2017). https://en.wikipedia.org/wiki/Tensor_processing_unit.

Zhang, Y., Duchi, J. C. and Wainwright, M. J. (2012). Communication-efficient algorithms for statistical optimization. In *Decision and Control (CDC), 2012 IEEE 51st Annual Conference on*, 6792.

Brazilian Journal of Probability and Statistics
2017, Vol. 31, No. 4, 701–731
https://doi.org/10.1214/17-BJPS364

Dynamics & sparsity in latent threshold factor models: A study in multivariate EEG signal processing

Jouchi Nakajima[a] **and Mike West**[b]

[a]*Bank for International Settlements*
[b]*Duke University*

Abstract. We discuss Bayesian analysis of multivariate time series with dynamic factor models that exploit time-adaptive sparsity in model parametrizations via the latent threshold approach. One central focus is on the transfer responses of multiple interrelated series to underlying, dynamic latent factor processes. Structured priors on model hyper-parameters are key to the efficacy of dynamic latent thresholding, and MCMC-based computation enables model fitting and analysis. A detailed case study of electroencephalographic (EEG) data from experimental psychiatry highlights the use of latent threshold extensions of time-varying vector autoregressive and factor models. This study explores a class of dynamic transfer response factor models, extending prior Bayesian modeling of multiple EEG series and highlighting the practical utility of the latent thresholding concept in multivariate, non-stationary time series analysis.

References

Aguilar, O., Prado, R., Huerta, G. and West, M. (1999). Bayesian inference on latent structure in time series (with discussion). In *Bayesian Statistics, Vol. 6* (J. M. Bernardo, J. O. Berger, A. P. Dawid and A. F. M. Smith, eds.) 3–26. Oxford: Oxford University Press. MR1723490

Aguilar, O. and West, M. (2000). Bayesian dynamic factor models and portfolio allocation. *Journal of Business and Economic Statistics* **18**, 338–357.

Bernanke, B., Boivin, J. and Eliasz, P. (2005). Measuring the effects of monetary policy: A factor-augmented vector autoregressive (FAVAR) approach. *The Quarterly Journal of Economics* **120**, 387–422.

Bhattacharya, A. and Dunson, D. B. (2011). Sparse Bayesian infinite factor models. *Biometrika* **98**, 291–306. MR2806429

Carvalho, C. M., Chang, J., Lucas, J. E., Nevins, J. R., Wang, Q. and West, M. (2008). High-dimensional sparse factor modeling: Applications in gene expression genomics. *Journal of the American Statistical Association* **103**, 1438–1456. MR2655722

Carvalho, C. M., Lopes, H. F. and Aguilar, O. (2011). Dynamic stock selection strategies: A structured factor model framework (with discussion). In *Bayesian Statistics, Vol. 9* (J. M. Bernardo, M. J. Bayarri, J. O. Berger, A. P. Dawid, D. Heckerman, A. F. M. Smith and M. West, eds.) 69–90. Oxford: Oxford University Press. MR3204454

Key words and phrases. Dynamic factor models, Dynamic sparsity, EEG time series, factor-augmented vector autoregression, impulse response, multivariate time series, sparse time-varying loadings, time-series decomposition, transfer response factor models.

Del Negro, M. and Otrok, C. M. (2008). Dynamic factor models with time-varying parameters: Measuring changes in international business cycles. Staff Report 326, Federal Reserve Bank of New York. DOI:10.2139/ssrn.1136163.

Doornik, J. A. (2006). *Ox: Object Oriented Matrix Programming*. London: Timberlake Consultants Press.

Dyro, F. M. (1989). *The EEG Handbook*. Boston: Little, Brown and Co.

Huerta, G. and West, M. (1999). Priors and component structures in autoregressive time series models. *Journal of the Royal Statistical Society, Series B* **61**, 881–899. MR1722245

Kimura, T. and Nakajima, J. (2016). Identifying conventional and unconventional monetary policy shocks: A latent threshold approach. *The BE Journals in Macroeconomics* **16**, 277–300.

Kitagawa, G. and Gersch, W. (1996). *Smoothness Priors Analysis of Time Series*. *Lecture Notes in Statistics* **116**. New York: Springer. MR1441074

Koop, G. and Korobilis, D. (2010). Bayesian multivariate time series methods for empirical macroeconomics. *Foundations and Trends in Econometrics* **3**, 267–358. DOI:10.1561/0800000013.

Koop, G. M. and Potter, S. (2004). Forecasting in dynamic factor models using Bayesian model averaging. *Econometrics Journal* **7**, 550–565. MR2103792

Lopes, H. F. and Carvalho, C. M. (2007). Factor stochastic volatility with time varying loadings and Markov switching regimes. *Journal of Statistical Planning and Inference* **137**, 3082–3091. MR2364152

Lopes, H. F. and West, M. (2004). Bayesian model assessment in factor analysis. *Statistica Sinica* **14**, 41–67. MR2036762

Lucas, J. E., Carvalho, C. M., Wang, Q., Bild, A. H., Nevins, J. R. and West, M. (2006). Sparse statistical modelling in gene expression genomics. In *Bayesian Inference for Gene Expression and Proteomics* (K. A. Do, P. Mueller and M. Vannucci, eds.) 155–176. Cambridge: Cambridge University Press. MR2269095

Lucas, J. E., Carvalho, C. M. and West, M. (2009). A Bayesian analysis strategy for cross-study translation of gene expression biomarkers. *Statistical Applications in Genetics and Molecular Biology* **8**, Article no. 11. MR2476389

Nakajima, J. and West, M. (2013a). Bayesian analysis of latent threshold dynamic models. *Journal of Business & Economic Statistics* **31**, 151–164. MR3055329

Nakajima, J. and West, M. (2013b). Bayesian dynamic factor models: Latent threshold approach. *Journal of Financial Econometrics* **11**, 116–153. DOI:10.1093/jjfinec/nbs013.

Nakajima, J. and West, M. (2015). Dynamic network signal processing using latent threshold models. *Digital Signal Processing* **47**, 6–15. MR3425313

Pitt, M. and Shephard, N. (1999). Time varying covariances: A factor stochastic volatility approach (with discussion). In *Bayesian Statistics, Vol. 6* (J. M. Bernardo, J. O. Berger, A. P. Dawid and A. F. M. Smith, eds.) 547–570. Oxford: Oxford University Press. MR1724873

Prado, R. (2010a). Characterization of latent structure in brain signals. In *Statistical Methods for Modeling Human Dynamics* (S. Chow, E. Ferrer and F. Hsieh, eds.) 123–153. New York: Routledge, Taylor and Francis.

Prado, R. (2010b). Multi-state models for mental fatigue. In *The Handbook of Applied Bayesian Analysis* (A. O'Hagan and M. West, eds.) 845–874. Oxford: Oxford University Press. MR2790366

Prado, R. and Huerta, G. (2002). Time-varying autoregressions with model order uncertainty. *Journal of Time Series Analysis* **23**, 599–618. MR1925266

Prado, R. and West, M. (2010). *Time Series Modeling, Computation, and Inference*. New York: Chapman & Hall/CRC.

Prado, R., West, M. and Krystal, A. D. (2001). Multichannel electroencephalographic analyses via dynamic regression models with time-varying lag-lead structure. *Journal of the Royal Statistical Society Series C Applied Statistics* **50**, 95–109.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P. and van der Linde, A. (2002). Bayesian measures of model complexity and fit (with discussion). *Journal of the Royal Statistical Society, Series B* **64**, 583–639.

Weiner, R. D. and Krystal, A. D. (1994). The present use of electroconvulsive therapy. *Annual Review of Medicine* **45**, 273–281.

West, M. (1997). Time series decomposition. *Biometrika* **84**, 489–494.

West, M. (2003). Bayesian factor regression models in the "large p, small n" paradigm. In *Bayesian Statistics, Vol. 7* (J. M. Bernardo, M. J. Bayarri, J. O. Berger, A. P. David, D. Heckerman, A. F. M. Smith and M. West, eds.) 723–732. Oxford: Oxford University Press. MR2003537

West, M. (2013). Bayesian dynamic modelling. In *Bayesian Theory and Applications, Vol. 8* (P. Damien, P. Dellaportes, N. G. Polson and D. A. Stephens, eds.) 145–166. Oxford: Oxford University Press. MR3221162

West, M. and Harrison, P. J. (1997). *Bayesian Forecasting and Dynamic Models*, 2nd ed. New York: Springer. MR1482232

West, M., Prado, R. and Krystal, A. D. (1999). Evaluation and comparison of EEG traces: Latent structure in nonstationary time series. *Journal of the American Statistical Association* **94**, 375–387.

Yoshida, R. and West, M. (2010). Bayesian learning in sparse graphical factor models via annealed entropy. *Journal of Machine Learning Research* **11**, 1771–1798. MR2653356

Zhou, X., Nakajima, J. and West, M. (2014). Bayesian forecasting and portfolio decisions using dynamic dependent factor models. *International Journal of Forecasting* **30**, 963–980. DOI:10.1016/j.ijforecast.2014.03.017.

Brazilian Journal of Probability and Statistics
2017, Vol. 31, No. 4, 732–745
https://doi.org/10.1214/17-BJPS374

Barker's algorithm for Bayesian inference with intractable likelihoods

Flávio B. Gonçalves[a], Krzysztof Łatuszyński[b,c] and Gareth O. Roberts[b]

[a]*Departamento de Estatística, Universidade Federal de Minas Gerais*
[b]*Department of Statistics, University of Warwick*
[c]*Alan Turing Institute*

Abstract. In this expository paper, we abstract and describe a simple MCMC scheme for sampling from intractable target densities. The approach has been introduced in Gonçalves, Łatuszyński and Roberts (2017a) in the specific context of jump-diffusions, and is based on the Barker's algorithm paired with a simple Bernoulli factory type scheme, the so called *2-coin algorithm*. In many settings, it is an alternative to standard Metropolis–Hastings pseudo-marginal method for simulating from intractable target densities. Although Barker's is well known to be slightly less efficient than Metropolis–Hastings, the key advantage of our approach is that it allows to implement the "marginal Barker's" instead of the extended state space pseudo-marginal Metropolis–Hastings, owing to the special form of the accept/reject probability. We shall illustrate our methodology in the context of Bayesian inference for discretely observed Wright–Fisher family of diffusions.

References

Andrieu, C. and Roberts, G. O. (2009). The pseudo-marginal approach for efficient Monte Carlo computations. *The Annals of Statistics* **37**, 697–725. MR2502648

Andrieu, C. and Vihola, M. (2015). Convergence properties of pseudo-marginal Markov chain Monte Carlo algorithms. *The Annals of Applied Probability* **25**, 1030–1077.

Asmussen, S., Glynn, P. and Thorisson, H. (1992). Stationarity detection in the initial transient problem. *ACM Transactions on Modeling and Computer Simulation* **2**, 130–157.

Barker, A. A. (1965). Monte Carlo calculations of the radial distribution functions for a protonelectron plasma. *Australian Journal of Physics* **18**, 119–133.

Beaumont, M. A. (2003). Estimation of population growth or decline in genetically monitored populations. *Genetics* **164**, 1139–1160.

Beskos, A., Papaspiliopoulos, O. and Roberts, G. O. (2006a). Retrospective exact simulation of diffusion sample paths with applications. *Bernoulli* **12**, 1077–1098.

Beskos, A., Papaspiliopoulos, O. and Roberts, G. O. (2008). A new factorisation of diffusion measure and sample path reconstruction. *Methodology and Computing in Applied Probability* **10**, 85–104.

Beskos, A., Papaspiliopoulos, O., Roberts, G. O. and Fearnhead, P. (2006b). Exact and computationally efficient likelihood-based inference for discretely observed diffusion processes (with discussion). *Journal of the Royal Statistical Society, Series B* **68**, 333–382.

Flegal, J. M. and Herbei, R. (2012). Exact sampling for intractable probability distributions via a Bernoulli factory. *Electronic Journal of Statistics* **6**, 10–37.

Key words and phrases. Intractable likelihood, Bayesian inference, Barker's algorithm, Bernoulli factory, 2-coin algorithm, stochastic differential equations, Wright–Fisher diffusion.

Gonçalves, F. B., Łatuszyński, K. G. and Roberts, G. O. (2017a). Exact Monte Carlo likelihood-based inference for jump-diffusion processes. Available at arXiv:1707.00332.

Gonçalves, F. B., Roberts, G. O. and Łatuszyński, K. G. (2017b). Exact Bayesian inference for diffusion driven Cox processes. In preparation.

Herbei, R. and Berliner, L. M. (2014). Estimating ocean circulation: An MCMC approach with approximated likelihoods via the Bernoulli factory. *Journal of the American Statistical Association* **109**, 944–954.

Huber, M. (2015). Optimal linear Bernoulli factories for small mean problems. Available at arXiv:1507.00843.

Huber, M. (2016). Nearly optimal Bernoulli factories for linear functions. *Combinatorics, Probability & Computing* **25**, 577–591.

Jacob, P. E. and Thiery, A. H. (2015). On nonnegative unbiased estimators. *The Annals of Statistics* **43**, 769–784.

Jenkins, P. A. and Spanó, D. (2016). Exact simulation of the Wright–Fisher diffusion. *The Annals of Applied Probability* **27**, 1478–1509. MR3678477

Keane, M. and O'Brien, G. (1994). A Bernoulli factory. *ACM Transactions on Modeling and Computer Simulation* **4**, 213–219.

Kloeden, P. and Platen, E. (1995). *Numerical Solution of Stochastic Differential Equations*. New York: Springer.

Łatuszyński, K., Kosmidis, I., Papaspiliopoulos, O. and Roberts, G. (2011). Simulating events of unknown probabilities via reverse time martingales. *Random Structures & Algorithms* **38**, 441–452.

Łatuszyński, K., Palczewski, J. and Roberts, G. (2017). Exact inference for a Markov switching diffusion model with discretely observed data. In preparation.

Łatuszyński, K. and Roberts, G. O. (2013). CLTs and asymptotic variance of time-sampled Markov chains. *Methodology and Computing in Applied Probability* **15**, 237–247.

Mira, A. (2001). Ordering and improving the performance of Monte Carlo Markov chains. *Statistical Science* **16**, 340–350. MR1888449

Mira, A. and Geyer, C. (1999). Ordering Monte Carlo Markov chains. Technical report, School of Statistics, Univ. Minnesota.

Nacu, Ş. and Peres, Y. (2005). Fast simulation of new coins from old. *The Annals of Applied Probability* **15**, 93–115.

Papaspiliopoulos, O. and Roberts, G. O. (2008). Retrospective Markov chain Monte Carlo methods for Dirichlet process hierarchical models. *Biometrika* **95**, 169–186.

Peskun, P. H. (1973). Optimum Monte Carlo sampling using Markov chains. *Biometrika* **60**, 607–612.

Roberts, G. and Rosenthal, J. (1997). Geometric ergodicity and hybrid Markov chains. *Electronic Communications in Probability* **2**, 13–25.

Schraiber, J. G., Griffiths, R. C. and Evans, S. N. (2013). Analysis and rejection sampling of Wright–Fisher diffusion bridges. *Theoretical Population Biology* **89**, 64–74.

Sermaidis, G., Papaspiliopoulos, O., Roberts, G. O., Beskos, A. and Fearnhead, P. (2013). Markov chain Monte Carlo for exact inference for diffusions. *Scandinavian Journal of Statistics* **40**, 294–321. MR3066416

Von Neumann, J. (1951). Various techniques used in connection with random digits. In *Monte Carlo Method, Vol. 12*. National Bureau of Standards.

Brazilian Journal of Probability and Statistics
2017, Vol. 31, No. 4, 746–764
https://doi.org/10.1214/17-BJPS379

A Bayesian approach for a zero modified Poisson model to predict match outcomes applied to the 2012–13 La Liga season

Katiane S. Conceição, Adriano K. Suzuki, Marinho G. Andrade and Francisco Louzada

University of São Paulo

Abstract. In any sports competition, strong interest is devoted to the knowledge on the team that will be champion. The result of a match, the chance of a team either qualifying for a specific tournament, or relegating, the best attack and defense are all topics of interest. This paper presents a Bayesian methodology for modeling the number of goals scored by a team based on Zero-Modified Poisson distribution. An important advantage of this distribution is the flexibility in modeling count data without previous knowledge of the sampling characteristic with respect to the frequency of zeros (inflated, standard, deflation). These characteristics are present in the data sets referring to the number of goals scored by different teams. Inference procedures and computational simulation studies are also discussed. The proposed methodology was applied to the 2012–13 La Liga and the results were compared with those of the Poisson model using the De Finetti measure an percentage of correct predictions.

References

Brillinger, D. R. (2008). Modeling game outcomes of the Brazilian 2006 series a championship as ordinal-valued. *Brazilian Journal of Probability and Statistics* **22**, 89–104. MR2575392

Chib, S. and Greenberg, E. (1995). Understanding the Metropolis–Hastings algorithm. *American Statistician* **49**, 327–335.

Conceição, K. S., Andrade, M. G. and Louzada, F. (2013). Zero-modified Poisson model: Bayesian approach, influence diagnostics and an application to a Brazilian leptospirosis notification data. *Biometrical Journal* **55**, 661–678.

Conceição, K. S., Andrade, M. G. and Louzada, F. (2014). On the zero-modified Poisson model: Bayesian analysis and posterior divergence measure. *Computational Statistics* **29**, 959–980. MR3266043

Cowless, M. K. and Carlin, B. P. (1996). Markov chain Monte Carlo convergence diagnostics: A comparative review. *Journal of the American Statistical Association* **91**, 883–904.

De Finetti, B. (1972). *Probability, Induction and Statistics*. London: John Wiley. MR0440638

Dietz, E. and Böhning, D. (2000). On estimation of the Poisson parameter in zero-modified Poisson models. *Computational Statistics & Data Analysis* **34**, 441–459.

Dyte, D. and Clarke, S. R. (2000). A ratings based Poisson model for world cup soccer simulation. *Journal of the Operational Research Society* **51**, 993–998.

Key words and phrases. Applied probability, OR in sports, ZMP model, Bayesian approach, simulation.

Gelman, A. and Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science* **7**, 457–511.

Karlis, D. and Ntzoufras, I. (2009). Bayesian modeling of football outcomes: Using the Skellam's distribution for the goal difference. *IMA Journal of Management Mathematics* **20**, 133–145.

Keller, J. B. (1994). A characterization of the Poisson distribution and the probability of winning a game. *American Statistician* **48**, 294–298.

Lee, A. (1997). Modeling scores in the Premier League: Is Manchester United really the best? *Chance* **10**, 15–19.

Plummer, M., Best, N., Cowles, K. and Vines, K. (2006). Output analysis and diagnostics for MCMC. http://cran.r-project.org/web/packages/coda/index.html.

Saraiva, E. F., Suzuki, A. K., Filho, C. A. O. and Louzada, F. (2016). Predicting football scores via Poisson regression model: Applications to the National Football League. *Communications for Statistical Applications and Methods* **23**, 297–319.

Suzuki, A. K., Salasar, L. E. B., Louzada-Neto, F. and Leite, J. G. (2009). A Bayesian approach for predicting match outcomes: The 2006 (association) football world cup. *Journal of the Operational Research Society* **61**, 1530–1539.

Volf, P. (2009). A random point process model for the score in sport matches. *IMA Journal of Management Mathematics* **20**, 121–131.

Brazilian Journal of Probability and Statistics
2017, Vol. 31, No. 4, 765–800
https://doi.org/10.1214/17-BJPS375

Studying the effective brain connectivity using multiregression dynamic models

Lilia Costa[a], Thomas Nichols[b,c,d] and Jim Q. Smith[e]

[a]*Universidade Federal da Bahia*
[b]*Oxford Big Data Institute*
[c]*Li Ka Shing Centre for Health Information*
[d]*Discovery, Nuffield Department of Population Health, University of Oxford*
[e]*The University of Warwick*

Abstract. The *Multiregression Dynamic Model* (MDM) is a multivariate graphical model for a multidimensional time series that allows the estimation of time-varying effective connectivity. An MDM is a state space model where connection weights reflect the contemporaneous interactions between brain regions. Because the marginal likelihood has a closed form, model selection across a large number of potential connectivity networks is easy to perform. With application of the Integer Programming Algorithm, we can quickly find optimal models that satisfy acyclic graph constraints and, due to a factorisation of the marginal likelihood, the search over all possible directed (acyclic or cyclic) graphical structures is even faster. These methods are illustrated using recent resting-state and steady-state task fMRI data.

References

Achterberg, T. (2007). Constraint integer programming. PhD thesis, TU Berlin.

Baba, K., Shibata, R. and Sibuya, M. (2004). Partial correlation and conditional correlation as measures of conditional independence. *Australian and New Zealand Journal of Statistics* **46**, 4, 657–664.

Bartlett, M. and Cussens, J. (2013). Advances in Bayesian network learning using integer programming. arXiv preprint. Available at arXiv:1309.6825.

Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B* **57**, 1, 289–300.

Chang, C., Thomason, M. E. and Glover, G. H. (2008). Mapping and correction of vascular hemodynamic latency in the BOLD signal. *NeuroImage* **43**, 90–102.

Costa, L., Smith, J., Nichols, T., Cussens, J., Duff, E. P. and Makin, T. R. (2015). Searching multiregression dynamic models of resting-state fMRI networks using integer programming. *Bayesian Analysis* **10**, 441–478.

Cowell, R. G. (2013). A simple greedy algorithm for reconstructing pedigrees. *Theoretical Population Biology* **83**, 55–63.

Cussens, J. (2010). SMaximum likelihood pedigree reconstruction using integer programming. *WCB@ ICLP* 8–19.

Cussens, J. (2012). Bayesian network learning with cutting planes. arXiv preprint. Available at arXiv:1202.3713.

Key words and phrases. Multiregression dynamic model, Bayesian network, effective connectivity, functional magnetic resonance imaging, integer programming algorithm.

David, O., Guillemain, I., Saillet, S., Reyt, S., Deransart, C., Segebarth, C. and Depaulis, A. (2008). Identifying neural drivers with functional MRI: An electrophysiological validation. *PLoS Biology* **6**, 2683–2697.

Duff, E., Tamar, M., Smith, S. M. and Woolrich, M. W. (2017). Disambiguating brain functional connectivity. bioRxiv. http://biorxiv.org/content/early/2017/01/25/103002.

Friston, K. J. (2011). Functional and Effective Connectivity: a review. *Brain Connectivity* **1**, 1, 13–36.

Friston, K. J., Harrison, L. and Penny, W. (2003). Dynamic causal modelling. *NeuroImage* **19**, 1273–1302.

Goldenberg, A., Zheng, A. X., Fienberg, S. E., Airoldi, E. M. and others (2010). A survey of statistical network models. *Foundations and Trends® in Machine Learning* **2**, 2, 129–233.

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* **37**, 424–438.

Griffanti, L., Salimi-Khorshidi, G., Beckmann, C. F., Auerbach, E. J., Douaud, G., Sexton, C. E., Zsoldos, E., Ebmeier, K. P., Filippini, N., Mackay, C. E. and Moeller, S. (2014). ICA-based artefact removal and accelerated fMRI acquisition for improved resting state network imaging. *NeuroImage* **95**, 232–247.

Havlicek, M., Jan, J., Brazdil, M. and Calhoun, V. D. (2010). Dynamic Granger causality based on Kalman filter for evaluation of functional network connectivity in fMRI data. *NeuroImage* **53**, 65–77.

Heckerman, D. (1998). A tutorial on learning with Bayesian networks. *Nato Asi Series D Behavioural And Social Sciences* **89**, 301–354.

Jeffreys, H. (1961). *Theory of Probability*, 3rd ed. London: Oxford University Press.

Jenkinson, M., Beckmann, C. F., Behrens, T. E., Woolrich, M. W. and Smith, S. M. (2012). FSL. *NeuroImage* **62**, 782–790.

Koster, J. T. (1996). Markov properties of nonrecursive causal models. *The Annals of Statistics* 2148–2177.

Marrelec, G., Krainik, A., Duffau, H., Pélégrini-Issac, M., Lehéricy, S., Doyon, J. and Benali, H. (2006). Partial correlation for functional brain interactivity investigation in functional MRI. *NeuroImage* **62**, 228–237.

Pearl, J. (2000). *Causality: Models, Reasoning, and Inference*. Cambridge: Cambridge University Press.

Penny, W., Ghahramani, Z. and Friston, K. (2005). Bilinear dynamical systems. *Philosophical Transactions of the Royal Society of London Series B, Biological Sciences* **360**, 983–993.

Petris, G., Petrone, S. and Campagnoli, P. (2009). *Dynamic Linear Models with R*. New York: Springer.

Poldrack, R. A., Mumford, J. A. and Nichols, T. E. (2011). *Handbook of fMRI Data Analysis*. Cambridge University Press.

Queen, C. M. and Albers, C. J. (2008). Forecast covariances in the linear multiregression dynamic model. *J. Forecast.* **27**, 175–191.

Queen, C. M. and Albers, C. J. (2009). Intervention and causality: Forecasting traffic flows using a dynamic Bayesian network. *Journal of the American Statistical Association* **104**, 669–681.

Queen, C. M. and Smith, J. Q. (1993). Multiregression dynamic models. *Journal of the Royal Statistical Society, Series B* **55**, 849–870.

R Core Team (2016). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Raichle, M. E. (2010). Two views of brain function. *Trends in Cognitive Sciences* **14**, 180–190.

Ramsey, J. D., Hanson, S. J., Hanson, C., Halchenko, Y. O., Poldrack, R. A. and Glymour, C. (2010). Six problems for causal inference from fMRI. *NeuroImage* **49**, 1545–1558.

Ridgway, G., Leite, A. B., Penny, W. and Friston, K. (2013). Stochastic DCM of the DMN using resting-state fMRI: test-retest reliability. figshare. https://doi.org/10.6084/m9.figshare.866771.v1.

Ryali, S., Supekar, K., Chen, T. and Menon, V. (2011). Multivariate dynamical systems models for estimating causal interactions in fMRI. *NeuroImage* **54**, 807–823.

Salimi-Khorshidi, G., Douaud, G., Beckmann, C. F., Glasser, M. F., Griffanti, L. and Smith, S. M. (2014). Automatic denoising of functional MRI data: Combining independent component analysis and hierarchical fusion of classifiers. *NeuroImage* **90**, 449–468.

Schwab, S., Harbord, R., Costa, L. and Nichols, T. E. (2017). multdyn: A package for Multiregression Dynamic Models (MDM). Available at https://github.com/schw4b/multdyn.

Shehzad, Z., Kelly, A. C., Reiss, P. T., Gee, D. G., Gotimer, K., Uddin, L. Q., Lee, S. H., Margulies, D. S., Roy, A. K., Biswal, B. B. and Petkova, E. (2009). The resting brain: Unconstrained yet reliable. *Cerebral Cortex* **19**, 2209–2229.

Sloane, N. J. A. and Plouffe, S. (1995). *The Encyclopedia of Integer Sequences*. Academic Press.

Smith, J. F., Pillai, A., Chen, K. and Horwitz, B. (2010). Identification and validation of effective connectivity networks in functional magnetic resonance imaging using switching linear dynamic systems. *NeuroImage* **52**, 1027–1040.

Smith, J. F., Pillai, A., Chen, K. and Horwitz, B. (2011). Effective connectivity modeling for fMRI: Six issues and possible solutions using linear dynamic systems. *Frontiers in Systems Neuroscience* **5**, 104.

Smith, J. Q. and Croft, J. (2003). Bayesian networks for discrete multivariate data: An algebraic approach to inference. *Journal of Multivariate Analysis* **84**, 387–402.

Smith, S. M., Bandettini, P. A., Miller, K. L., Behrens, T. E. J., Friston, K. J., David, O., Liue, T., Woolricha, M. W. and Nichols, T. E. (2012). The danger of systematic bias in group-level FMRI-lag-based causality estimation. *NeuroImage* **59**, 1228–1229.

Smith, S. M., Fox, P. T., Miller, K. L., Glahn, D. C., Fox, P. M., Mackay, C. E., Filippini, N., Watkins, K. E., Toro, R., Laird, A. R. and Beckmann, C. F. (2009). Correspondence of the brain's functional architecture during activation and rest. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 13040–13045.

Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., Bannister, P. R., De Luca, M., Drobnjak, I., Flitney, D. E., Niazy, R., Saunders, J., Vickers, J., Zhang, Y., De Stefano, N., Brady, J. M. and Matthews, P. M. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage* **23**, 208–219.

Spirtes, P. (1995). Directed cyclic graphical representations of feedback models. In *Uncertainty in Artificial Intelligence* **11** (P. Besnard and S. Hanks, eds.) 491–498. Morgan Kaufmann.

Spirtes, P., Glymour, C. N. and Scheines, R. (2000). *Causation, Prediction, and Search*, 2nd ed. Cambridge, MA: MIT Press.

Sporns, O. (2010). *Networks of the Brain*, 1st ed. MIT Press.

Stephan, K. E., Kasper, L., Harrison, L. M., Daunizeau, J., den Ouden, H. E., Breakspear, M. and Friston, K. J. (2008). Nonlinear dynamic causal models for fMRI. *NeuroImage* **42**, 649–662.

Valdés-Sosa, P. A., Roebroeck, A., Daunizeau, J. and Friston, K. (2011). Effective connectivity: Influence, causality and biophysical modeling. *NeuroImage* **58**, 339–361.

West, M. and Harrison, P. J. (1997). *Bayesian Forecasting and Dynamic Models*, 2nd ed. New York: Springer.

Williams, H. P. (2009). *Logic and Integer Programming*. Springer. ISBN 978-0-387-92279-9.

Woolrich, M. W., Jbabdi, S., Patenaude, B., Chappell, M., Makni, S., Behrens, T., Beckmann, C., Jenkinson, M. and Smith, S. M. (2009). Bayesian analysis of neuroimaging data in FSL. *NeuroImage* **45**, S173–S186.

Brazilian Journal of Probability and Statistics
2017, Vol. 31, No. 4, 801–820
https://doi.org/10.1214/17-BJPS378

A Bayesian approach to extended models for exceedance

Fernando Ferraz do Nascimento[a] **and Marcelo Bourguignon Pereira**[b]

[a]*Universidade Federal do Piauí, Brazil*
[b]*Universidade Federal do Rio Grande do Norte, Brazil*

Abstract. In extreme value theory, the generalized Pareto distribution (GPD) is a family of continuous distribution used to model the tail of the distribution to values higher than a threshold u. Several works have used this method to approximate the tail of distribution. In this paper, we propose two extensions of GPD, including an additional shape parameter, to provide a more flexible distribution for exceedance. Some properties of these approximations are presented. Inference for these extensions were performed under the Bayesian paradigm, and the results showed fit improvement when compared with the standard GPD in applications to environmental and financial data.

References

Castillo, E. and Hadi, A. S. (1997). Fitting the generalized Pareto distribution to data. *J. Amer. Statist. Assoc.* **92**, 1609–1620. MR1615270

Coles, S. (2001). *An Introduction to Statistical Modelling of Extreme Values*. *Springer Series in Statistics*. London: Springer-Verlag London, Ltd. MR1932132

Cordeiro, G. M., Alizadeh, M. and Marinho, R. D. M. (2016). The type I half-logistic family of distributions. *Journal of Statistical Computation and Simulation* **86**, 707–728.

Davison, A. C. and Smith, R. L. (1990). Models for exceedances over high thresholds (with discussion). *Journal of the Royal Statistical Society, Series B* **52**, 393–442.

Do Nascimento, F. F., Gamerman, D. and Lopes, H. F. (2012). A semiparametric Bayesian approach to extreme value estimation. *Stat. Comput.* **22**, 661–675. MR2865043

Embrechts, P., Kluppelberg, C. and Mikosch (1997). *Modelling Extremal Events for Insurance and Finance*. New York: Springer.

Eugene, N., Lee, C. and Famoye, F. (2002). Beta-normal distribution and its applications. *Communications in Statistics Theory and Methods* **31**, 497–512.

Eugenia Castellanos, M. and Cabras, S. (2007). A default Bayesian procedure for the generalized Pareto distribution. *J. Statisti. Plann. Inference* **137**, 373–483. MR2298951

Ferreira, A. and de Haan, L. (2006). *Extreme Value Theory. And Introduction*. New York: Springer.

Gamerman, D. and Lopes, H. F. (2006). *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*, 2nd ed. Baton Rouge: Chapman and Hall/CRC.

Gupta, R. D. and Kundu, D. (2001). Exponentiated exponential family: An alternative to gamma and Weibull distributions. *Biometrical Journal* **43**, 117–130.

Jenkinson, A. F. (1955). The frequency distribution of the annual maximum (or minimum) values of meteorological events. *Quarterly Journal of the Royal Meteorology Society* **81**, 158–171.

Marshall, A. and Olkin, I. (1997). A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families. *Biometrika* **84**, 641–652.

Key words and phrases. Extreme value theory, exceedance analysis, generalized Pareto distribution, generalized classes of distributions, MCMC.

McDonald, J. B. and Xu, Y. J. (1993). A generalization of the beta distribution with applications. *Journal of Econometrics* **66**, 133–152.

Mudholkar, G. S. and Hutson, A. D. (1996). The exponentiated Weibull family: Some properties and a flood data application. *Communications in Statistics Theory and Methods* **25**, 3059–3083.

Nascimento, F. F., Bourguignon, M. and Leão, J. (2016). Extended generalized extreme value distribution with applications in environmental data. *Hacettepe Journal of Mathematics and Statistics* **45**, 1847–1864.

Nascimento, F. F., Gamerman, D. and Lopes, H. F. (2016). Time-varying extreme pattern with dynamic models. *Test* **25**, 131–149.

Ortega, E. M. M., Cordeiro, G. M. and Kattan, M. W. (2012). The negative binomial beta Weibull regression model to predict the cure of prostate cancer. *Journal of Applied Statistics* **39**, 1191–1210.

Papastathopoulos, I. and Tawn, J. A. (2013). Extended generalized Pareto models for tail estimation. *Journal of Statistical Planning and Inference* **143**, 131–143.

Pickands, J. (1975). Statistical inference using extreme order statistics. *The Annals of Statistics* **3**, 119–131.

Pickands, J. (1986). The continuous and differentiable domains of attraction in extreme value theory. *Annals of Probability* **14**, 996–1004.

Ristić, M. M. and Balakrishnan, N. (2012). The gamma exponentiated exponential distribution. *Journal of Statistical Computation and Simulation* **8**, 1191–1206.

Roberts, G. O. and Rosenthal, J. S. (2009). Examples of adaptive MCMC. *Journal of Computational and Graphical Statistics* **18**, 349–367.

Santos, T. R., Gamerman, D. and Franco, G. C. (2017). Reliability analysis via non-Gaussian state-space models. *IEEE Transactions on Reliability* **66**, 309–318.

Smith, R. L. (1985). Maximum likelihood estimation in a class of nonregular cases. *Biometrika* **72**, 67–90.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P. and Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society, Series B, Methodological* **64**, 583–639.

von Mises, R. (1954). La distribution de la plus grande de nvaleurs. *American Mathematical Society* **2**, 271–294.

Zografos, K. and Balakrishnan, N. (2009). On families of beta and generalized gamma-generated distributions and associated inference. *Statistical Methodology* **6**, 344–362.

AIMS AND SCOPE

The Brazilian Journal of Probability and Statistics aims to publish high quality research papers in applied probability, applied statistics, computational statistics, mathematical statistics, probability theory and stochastic processes.

More specifically, the following types of contributions will be considered:

(i) Original articles dealing with methodological developments, comparison of competing techniques or their computational aspects;
(ii) Original articles developing theoretical results;
(iii) Articles that contain novel applications of existing methodologies to practical problems. For these papers the focus is in the importance and originality of the applied problem, as well as, applications of the best available methodologies to solve it.
(iv) Survey articles containing a thorough coverage of topics of broad interest to probability and statistics. The journal will occasionally publish book reviews, invited papers and essays on the teaching of statistics.

GENERAL INFORMATION

Journal Homepage: http://www.imstat.org/bjps

Submissions: Manuscripts for *Brazilian Journal of Probability and Statistics* should be submitted online. Authors may access the Electronic Journals Management System (EJMS) at http://www.e-publications.org/ims/submission.

Preparation of Manuscripts: http://www.imstat.org/bjps/mansub.html

Permissions Policy. Authorization to photocopy items for internal or personal use is granted by the Institute of Mathematical Statistics. For multiple copies or reprint permission, contact The Copyright Clearance Center, 222 Rosewood Drive, Danvers, Massachusetts 01923. Telephone (978) 750-8400. http://www.copyright.com. If the permission is not found at the Copyright Clearance Center, please contact the IMS Business Office: ims@imstat.org.

Correspondence. Mail concerning membership, subscriptions, nonreceipt claims, copyright permissions, advertising or back issues should be sent to the IMS Dues and Subscription Office, 9650 Rockville Pike, Suite L 2310, Bethesda, Maryland 20814-3998. Mail concerning submissions or editorial content should be sent to the Editor at nancy@ime.unicamp.br. Mail concerning the production of this journal should be sent to: Geri Mattson at bjps@mattsonpublishing.com.

Individual and Organizational Memberships:
http://www.imstat.org/membership

Individual and General Subscriptions:
http://www.imstat.org/publications/subscriptions.htm

Electronic Access to IMS Journals:
http://www.imstat.org/publications/eaccess.htm

The *Brazilian Journal of Probability and Statistics* is an IMS supported journal:
http://imstat.org/publications/supported.html