Contents

Contributions to the Special Volume of the XIV EBEB ("Encontro Brasileiro de Estatística Bayesiana")

Articles

BJPS is an IMS supported journal   ISSN 0103-0752   Brazilian Statistical Association

Brazilian Journal of Probability and Statistics

Volume 33 • Number 4 • November 2019

Brazilian Journal of Probability and Statistics

Volume 33 • Number 4 • November 2019

www.imstat.org/bjps

ISSN 0103-0752 (Print) ISSN 2317-6199 (Online), Volume 33, Number 4, November 2019. 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
2019, Vol. 33, No. 4, 685
https://doi.org/10.1214/19-BJPS334PRE

Preface

Brazilian Journal of Probability and Statistics
2019, Vol. 33, No. 4, 686–705
https://doi.org/10.1214/19-BJPS444

Spatiotemporal point processes: regression, model specifications and future directions

Dani Gamerman

UFRJ

Abstract. Point processes are one of the most commonly encountered observation processes in Spatial Statistics. Model-based inference for them depends on the likelihood function. In the most standard setting of Poisson processes, the likelihood depends on the intensity function, and can not be computed analytically. A number of approximating techniques have been proposed to handle this difficulty. In this paper, we review recent work on exact solutions that solve this problem without resorting to approximations. The presentation concentrates more heavily on discrete time but also considers continuous time. The solutions are based on model specifications that impose smoothness constraints on the intensity function. We also review approaches to include a regression component and different ways to accommodate it while accounting for additional heterogeneity. Applications are provided to illustrate the results. Finally, we discuss possible extensions to account for discontinuities and/or jumps in the intensity function.

References

Adams, R. P., Murray, I. and Mackay, D. J. C. (2009). Tractable nonparametric Bayesian inference in Poisson processes with Gaussian process intensities. In *Proc. 26th Int. Conf. Machine Learning* (L. Bottou and M. Littman, eds.) Montreal: Omnipress.

Baddeley, A. and Turner, R. (2005). Spatstat: An R package for analyzing spatial point patterns. *Journal of Statistical Software* **12**, 1–42.

Benes, V., Boldak, K., Møller, J. and Waagepetersen, R. (2005). A case study on point process modelling in disease mapping. *Image Analysis and Stereology* **24**, 159–168.

Brown, P. E. (2015). Model-based Geostatistics the easy way. *Journal of Statistical Software* **63**.

Cox, D. R. (1955). Some statistical methods connected with series of events. *Journal of the Royal Statistical Society, Series B* **17**, 129–164. MR0092301

Dias, R., Ferreira, C. S. and Garcia, N. L. (2008). Penalized maximum likelihood estimation for a function of the intensity of a Poisson point process. *Statistical Inference for Stochastic Processes* **11**, 11–34. MR2357551

Diggle, P. J. and Brix, A. (2001). Spatio-temporal prediction for log-Gaussian Cox processes. *Journal of the Royal Statistical Society, Series B* **63**, 823–841. MR1872069

Diggle, P. J., Guan, Y., Hart, A. C., Paize, F. and Stanton, M. (2010). Estimating individual-level risk in spatial epidemiology using spatially aggregated information on the population at risk. *Journal of the American Statistical Association* **105**, 1394–1402. MR2796558

Frühwirth-Schnatter, S. and Wagner, H. (2006). Auxiliary mixture sampling for parameter-driven models of time series of counts with applications to state space modelling. *Biometrika* **93**, 827–841. MR2285074

Key words and phrases. Data augmentation, discretization, dynamic, Gaussian processes, partition models, spatial interpolation.

Gamerman, D. (1998). Markov chain Monte Carlo for dynamic generalised linear models. *Biometrika* **85**, 215–227. MR1627273

Gamerman, D. (2010). Dynamic spatial models including spatial time series. In *Handbook of Spatial Statistics* (E. A. E. Gelfand, P. Diggle, P. Guttorp and M. Fuentes, eds.) 437–448. Londres: CRC/Chapman & Hall. MR2730959

Gamerman, D. and Lopes, H. F. (2006). Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference. 2a edição. Londres: Chapman and Hall. MR2260716

Gamerman, D., Moreira, A. R. B. and Rue, H. (2003). Space-varying regression models: Specifications and simulation. *Computational Statistics & Data Analysis* **42**, 513–533. MR2005406

Gamerman, D., Santos, T. R. and Franco, G. C. (2013). A non-Gaussian family of state-space models with exact marginal likelihood. *Journal of Time Series Analysis* **34**, 625–645. MR3127211

Gelfand, A. E., Banerjee, S. and Gamerman, D. (2005). Spatial process modelling for univariate and multivariate dynamic spatial data. *EnvironMetrics* **16**, 465–479. MR2147537

Gonçalves, F. B. and Gamerman, D. (2018). Exact Bayesian inference in spatio-temporal Cox processes driven by multivariate Gaussian processes. *Journal of the Royal Statistical Society, Series B* **80**, 157–175. MR3744716

Gramacy, R. and Lee, H. K. H. (2008). Bayesian treed Gaussian process models with an application to computer modeling. *Journal of the American Statistical Association* **103**, 1119–1130. MR2528830

Kim, H.-M., Mallick, B. K. and Holmes, C. C. (2005). Analyzing nonstationary spatial data using piecewise Gaussian processes. *Journal of the American Statistical Association* **100**, 653–668. MR2160567

Lewis, P. A. W. and Shedler, G. S. (1979). Simulation of a nonhomogeneous Poisson process by thinning. *Naval Research Logistics Quarterly* **26**, 403–413. MR0546120

Liang, S., Carlin, B. and Gelfand, A. (2008). Analysis of Minnesota colon and rectum cancer point patterns with spatial and nonspatial covariate information. *Annals of Applied Statistics* **3**, 943–962. MR2750381

Liang, W. W. J., Colvin, J. B., Sansó, B. and Lee, H. K. H. (2014). Modeling and anomalous cluster detection for point processes using process convolutions. *Journal of Computational and Graphical Statistics* **23**, 129–150. MR3173764

Moller, J., Syversveen, A. R. and Waagepetersen, R. P. (1998). Log Gaussian Cox processes. *Scandinavian Journal of Statistics* **25**, 451–482. MR1650019

Pinto, J. A. Jr. (2014). Point processes with spatio-temporal heterogeneity. Unpublished Ph.D. thesis, Universidade Federal do Rio de Janeiro. Brazil. (in Portuguese).

Pinto, J. A. Jr., Gamerman, D., Paez, M. S. and Alves, R. H. F. (2015). Point pattern analysis with spatially varying covariate effects, applied to the study of cerebrovascular deaths. *Statistics in Medicine* **34**, 1214–1226. MR3322745

Reis, E. A., Gamerman, D., Paez, M. S. and Martins, T. G. (2013). Bayesian dynamic models for space–time point processes. *Computational Statistics & Data Analysis* **60**, 146–156. MR3007026

Shirota, S. and Banerjee, S. (2018). Scalable inference for space-time Gaussian Cox processes. preprint, ArXiV.

Taylor, B., Davies, T. M., Rowlingson, B. S. and Diggle, P. J. (2013). lgcp: An R Package for Inference with Spatial and Spatio-Temporal Log-Gaussian Cox Processes. *Journal of Statistical Software* **52**. MR3161587

Vivar, J. C. and Ferreira, M. A. R. (2009). Spatio-temporal models for Gaussian areal data. *Journal of Computational and Graphical Statistics* **18**, 658–674. MR2751645

Wikle, C. K. and Cressie, N. (1999). A dimension-reduced approach to space-time Kalman filtering. *Biometrika* **86**, 815–829. MR1741979

Brazilian Journal of Probability and Statistics
2019, Vol. 33, No. 4, 706–733
https://doi.org/10.1214/19-BJPS446

Keeping the balance—Bridge sampling for marginal likelihood estimation in finite mixture, mixture of experts and Markov mixture models

Sylvia Frühwirth-Schnatter

Vienna University of Economics and Business (WU), Austria

Abstract. Finite mixture models and their extensions to Markov mixture and mixture of experts models are very popular in analysing data of various kind. A challenge for these models is choosing the number of components based on marginal likelihoods. The present paper suggests two innovative, generic bridge sampling estimators of the marginal likelihood that are based on constructing balanced importance densities from the conditional densities arising during Gibbs sampling. The full permutation bridge sampling estimator is derived from considering all possible permutations of the mixture labels for a subset of these densities. For the double random permutation bridge sampling estimator, two levels of random permutations are applied, first to permute the labels of the MCMC draws and second to randomly permute the labels of the conditional densities arising during Gibbs sampling. Various applications show very good performance of these estimators in comparison to importance and to reciprocal importance sampling estimators derived from the same importance densities.

References

Aßmann, C. and Boysen-Hogrefe, J. (2011). A Bayesian approach to model-based clustering for binary panel probit models. *Computational Statistics & Data Analysis* **55**, 261–279. MR2736553

Berger, J. O. and Jefferys, W. H. (1992). Sharpening Ockham's razor on a Bayesian strop. *American Statistician* **80**, 64–72.

Berkhof, J., van Mechelen, I. and Gelman, A. (2003). A Bayesian approach to the selection and testing of mixture models. *Statistica Sinica* **13**, 423–442. MR1977735

Celeux, G., Frühwirth-Schnatter, S. and Robert, C. P. (2019). Model selection for mixture models—perspectives and strategies. In *Handbook of Mixture Analysis* (S. Frühwirth-Schnatter, G. Celeux and C. P. Robert, eds.) 117–154. Boca Raton, FL: CRC Press. MR3889692

Chib, S. (1995). Marginal likelihood from the Gibbs output. *Journal of the American Statistical Association* **90**, 1313–1321. MR1379473

DiCiccio, T. J., Kass, R. E., Raftery, A. and Wasserman, L. (1997). Computing Bayes factors by combining simulation and asymptotic approximations. *Journal of the American Statistical Association* **92**, 903–915. MR1482122

Diebolt, J. and Robert, C. P. (1994). Estimation of finite mixture distributions through Bayesian sampling. *Journal of the Royal Statistical Society, Series B* **56**, 363–375. MR1281940

Key words and phrases. Markov chain Monte Carlo, model-based clustering, Gaussian mixtures, hierarchical priors, permutation sampling, importance sampling.

Diggle, P. J., Heagerty, P., Liang, K.-Y. and Zeger, S. L. (2002). *Analysis of Longitudinal Data*, 2nd ed. Oxford: Oxford University Press. MR2049007

Escobar, M. D. and West, M. (1998). Computing nonparametric hierarchical models. In *Practical Nonparametric and Semiparametric Bayesian Statistics* (D. Dey, P. Müller and D. Sinha, eds.), *Lecture Notes in Statistics* **133**, 1–22. Berlin: Springer. MR1630073

Frühwirth-Schnatter, S. (1995). Bayesian model discrimination and Bayes factors for linear Gaussian state space models. *Journal of the Royal Statistical Society, Series B* **57**, 237–246. MR1325388

Frühwirth-Schnatter, S. (2001). Markov chain Monte Carlo estimation of classical and dynamic switching and mixture models. *Journal of the American Statistical Association* **96**, 194–209. MR1952732

Frühwirth-Schnatter, S. (2004). Estimating marginal likelihoods for mixture and Markov switching models using bridge sampling techniques. *Econometrics Journal* **7**, 143–167. MR2076630

Frühwirth-Schnatter, S. (2006). *Finite Mixture and Markov Switching Models*. New York: Springer. MR2265601

Frühwirth-Schnatter, S. (2011). Panel data analysis—A survey on model-based clustering of time series. *Advances in Data Analysis and Classification* **5**, 251–280. MR2860101

Frühwirth-Schnatter, S. (2019). Applied Bayesian Mixture Modelling. Implementations in MATLAB using the package bayesf (Version 4.0). Available at https://www.wu.ac.at/statmath/faculty-staff/faculty/sfruehwirthschnatter/.

Frühwirth-Schnatter, S., Celeux, G. and Robert, C. P., eds. (2019). *Handbook of Mixture Analysis*. Boca Raton, FL: CRC Press. MR3889980

Frühwirth-Schnatter, S. and Frühwirth, R. (2010). Data augmentation and MCMC for binary and multinomial logit models. In *Statistical Modelling and Regression Structures—Festschrift in Honour of Ludwig Fahrmeir* (T. Kneib and G. Tutz, eds.) 111–132. Heidelberg: Physica-Verlag. MR2664631

Frühwirth-Schnatter, S., Frühwirth, R., Held, L. and Rue, H. (2009). Improved auxiliary mixture sampling for hierarchical models of non-Gaussian data. *Statistics and Computing* **19**, 479–492. MR2565319

Frühwirth-Schnatter, S. and Kaufmann, S. (2008). Model-based clustering of multiple time series. *Journal of Business & Economic Statistics* **26**, 78–89. MR2422063

Frühwirth-Schnatter, S. and Malsiner-Walli, G. (2019). From here to infinity: Sparse finite versus Dirichlet process mixtures in model-based clustering. *Advances in Data Analysis & Classification* **13**, 33–64. MR3935190

Frühwirth-Schnatter, S., Pamminger, C., Weber, A. and Winter-Ebmer, R. (2012). Labor market entry and earnings dynamics: Bayesian inference using mixtures-of-experts Markov chain clustering. *Journal of Applied Econometrics* **27**, 1116–1137. MR3041877

Frühwirth-Schnatter, S., Pittner, S., Weber, A. and Winter-Ebmer, R. (2018). Analysing plant closure effects using time-varying mixture-of-experts Markov chain clustering. *Annals of Applied Statistics* **12**, 1786–1830. MR3852698

Frühwirth-Schnatter, S. and Pyne, S. (2010). Bayesian inference for finite mixtures of univariate and multivariate skew normal and skew-t distributions. *Biostatistics* **11**, 317–336.

Gelfand, A. E. and Dey, D. K. (1994). Bayesian model choice: Asymptotics and exact calculations. *Journal of the Royal Statistical Society, Series B* **56**, 501–514. MR1278223

Gelfand, A. E. and Smith, A. F. M. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association* **85**, 398–409. MR1141740

Geweke, J. (1989). Bayesian inference in econometric models using Monte Carlo integration. *Econometrica* **57**, 1317–1339. MR1035115

Gormley, I. C. and Frühwirth-Schnatter, S. (2019). Mixture of experts models. In *Handbook of Mixture Analysis* (S. Frühwirth-Schnatter, G. Celeux and C. P. Robert, eds.) 271–307. Boca Raton, FL: CRC Press. MR3889697

Lee, J. E. and Robert, C. P. (2016). Importance sampling schemes for evidence approximation in mixture models. *Bayesian Analysis* **11**, 573–597. MR3472003

Lee, K., Marin, J.-M., Mengersen, K. and Robert, C. (2009). Bayesian inference on mixtures of distributions. In *Perspectives in Mathematical Sciences I: Probability and Statistics* (N. N. Sastry, M. Delampady and B. Rajeev, eds.) 165–202. Singapore: World Scientific. MR2581744

Leroux, B. G. and Puterman, M. L. (1992). Maximum-penalized-likelihood estimation for independent and Markov-dependent mixture models. *Biometrics* **48**, 545–558.

Malsiner Walli, G., Frühwirth-Schnatter, S. and Grün, B. (2016). Model-based clustering based on sparse finite Gaussian mixtures. *Statistics and Computing* **26**, 303–324. MR3439375

Meng, X.-L. and Schilling, S. (1996). Fitting full-information item factor models and an empirical investigation of bridge sampling. *Journal of the American Statistical Association* **91**, 1254–1267.

Meng, X.-L. and Wong, W. H. (1996). Simulating ratios of normalizing constants via a simple identity: A theoretical exploration. *Statistica Sinica* **6**, 831–860. MR1422406

Polson, N. G., Scott, J. G. and Windle, J. (2013). Bayesian inference for logistic models using Pólya–Gamma latent variables. *Journal of the American Statistical Association* **108**, 1339–1349. MR3174712

Richardson, S. and Green, P. J. (1997). On Bayesian analysis of mixtures with an unknown number of components. *Journal of the Royal Statistical Society, Series B* **59**, 731–792. MR1483213

Robert, C. P. and Casella, G. (1999). *Monte Carlo Statistical Methods*. *Springer Series in Statistics*. New York/Berlin/Heidelberg: Springer. MR1707311

Rousseau, J., Grazian, C. and Lee, J. E. (2019). Bayesian mixture models: Theory and methods. In *Handbook of Mixture Analysis* (S. Frühwirth-Schnatter, G. Celeux and C. P. Robert, eds.) 53–72. Boca Raton, FL: CRC Press. MR3889980

Schwerdt, G., Ichino, A., Ruf, O., Winter-Ebmer, R. and Zweimüller, J. (2010). Does the color of the collar matter? Employment and earnings after plant closure. *Economics Letters* **108**, 137–140.

Zweimüller, J., Winter-Ebmer, R., Lalive, R., Kuhn, A., Wuellrich, J.-P., Ruf, O. and Büchi, S. (2009). The Austrian Social Security Database (ASSD). Working Paper 0903, NRN: The Austrian Center for Labor Economics and the Analysis of the Welfare State, Linz, Austria.

Brazilian Journal of Probability and Statistics
2019, Vol. 33, No. 4, 734–744
https://doi.org/10.1214/19-BJPS435

The limiting distribution of the Gibbs sampler for the intrinsic conditional autoregressive model

Marco A. R. Ferreira

Virginia Tech

Abstract. We study the limiting behavior of the one-at-a-time Gibbs sampler for the intrinsic conditional autoregressive model with centering on the fly. The intrinsic conditional autoregressive model is widely used as a prior for random effects in hierarchical models for spatial modeling. This model is defined by full conditional distributions that imply an improper joint "density" with a multivariate Gaussian kernel and a singular precision matrix. To guarantee propriety of the posterior distribution, usually at the end of each iteration of the Gibbs sampler the random effects are centered to sum to zero in what is widely known as centering on the fly. While this works well in practice, this informal computational way to recenter the random effects obscures their implied prior distribution and prevents the development of formal Bayesian procedures. Here we show that the implied prior distribution, that is, the limiting distribution of the one-at-a-time Gibbs sampler for the intrinsic conditional autoregressive model with centering on the fly is a singular Gaussian distribution with a covariance matrix that is the Moore–Penrose inverse of the precision matrix. This result has important implications for the development of formal Bayesian procedures such as reference priors and Bayes-factor-based model selection for spatial models.

References

Bell, B. S. and Broemeling, L. D. (2000). A Bayesian analysis for spatial processes with application to disease mapping. *Statistics in Medicine* **19**, 957–974.

Besag, J. and Kooperberg, C. (1995). On conditional and intrinsic autoregressions. *Biometrika* **82**, 733–746. MR1380811

Besag, J., York, J. and Mollié, A. (1991). Bayesian image restoration, with two applications in spatial statistics. *Annals of the Institute of Statistical Mathematics* **43**, 1–20. MR1105822

Clayton, D. and Kaldor, J. (1987). Empirical Bayes estimates of age-standardized relative risks for use in disease mapping. *Biometrics* **43**, 671–681.

De Oliveira, V. and Ferreira M. A. R. (2011). Maximum likelihood and restricted maximum likelihood estimation for a class of Gaussian Markov random fields. *Metrika* **74**, 167–183. MR2822154

Ferreira, M. A. R., Bi, Z., West, M., Lee, H. and Higdon, D. (2003). Multi-modelling of 1-D permeability fields. In *Bayesian Statistics 7* (J. M. Bernardo, M. J. Bayarri, J. O. Berger, A. P. Dawid, D. Heckerman, A. F. M. Smith and M. West, eds.) 519–527. Oxford: Oxford University Press. MR2003196

Ferreira M. A. R. and De Oliveira, V. (2007). Bayesian reference analysis for Gaussian Markov random fields. *Journal of Multivariate Analysis* **98**, 789–812. MR2322129

Key words and phrases. Areal data, ICAR models, Markov random fields, spatial data.

Ferreira, M. A. R. and Lee, H. K. H. (2007). *Multiscale Modeling: A Bayesian Perspective*. *Springer Series in Statistics*. New York: Springer. MR2348732

Gamerman, D. and Lopes, H. F. (2006). *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*, 2nd ed. Boca Raton: CRC Press. MR2260716

Gelfand, A. E. and Smith, A. F. M. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association* **85**, 398–409. MR1141740

Goicoa, T., Ugarte, M. D., Etxeberria, J. and Militino, A. F. (2016). Age–space–time CAR models in Bayesian disease mapping. *Statistics in Medicine* **35**, 2391–2405. MR3513520

Keefe, M. J., Ferreira, M. A. R. and Franck, C. T. (2018). On the formal specification of sum-zero constrained intrinsic conditional autoregressive models. *Spatial Statistics* **24**, 54–65. MR3795058

Keefe, M. J., Ferreira, M. A. R. and Franck, C. T. (2019). Objective Bayesian analysis for Gaussian hierarchical models with intrinsic conditional autoregressive priors. *Bayesian Analysis* **14**, 181–209. MR3910043

Lee, H. K. H., Higdon, D., Bi, Z., Ferreira, M. A. R. and West, M. (2002). Markov random field models for high-dimensional parameters in simulations of fluid flow in porous media. *Technometrics* **44**, 230–241. MR1951639

Liu, Z., Berrocal, V. J., Bartsch, A. J. and Johnson, T. D. (2016). Pre-surgical fMRI data analysis using a spatially adaptive conditionally autoregressive model. *Bayesian Analysis* **11**, 599–625. MR3472004

Prado, R. and West, M. (2010). *Time Series: Modeling, Computation, and Inference*. Boca Raton: CRC Press. MR2655202

Reinsel, G. C. (1997). *Elements of Multivariate Time Series Analysis*, 2nd ed. Berlin: Springer. MR1451875

Robert, C. and Casella, G. (2005). *Monte Carlo Statistical Methods*, 2nd ed. New York: Springer. MR1707311

Wu, H. H., Ferreira, M. A. R., Elkhouly, M. and Ji, T. (2019). Nonlocal priors for variable selection in generalized linear models. *Sankhya Series A*. DOI:10.1007/s13171-018-0151-9. To appear.

Wu, H. H., Ferreira, M. A. R. and Gompper, M. E. (2016). Consistency of hyper-g-prior-based Bayesian variable selection for generalized linear models. *Brazilian Journal of Probability and Statistics* **30**, 691–709. MR3582395

Brazilian Journal of Probability and Statistics
2019, Vol. 33, No. 4, 745–755
https://doi.org/10.1214/19-BJPS442

Bayesian hypothesis testing: Redux

Hedibert F. Lopes[a] **and Nicholas G. Polson**[b]

[a]*Insper Institute of Education and Research*
[b]*University of Chicago Booth School of Business*

Abstract. Bayesian hypothesis testing is re-examined from the perspective of an *a priori* assessment of the test statistic distribution under the alternative. By assessing the distribution of an observable test statistic, rather than prior parameter values, we revisit the seminal paper of Edwards, Lindman and Savage (*Psychol. Rev.* **70** (1963) 193–242). There are a number of important take-aways from comparing the Bayesian paradigm via Bayes factors to frequentist ones. We provide examples where evidence for a Bayesian strikingly supports the null, but leads to rejection under a classical test. Finally, we conclude with directions for future research.

References

Bartlett, M. S. (1957). Comment on “A statistical paradox”. *Biometrika* **44**, 533–534. MR0086727

Berger, J. O. (2003). Could Fisher, Jeffreys and Neyman have agreed on testing? *Statistical Science* **18**, 1–32. MR1997064

Berger, J. O. and Delampady, M. (1987). Testing precise hypothesis (with discussion). *Statistical Science* **2**, 317–335. MR0920141

Berger, J. O. and Jefferys, W. H. (1992). The application of robust Bayesian analysis to hypothesis testing and Occam’s razor. *Journal of the Italian Statistical Society* **1**, 17–32.

Berger, J. O. and Sellke, T. (1987). Testing of a point null hypothesis: The irreconcilability of significance levels and evidence (with discussion). *Journal of the American Statistical Association* **82**, 112–139. MR0883340

Berkson, J. (1938). Some difficulties of interpretation encountered in the application of the chi-square test. *Journal of the American Statistical Association* **33**, 526–542. MR0002074

Connolly, R. A. (1991). A posterior odds analysis of the weekend effect. *Journal of Econometrics* **49**, 51–104.

Edwards, W., Lindman, H. and Savage, L. J. (1963). Bayesian statistical inference for psychological research. *Psychological Review* **70**, 193–242.

Efron, B. and Gous, A. (2001). Scales of evidence for model selection: Fisher versus Jeffreys. In *Model Selection*, 208–246. Beachwood: Institute of Mathematical Statistics. MR2000754

Etz, A. and Wagenmakers, E.-J. (2017). J. B. S. Haldane’s contribution to the Bayes factor hypothesis test. *Statistical Science* **32**, 313–329. MR3648962

Gelman, A., Jakulin, A., Pitau, M. G. and Su, Y.-S. (2008). A weakly informative default prior distribution for logistic and other regression models. *Annals of Applied Statistics* **2**, 1360–1383. MR2655663

Good, I. J. (1992). The Bayes/non-Bayes compromise: A brief review. *Journal of the American Statistical Association* **87**, 597–606. MR1185188

Key words and phrases. Bayesian, hypothesis testing, Bayes factor, p-value, test statistic, multiple comparisons.

Hartigan, J. A. (2003). Akaike–Jeffreys priors. Technical report, Yale Univ.
Held, L. and Ott, M. (2018). On p-values and Bayes factors. *Annual Review of Statistics and Its Application* **5**, 393–419. MR3774753
Jefferys, W. H. and Berger, J. O. (1992). Ockham's razor and Bayesian analysis. *American Scientist* **80**, 64–72.
Jeffreys, H. (1957). *Scientific Inference*. Cambridge: Cambridge University Press.
Jeffreys, H. (1961). *Theory of Probability*. London: Oxford University Press. MR0187257
Jeffreys, H. (1961). *Theory of Probability*, 3rd ed. *Oxford Classic Texts in the Physical Sciences*. Oxford: Oxford Univ. Press. MR1647885
Johnson, V. (2005). Bayes factors based on test statistics. *Journal of the Royal Statistical Society, Series B* **67**, 689–701. MR2210687
Johnson, V. (2008). Properties of Bayes factors based on test statistics. *Scandinavian Journal of Statistics* **35**, 354–368. MR2418746
Kass, R. E. and Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association* **90**, 773–795. MR3363402
Lehmann, E. L. (1959). *Testing Statistical Hypotheses*. New York: Wiley. MR0107933
Lindley, D. V. (1957). A statistical paradox. *Biometrika* **44**, 187–192. MR0087273
Lopes, H. F. and West, M. (2004). Bayesian model assessment in factor analysis. *Statistica Sinica* **14**, 41–67. MR2036762
Polson, N. G. and Roberts, G. O. (1994). Bayes factors for discrete observations from diffusion processes. *Biometrika* **81**, 11–26. MR1279652
Savage, L. J. (1962). Subjective probability and statistical practice. In *The Foundations of Statistical Inference: A Discussion* (L. J. Savage et al., eds.). New York: Wiley. MR0063582
Scott, J. G. and Berger, J. O. (2010). Bayes and empirical-Bayes multiplicity adjustment in the variable-selection problem. *The Annals of Statistics* **38**, 2587–2619. MR2722450
Zellner, A. and Siow, A. (1979). Posterior odds ratio for selected regression hypotheses. In *Bayesian Statistics* (J. M. Bernardo, M. H. De Groot, D. V. Lindley and A. F. M. Smith, eds.), *Proceedings of the First International Meeting*, 585–603. Valencia: University Press. MR0638871

Brazilian Journal of Probability and Statistics
2019, Vol. 33, No. 4, 756–781
https://doi.org/10.1214/19-BJPS437

Time series of count data: A review, empirical comparisons and data analysis

Glaura C. Franco[a], **Helio S. Migon**[b] **and Marcos O. Prates**[a]

[a]*Department of Statistics, Universidade Federal de Minas Gerais*
[b]*Department of Statistical Methods, Universidade Federal do Rio de Janeiro*

Abstract. Observation and parameter driven models are commonly used in the literature to analyse time series of counts. In this paper, we study the characteristics of a variety of models and point out the main differences and similarities among these procedures, concerning parameter estimation, model fitting and forecasting. Alternatively to the literature, all inference was performed under the Bayesian paradigm. The models are fitted with a latent AR(p) process in the mean, which accounts for autocorrelation in the data. An extensive simulation study shows that the estimates for the covariate parameters are remarkably similar across the different models. However, estimates for autoregressive coefficients and forecasts of future values depend heavily on the underlying process which generates the data. A real data set of bankruptcy in the United States is also analysed.

References

Andrieu, C., Doucet, A. and Holenstein, R. (2010). Particle Markov chain Monte Carlo methods (with discussion). *Journal of the Royal Statistical Society, Series B, Statistical Methodology* **72**, 1–33. MR2758115

Benjamin, M. A., Rigby, R. A. and Stasinopoulos, D. M. (2003). Generalized autoregressive moving average models. *Journal of the American Statistical Association* **98**, 214–223. MR1965687

Box, G. E. P. and Jenkins, G. M. (1976). *Time Series Analysis: Forecasting and Control*. San Francisco: Holden-Day. MR0436499

Chan, K. and Ledolter, J. (1995). Monte Carlo EM estimation for time series models involving counts. *Journal of the American Statistical Association* **90**, 242–251. MR1325132

Cox, D. R. (1981). Statistical analysis of time series: Some recent developments. *Scandinavian Journal of Statistics* **8**, 93–115. MR0623586

Creal, D., Koopman, S. J. and Lucas, A. (2013). Generalized autoregressive score models with applications. *Journal of Applied Econometrics* **28**, 777–795. MR3083971

Creal, D. D. (2012). A survey of sequential Monte Carlo methods for economics and finance. *Econometric Reviews* **31**, 245–296.

Czado, C. and Kolbe, A. (2004). Empirical study of intraday option price changes using extended count regression models. Germany. Collaborative Research Center 386, Discussion Paper 403. Ludwig-Maximilians-Universitat.

Davis, R. A., Dunsmuir, W. T. M. and Streett, S. B. (2003). Observation-driven models for Poisson counts. *Biometrika* **90**, 777–790. MR2024757

Key words and phrases. Observation driven model, parameter driven model, autoregressive processes, Bayesian inference.

Davis, R. A., Dunsmuir, W. T. M. and Streett, S. B. (2005). Maximum likelihood estimation for an observation driven model for Poisson counts. *Methodology and Computing in Applied Probability* **7**, 149–159. MR2189786

Davis, R. A., Dunsmuir, W. T. M. and Wang, Y. (1999). Modelling time series of counts. In *Asymptotics, Nonparametrics, and Time Series: A Tribute to Madan Lal Puri* (S. Ghosh, ed.) 63–113. NewYork: Marcel Dekker.

Diebold, F. X. and Schuermann, T. (1996). Exact maximum likelihood estimation of observation-driven econometric models. NBER Technical Working Paper No. 194.

Durbin, J. and Koopman, S. J. (2002). A simple and efficient simulation smoother for state space time series analysis. *Biometrika* **89**, 603–615. MR1929166

Gamerman, D. and Lopes, H. F. (2006). *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*. London: Chapman and Hall. MR2260716

Gamerman, D., Santos, T. R. and Franco, G. C. (2013). A non-Gaussian family of state-space models with exact marginal likelihood. *Journal of Time Series Analysis* **34**, 625–645. MR3127211

Gelman, A. (1996). Inference and monitoring convergence. In *Markov Chain Monte Carlo in Practice* (W. R. Gilks, S. Richardson and D. J. Spiegelhalter, eds.) 131–143. London: Chapman and Hall. MR1397966

Goudie, R. J. B., Turner, R. M., De Angelis, D., Thomas, A. and MultiBUGS (2017). A parallel implementation of the BUGS modelling framework for faster Bayesian inference. Available at https://arxiv.org/abs/1704.03216.

Harvey, A. C. and Fernandez, C. (1989). Time series models for count or qualitative observations. *Journal of Business and Economic Statistics* **7**, 407–417.

Jung, R. C., Kukuk, M. and Liesenfeld, R. (2006). Time series of count data: Modeling, estimation and diagnostics. *Computational Statistics & Data Analysis* **51**, 2350–2364. MR2307505

Jung, R. C. and Tremayne, A. R. (2011). Useful models for time series of counts or simply wrong ones?. *AStA Advances in Statistical Analysis* **95**, 59–91. MR2775582

Koopman, S. J., Lucas, A. and Scharth, M. (2016). Predicting time-varying parameters with parameter-driven and observation-driven models. *Review of Economics and Statistics* **98**, 97–110.

Li, W. K. (1994). Time series models based on generalized linear models: Some further results. *Biometrics* **50**, 506–511.

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.

Meyer, R. and Yu, J. (2000). BUGS for a Bayesian analisys of stochastic volatility models. *Econometrics Journal* **3**, 198–215.

Migon, H. S., Gamerman, D., Lopes, H. F. and Ferreira, M. A. R. (2005). *Dynamic Models*. *Handbook of Statistics* **25**. Amsterdam: Elsevier. MR2490539

R Core Team (2018). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Rue, H., Martino, S. and Chopin, N. (2009). Approximate Bayesian inference for latent Gaussian models using integrated nested Laplace approximations (with discussion). *Journal of the Royal Statistical Society, Series B, Statistical Methodology* **71**, 319–392. MR2649602

Shephard, N. (1995). Generalized linear autoregressions. Oxford, Nuffield College. Working Paper.

Shephard, N. and Pitt, M. K. (1995). Likelihood analysis of non-Gaussian measurement time series. *Biometrika* **84**, 653–667. MR1603940

Smith, R. L. and Miller, J. E. (1986). A non-Gaussian state space model and application to prediction of records. *Journal of the Royal Statistical Society, Series B, Statistical Methodology* **48**, 79–88. MR0848053

Souza, M. A. O. and Migon, H. S. (2018). Extended dynamic generalized linear models: The two-parameter exponential family. *Computational Statistics & Data Analysis* **121**, 164–179. MR3759205

Tierney, L. and Kadane, J. B. (1986). Accurate approximations for posterior moments and marginal densities. *Journal of the American Statistical Association* **81**, 82–86. MR0830567

West, M., Harrison, J. and Migon, H. S. (1985). Dynamic generalized linear models and Bayesian forecasting. *Journal of the American Statistical Association* **80**, 73–83. MR0786598

Zeger, S. L. (1988). A regression model for time series of counts. *Biometrika* **75**, 621–629. MR0995107

Zeger, S. L. and Qaqish, B. (1988). Markov regression models for time series: A quasi-likelihood approach. *Biometrics* **44**, 1019–1031. MR0980997

Brazilian Journal of Probability and Statistics
2019, Vol. 33, No. 4, 782–800
https://doi.org/10.1214/19-BJPS443

Bayesian modelling of the abilities in dichotomous IRT models via regression with missing values in the covariates

Flávio B. Gonçalves and Bárbara C. C. Dias

Universidade Federal de Minas Gerais

Abstract. Educational assessment usually considers a contextual questionnaire to extract relevant information from the applicants. This may include items related to socio-economical profile as well as items to extract other characteristics potentially related to applicant's performance in the test. A careful analysis of the questionnaires jointly with the test's results may evidence important relations between profiles and test performance. The most coherent way to perform this task in a statistical context is to use the information from the questionnaire to help explain the variability of the abilities in a joint model-based approach. Nevertheless, the responses to the questionnaire typically present missing values which, in some cases, may be missing not at random. This paper proposes a statistical methodology to model the abilities in dichotomous IRT models using the information of the contextual questionnaires via linear regression. The proposed methodology models the missing data jointly with the all the observed data, which allows for the estimation of the former. The missing data modelling is flexible enough to allow the specification of missing not at random structures. Furthermore, even if those structures are not assumed a priori, they can be estimated from the posterior results when assuming missing (completely) at random structures a priori. Statistical inference is performed under the Bayesian paradigm via an efficient MCMC algorithm. Simulated and real examples are presented to investigate the efficiency and applicability of the proposed methodology.

References

Alves, M. T. G., Xavier, F. P., Barbosa, L. E., de Figueiredo Caldeira, B., Silva, C. A. S. and Soares, J. F. (2015). Fatores contextuais das escolas de educação básica brasileiras: Dados, métodos e aplicações. *Reuniões da ABAVE* **8**, 57–76.

da Silva Fernandes, N., Soares, T. M., Pena, A. C. and Cunha, I. C. (2010). O conhecimento do professor em avaliação educacional e a proficiência do aluno. *Estudos em Avaliação Educacional* **21**, 569–590.

Fox, J. (2005). Multilevel irt using dichotomous and polytomous response data. *British Journal of Mathematical & Statistical Psychology* **58**, 145–172.

Gonçalves, F. B., Dias, B. C. C. and Soares, T. M. (2018). Bayesian item response model: A generalised approach for the abilities' distribution using mixtures. *Journal of Statistical Computation and Simulation* **88**, 967–981.

Laros, J. A. and Marciano, J. L. (2008). Índices educacionais associados à proficiência em língua portuguesa: Um estudo multinível. *Avaliação Psicológica* **7**, 371–389.

Key words and phrases. 3PNO model, Bayesian inference, MCMC.

Richardson, S. and Green, P. J. (1997). On Bayesian analysis of mixtures with an unknown number of components (with discussion). *Journal of the Royal Statistical Society, Series B* **59**, 731–792.

Rubin, D. B. (1976). Inference and missing data. *Biometrika* **63**, 581–592.

Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. *Psychometrika monograph supplement.*

Soares, J. F. and Alves, M. T. G. (2003). Desigualdades raciais no sistema brasileiro de educação básica. *Educação e Pesquisa* **29**, 147–165.

Zwinderman, A. H. (1991). A generalized rasch model for manifest predictors. *Psychometrika* **56**, 589–600.

Brazilian Journal of Probability and Statistics
2019, Vol. 33, No. 4, 801–825
https://doi.org/10.1214/19-BJPS445

Option pricing with bivariate risk-neutral density via copula and heteroscedastic model: A Bayesian approach

Lucas Pereira Lopes[a]**, Vicente Garibay Cancho**[b] **and Francisco Louzada**[b]

[a]*Universidade de São Paulo e Universidade Federal de São Carlos—ICMC USP/UFSCar*
[b]*Universidade de São Paulo—ICMC USP*

Abstract. Multivariate options are adequate tools for multi-asset risk management. The pricing models derived from the pioneer Black and Scholes method under the multivariate case consider that the asset-object prices follow a Brownian geometric motion. However, the construction of such methods imposes some unrealistic constraints on the process of fair option calculation, such as constant volatility over the maturity time and linear correlation between the assets. Therefore, this paper aims to price and analyze the fair price behavior of the call-on-max (bivariate) option considering marginal heteroscedastic models with dependence structure modeled via copulas. Concerning inference, we adopt a Bayesian perspective and computationally intensive methods based on Monte Carlo simulations via Markov Chain (MCMC). A simulation study examines the bias, and the root mean squared errors of the posterior means for the parameters. Real stocks prices of Brazilian banks illustrate the approach. For the proposed method is verified the effects of strike and dependence structure on the fair price of the option. The results show that the prices obtained by our heteroscedastic model approach and copulas differ substantially from the prices obtained by the model derived from Black and Scholes. Empirical results are presented to argue the advantages of our strategy.

References

Abanto-Valle, C. A., Lachos, V. H. and Dey, K. (2015). Bayesian estimation of a skew-student-t stochastic volatility model. *Methodology and Computing in Applied Probability* **17**, 721–738. MR3377857

Ausin, M. C. and Lopes, H. F. (2010). Time-varying joint distribution through copulas. *Computational Statistics & Data Analysis* **54**, 2383–2399. MR2720449

Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy* **81**, 637–654. MR3363443

Cherubini, U. and Luciano, E. (2002). Bivariate option pricing with copulas. *Applied Mathematical Finance* **9**, 69–85.

Chiou, S. C. and Tsay, R. S. (2008). A copula-based approach to option pricing and risk assessment. *Journal of Data Science* **6**, 273–301.

Delatola, E. I. and Griffin, J. E. (2011). Bayesian nonparametric modelling of the return distribution with stochastic volatility. *Bayesian Analysis* **6**, 901–926. MR2869968

Delbaen, F. and Schachermayer, W. (1994). A general version of the fundamental theorem of asset pricing. *Mathematische Annalen* **300**, 463–520. MR1304434

Key words and phrases. Option pricing, heterocedastic, copula, Bayesian inference.

Duan, J. C. (1995). The GARCH option pricing model. *Mathematical Finance* **5**, 13–32. MR1322698

Embrechts, P., McNeil, A. and Straumann, D. (2002). Correlation and dependence in risk management: Properties and pitfalls. *Risk Management: Value at Risk and Beyond*. MR1892190

Fonseca, T. C., Migon, H. S. and Ferreira, M. A. (2012). Bayesian analysis based on the Jeffreys prior for the hyperbolic distribution. *Brazilian Journal of Probability and Statistics*, 327–343. MR2949082

Forbes, K. J. and Rigobon, R. (2002). No contagion, only interdependence: Measuring stock market comovements. *The Journal of Finance* **57**, 2223–2261.

Franses, P. H. and Van Dijk, D. (2000). *Non-linear Time Series Models in Empirical Finance*. Cambridge: Cambridge University Press. MR1739079

French, K. R., William, S. G. and Stambaugh, R. F. (1987). Expected stock returns and volatility. *Journal of Financial Economics* **19**, 3–29.

Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In *Bayesian Statistics 4* (J. M. Bernardo, J. O. Berger, A. P. Dawid and A. F. M. Smith, eds.) 169–193. London: Oxford University Press. MR1380276

Haug, E. G. (2007). *The Complete Guide to Option Pricing Formulas, Vol. 2*. New York: McGraw-Hill.

Hull, J. (1992). *Introduction to Futures and Options Markets*. Englewood Cliffs, NJ: Prentice Hall.

Hull, J. and White, A. (1987). The pricing of options on assets with stochastic volatilities. *The Journal of Finance* **42**, 281–300.

Hürlimann, W. (2004). Fitting bivariate cumulative returns with copulas. *Computational Statistics & Data Analysis* **45**, 355–372. MR2045637

Johnson, H. and David, S. (1987). Option pricing when the variance is changing. *Journal of Financial and Quantitative Analysis*, 143–151.

Kang, T. and Brorsen, B. W. (1993). GARCH option pricing with asymmetry. In *Proceedings of the NCR-134 Conference on Applied Commodity, Forecasting, and Market Risk Management*.

Klugman, S. A. and Parsa, R. (1999). Fitting bivariate loss distributions with copulas. *Insurance Mathematics & Economics* **24**(1–2), 139–148. MR1710816

Leão, W. L., Abanto-Valle, C. A. and Chen, M. H. (2017). Bayesian analysis of stochastic volatility-in-mean model with leverage and asymmetrically heavy-tailed error using generalized hyperbolic skew student's t-distribution. *Statistics and its Interface* **10**, 529–541. MR3662770

Lopes, L. P. and Pessanha, G. R. G. (2018). Análise de dependência entre mercados financeiros: Uma abordagem do modelo copula-GARCH. *Revista de Financas e Contabilidade da Unimep.*

Madan, D. B. and Milne, F. (1991). Option pricing with VG martingale components. *Mathematical Finance* **1**, 39–55.

Margrabe, W. (1978). The value of an option to exchange one asset for another. *The Journal of Finance* **33**, 177–186.

Merton, R. C. (1973). Theory of rational option pricing. *The Bell Journal of Economics and Management Science*, 141–183. MR0496534

Nelsen, R. B. (2006) *An Introduction to Copulas*, 2nd ed. New York: Springer Science Business Media. MR2197664

Rosenberg, J. V. (2002). Nonparametric pricing of multivariate contingent claims. Technical Report.

Rossi, J. L., Ehlers, R. S. and Andrade, M. G. (2012). Copula-GARCH Model Selection: A Bayesian Approach. University of São Paulo, Technical Report 88.

Sanfins, M. A. and Valle, G. (2012). On the copula for multivariate extreme value distributions. *Brazilian Journal of Probability and Statistics* **26**, 288–305. MR2911707

Schmidt, T. (2007). *Coping with Copulas. Copulas-from Theory to Application in Finance*. Risk Books.

Sharifonnasabi, Z., Alamatsaz, M. H. and Kazemi, I. (2018). A large class of new bivariate copulas and their properties. *Brazilian Journal of Probability and Statistics* **32**, 497–524. MR3812379

Shimko, D. C. (1994). Options on futures spreads: Hedging, speculation, and valuation. *The Journal of Futures Markets* **14**, 183–213.

Sklar, A. (1959). Fonctions de repartition a n dimensions et leures marges. *Publications de L'Institut de Statistique de L'Université de Paris* **8**, 229–231. MR0125600

Spiegelhalter, D. J., et al (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society, Series B, Statistical Methodology* **64**, 583–639. MR1979380

Stulz, R. (1982). Options on the minimum or the maximum of two risky assets: Analysis and applications. *Journal of Financial Economics* **10**(2), 161–185.

Zhang, J. and Guegan, D. (2008). Pricing bivariate option under GARCH processes with time-varying copula. *Insurance Mathematics & Economics* **42**, 1095–1103. MR2435379

Brazilian Journal of Probability and Statistics
2019, Vol. 33, No. 4, 826–860
https://doi.org/10.1214/19-BJPS447

Bayesian approach for the zero-modified Poisson–Lindley regression model

Wesley Bertoli[a], Katiane S. Conceição[b], Marinho G. Andrade[b] and Francisco Louzada[b]

[a]*Federal University of Technology—Paraná*
[b]*University of São Paulo*

Abstract. The primary goal of this paper is to introduce the zero-modified Poisson–Lindley regression model as an alternative to model overdispersed count data exhibiting inflation or deflation of zeros in the presence of covariates. The zero-modification is incorporated by considering that a zero-truncated process produces positive observations and consequently, the proposed model can be fitted without any previous information about the zero-modification present in a given dataset. A fully Bayesian approach based on the *g-prior* method has been considered for inference concerns. An intensive Monte Carlo simulation study has been conducted to evaluate the performance of the developed methodology and the maximum likelihood estimators. The proposed model was considered for the analysis of a real dataset on the number of bids received by 126 U.S. firms between 1978–1985, and the impact of choosing different *prior* distributions for the regression coefficients has been studied. A sensitivity analysis to detect influential points has been performed based on the Kullback–Leibler divergence. A general comparison with some well-known regression models for discrete data has been presented.

References

Angers, J. F. and Biswas, A. (2003). A Bayesian analysis of zero-inflated generalized Poisson model. *Computational Statistics & Data Analysis* **42**, 37–46. MR1963008

Bahn, G. D. and Massenburg, R. (2008). Deal with excess zeros in the discrete dependent variable, the number of homicide in Chicago census tract. In *Joint Statistical Meetings of the American Statistical Association*, 3905–3912.

Bazán, J. L., Torres-Avilés, F., Suzuki, A. K. and Louzada, F. (2017). Power and reversal power links for binary regressions: An application for motor insurance policyholders. *Applied Stochastic Models in Business and Industry* **33**, 22–34. MR3615258

Bertoli, W., Conceição, K. S., Andrade, M. G. and Louzada, F. (2018). On the zero-modified Poisson–Shanker regression model and its application to fetal deaths notification data. *Computational Statistics* **33**, 807–836. MR3809623

Beuf, K. D., Schrijver, J. D., Thas, O., Criekinge, W. V., Irizarry, R. A. and Clement, L. (2012). Improved base-calling and quality scores for 454 sequencing based on a hurdle Poisson model. *BMC Bioinformatics* **13**, 303.

Key words and phrases. Bayesian inference, influential points, overdispersion, Poisson–Lindley distribution, regression models, zero-modified data.

Bohara, A. K. and Krieg, R. G. (1996). A zero-inflated Poisson model of migration frequency. *International Regional Science Review* **19**, 211–222.

Brooks, S. P. (2002). Discussion on the paper by Spiegelhalter, Best, Carlin and van der Linde.

Bulmer, M. G. (1974). On fitting the Poisson–Lognormal distribution to species-abundance data. *Biometrics* 101–110.

Carlin, B. P. and Louis, T. A. (2010). *Bayes and Empirical Bayes Methods for Data Analysis*. Boca Raton: Chapman & Hall/CRC.

Chen, M. H. and Ibrahim, J. G. (2003). Conjugate priors for generalized linear models. *Statistica Sinica* 461–476. MR1977737

Cohen, A. C. (1960). An extension of a truncated Poisson distribution. *Biometrics* **16**, 446–450. MR0121900

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. MR3097373

Conceição, K. S., Louzada, F., Andrade, M. G. and Helou, E. S. (2017). Zero-modified power series distribution and its hurdle distribution version. *Journal of Statistical Computation and Simulation* **87**, 1842–1862. MR3635033

Csiszár, I. (1967). Information-type measures of difference of probability distributions and indirect observations. *Studia Scientiarum Mathematicarum Hungarica* **2**, 299–318. MR0219345

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

Geisser, S. (1993). *Predictive Inference: An Introduction*. *Chapman & Hall/CRC Monographs on Statistics & Applied Probability.* Boca Raton: CRC Press. MR1252174

Gelman, A. and Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science* 457–472.

Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. *Bayesian Statistics* **4**, 169–193.

Ghitany, M. E., Al-Mutairi, D. K. and Nadarajah, S. (2008). Zero-truncated Poisson–Lindley distribution and its application. *Mathematics and Computers in Simulation* **79**, 279–287. MR2477530

Ghosh, S. K., Mukhopadhyay, P. and Lu, J. C. (2006). Bayesian analysis of zero-inflated regression models. *Journal of Statistical Planning and Inference* **136**, 1360–1375. MR2253768

Gupta, M. and Ibrahim, J. G. (2009). An information matrix prior for Bayesian analysis in generalized linear models with high dimensional data. *Statistica Sinica* **19**, 1641–1663. MR2589202

Gurmu, S. and Trivedi, P. K. (1996). Excess zeros in count models for recreational trips. *Journal of Business & Economic Statistics* **14**, 469–477.

Heilbron, D. C. and Gibson, D. R. (1990). Shared needle use and health beliefs concerning AIDS: Regression modeling of zero-heavy count data. Poster session. In *Proceedings of the Sixth International Conference on AIDS*. *San Francisco, CA*.

Hörmann, W., Leydold, J. and Derflinger, G. (2004). *Automatic Nonuniform Random Variate Generation*. Berlin: Springer. MR2143197

Hu, M. C., Pavlicova, M. and Nunes, E. V. (2011). Zero-inflated and hurdle models of count data with extra zeros: Examples from an HIV-risk reduction intervention trial. *The American Journal of Drug and Alcohol Abuse* **37**, 367–375.

Jaggia, S. and Thosar, S. (1993). Multiple bids as a consequence of target management resistance: A count data approach. *Review of Quantitative Finance and Accounting* **3**, 447–457.

Karlis, D. and Xekalaki, E. (2005). Mixed Poisson distributions. *International Statistical Review* **73**, 35–58.

Lambert, D. (1992). Zero-inflated Poisson regression, with an application to defects in manufacturing. *Technometrics* **34**, 1–14.

Lindley, D. V. (1958). Fiducial distributions and Bayes' theorem. *Journal of the Royal Statistical Society, Series B (Methodological)* 102–107. MR0095550

McCulloch, R. E. (1989). Local model influence. *Journal of the American Statistical Association* **84**, 473–478.

McDowell, A. (2003). From the help desk: Hurdle models. *Stata Journal* **3**, 178–184.

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

Mullahy, J. (1986). Specification and testing of some modified count data models. *Journal of Econometrics* **33**, 341–365. MR0867980

Ngatchou-Wandji, J. and Paris, C. (2011). On the zero-inflated count models with application to modelling annual trends in incidences of some occupational allergic diseases in France. *Journal of Data Science* **9**, 639–659. MR2838666

Oliveira, M., Einbeck, J., Higueras, M., Ainsbury, E., Puig, P. and Rothkamm, K. (2016). Zero-inflated regression models for radiation-induced chromosome aberration data: A comparative study. *Biometrical Journal* **58**, 259–279. MR3499114

Peng, F. and Dey, D. K. (1995). Bayesian analysis of outlier problems using divergence measures. *Canadian Journal of Statistics* **23**, 199–213.

R Development Core Team (2017). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Ridout, M., Demétrio, C. G. B. and Hinde, J. (1998). Models for count data with many zeros. In *Proceedings of the XIXth International Biometric Conference, Vol. 19*, 179–192.

Roberts, G. O., Gelman, A. and Gilks, W. R. (1997). Weak convergence and optimal scaling of random walk Metropolis algorithms. *The Annals of Applied Probability* **7**, 110–120. MR1428751

Rodrigues, J. (2003). Bayesian analysis of zero-inflated distributions. *Communications in Statistics Theory and Methods* **32**, 281–289. MR1958514

Saffar, S. E., Adnan, R. and Greene, W. (2012). Parameter estimation on hurdle Poisson regression model with censored data. *Jurnal Teknologi* **57**, 189–198.

Sankaran, M. (1970). The discrete Poisson–Lindley distribution. *Biometrics* **26**, 145–149.

Shaban, S. A. (1981). On the discrete Poisson–Inverse Gaussian distribution. *Biometrical Journal* **23**, 297–303. MR0629707

Shanker, R. (2016a). The discrete Poisson–Shanker distribution. *Jacobs Journal of Biostatistics* **1**, 1–7.

Shanker, R. (2016b). The discrete Poisson–Sujatha distribution. *International Journal of Probability and Statistics* **5**, 1–9.

Shanker, R. and Fesshaye, H. (2016). On zero-truncation of Poisson, Poisson–Lindley and Poisson–Sujatha distributions and their applications. *Biometrics & Biostatistics International Journal* **3**, 1–13.

Umbach, D. (1981). On inference for a mixture of a Poisson and a degenerate distribution. *Communications in Statistics Theory and Methods* **10**, 299–306. MR0612561

Wagh, Y. S. and Kamalja, K. K. (2018). Zero-inflated models and estimation in zero-inflated Poisson distribution. *Communications in Statistics - Simulation and Computation* **47**, 2248–2265. MR3843056

Zamani, H. and Ismail, N. (2010). Negative Binomial–Lindley distribution and its application. *Journal of Mathematics and Statistics* **6**, 4–9.

Zellner, A. (1986). On assessing prior distributions and Bayesian regression analysis with g-prior distributions. *Bayesian inference and decision techniques: Essays in Honor of Bruno De Finetti* **6**, 233–243. MR0881437

Zorn, C. J. W. (1996). Evaluating zero-inflated and hurdle Poisson specifications. *Midwest Political Science Association* **18**, 1–16.

Brazilian Journal of Probability and Statistics
2019, Vol. 33, No. 4, 861–893
https://doi.org/10.1214/18-BJPS424

Subjective Bayesian testing using calibrated prior probabilities

Dan J. Spitzner

University of Virginia

Abstract. This article proposes a calibration scheme for Bayesian testing that coordinates analytically-derived statistical performance considerations with expert opinion. In other words, the scheme is effective and meaningful for incorporating *objective* elements into *subjective* Bayesian inference. It explores a novel role for default priors as anchors for calibration rather than substitutes for prior knowledge. Ideas are developed for use with multiplicity adjustments in multiple-model contexts, and to address the issue of prior sensitivity of Bayes factors. Along the way, the performance properties of an existing multiplicity adjustment related to the Poisson distribution are clarified theoretically. Connections of the overall calibration scheme to the Schwarz criterion are also explored. The proposed framework is examined and illustrated on a number of existing data sets related to problems in clinical trials, forensic pattern matching, and log-linear models methodology.

References

Bartlett, M. S. (1957). Comment on D. V. Lindley's statistical paradox. *Biometrika* **44**, 533–534.

Bayarri, M. J., Berger, J. O., Forte, A. and García-Donato, G. (2012). Criteria for Bayesian model choice with application to variable selection. *The Annals of Statistics* **40**, 1550–1577. MR3015035

Berger, J. and Pericchi, L. (2004). Training samples in objective model selection. *The Annals of Statistics* **32**, 841–869. MR2065191

Berger, J. O. and Pericchi, L. (1996). The intrinsic Bayes factor for model selection and prediction. *Journal of the American Statistical Association* **91**, 109–122. MR1394065

Berry, D. A. and Hochberg, Y. (1999). Bayesian perspectives on multiple comparisons. *Journal of Statistical Planning and Inference* **82**, 215–227. MR1736444

Berry, S. M. and Berry, D. A. (2004). Accounting for multiplicities in assessing drug safety: A three-level hierarchical mixture model. *Biometrics* **60**, 418–426. MR2066276

Billingsley, P. (1995). *Probability and Measure*, 3rd ed. New York: Wiley. MR1324786

Bollen, K., Ray, S., Zavisca, J. and Harden, J. J. (2012). A comparison of Bayes factor approximation methods including two new methods. *Sociological Methods and Research*. In press. MR3190708

Box, G. E. P. and Tiao, G. C. (1992). *Bayesian Inference in Statistical Analysis*. Reading, MA: Addison-Wesley. MR0418321

Casella, G., Girón, F. J., Martinez, M. L. and Moreno, E. (2009). Consistency of Bayesian procedures for variable selection. *The Annals of Statistics* **37**, 1207–1228. MR2509072

Castillo, I., Schmidt-Hieber, J. and van der Vaart, A. (2015). Bayesian linear regression with sparse priors. *The Annals of Statistics* **43**, 1986–2018. MR3375874

Key words and phrases. Subjective Bayes, Bayesian testing, Bayes factors, default priors, high-dimensional statistics, multiplicity, variable selection, Schwarz criterion.

Dellaportas, P., Forster, J. J. and Ntzoufras, I. (2012). Joint specification of model space and parameter space prior distributions. *Statistical Science* **27**, 232–246. MR2963994

Fan, J. and Lv, J. (2008). Sure independence screening for ultrahigh dimensional feature space. *Journal of the Royal Statistical Society, Series B* **70**, 849–911. MR2530322

Fan, J. and Lv, J. (2010). A selective overview of variable selection in high dimensional feature space. *Statistica Sinica* **20**, 101–148. MR2640659

Fouskakis, D., Ntzoufras, I. and Perrakis, K. (2018). Power-expected-posterior priors for generalized linear models. *Bayesian Analysis* **13**, 721–748. MR3807864

Good, I. J. (1950). *Probability and the Weighing of Evidence*. London: Griffin. MR0041366

Ibrahim, J. G. and Chen, M. H. (2000). Power prior distributions for regression models. *Statistical Science* **15**, 46–60. MR1842236

Jeffreys, H. (1961). *Theory of Probability*, 3rd ed. Oxford: Oxford University Press. MR0187257

Kass, R. E. and Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association* **90**, 773–795. MR3363402

Kass, R. E. and Wasserman, L. (1995). A reference Bayesian test for nested hypotheses and its relationship to the Schwarz criterion. *Journal of the American Statistical Association* **90**, 928–934. MR1354008

Lavine, M. and Schervish, M. J. (1999). Bayes factors: What they are and what they are not. *American Statistician* **53**, 119–122. MR1707756

Liang, F., Paulo, R., Molina, G., Clyde, C. A. and Berger, J. O. (2008). Mixtures of g priors for Bayesian variable selection. *Journal of the American Statistical Association* **103**, 410–423. MR2420243

Lindley, D. V. (1957). A statistical paradox. *Biometrika* **44**, 187–192.

Lindley, D. V. (1977). A problem in forensic science. *Biometrika* **64**, 207–213. MR0518935

Lund, S. P. and Iyer, H. (2017). Likelihood ratio as weight of forensic evidence: a closer look. *Journal of Research of National Institute of Standards and Technology* **122**, 1–32.

Moreno, E. and Pericchi, L. R. (2014). Intrinsic priors for objective Bayesian model selection. In *Bayesian Model Comparison* (I. Jeliazkov and D. J. Poirier, eds.) 279–300. Emerald Group Publishing Limited.

Müller, P., Parmigiani, G. and Rice, K. (2007). FDR and Bayesian multiple comparison rules. In *Bayesian Statistics 8*, 349–370. Oxford: Oxford University Press. MR2433200

Narisetty, N. N. and He, X. (2014). Bayesian variable selection with shrinking and diffusing priors. *The Annals of Statistics* **42**, 789–817. MR3210987

O'Hagan, A. (1995). Fractional Bayes factors for model comparisons. *Journal of the Royal Statistical Society, Series B* **57**, 99–138. MR1325379

Pérez, J. M. and Berger, J. O. (2002). Expected-posterior prior distributions for model selection. *Biometrika* **89**, 491–511. MR1929158

Raftery, A. E. (1993). Approximate Bayes factors and accounting for model uncertainty in generalized linear models. Technical Report 255, Dept. Statistics, Univ. Washington.

Robert, C. P. (1993). A note on Jeffreys–Lindley paradox. *Statistica Sinica* **3**, 603–608. MR1243404

Robert, C. P. and Casella, G. (1999). *Monte Carlo Statistical Methods*. New York: Springer. MR1707311

Schwartz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics* **6**, 461–464. MR0468014

Scott, J. G. and Berger, J. O. (2010). Bayes and empirical Bayes multiplicity adjustment in the variable selection problem. *The Annals of Statistics* **38**, 2587–2619. MR2722450

Spitzner, D. J. (2008). An asymptotic viewpoint on high-dimensional Bayesian testing. *Bayesian Analysis* **3**, 121–160. MR2383254

Spitzner, D. J. (2011). Neutral-data comparisons for Bayesian testing. *Bayesian Analysis* **6**, 603–638. MR2869959

Tierney, L. and Kadane, J. B. (1986). Accurate approximations for posterior moments and marginal densities. *Journal of the American Statistical Association* **81**, 82–86. MR0830567

Wilson, M. A., Iversen, E. S., Clyde, M. A., Schmidler, S. C. and Schildkraut, J. M. (2010). Bayesian model search and multilevel inference for SNP association studies. *Annals of Applied Statistics* **4**, 1342–1364. MR2758331

Womack, A. J., Fuentes, C. and Taylor-Rodriguez, D. (2015). Model space priors for objective sparse Bayesian regression. Preprint. Available at https://arxiv.org/abs/1511.04745.

Zellner, A. (1986). On assessing prior distributions and Bayesian regression analysis using g-prior distributions. In *Bayesian Inference and Decision Techniques: Essays in Honor of Bruno de Finetti* (P. Goel and A. Zellner, eds.) 233–243. Amsterdam: North-Holland. MR0881437

Brazilian Journal of Probability and Statistics
2019, Vol. 33, No. 4, 894–914
https://doi.org/10.1214/18-BJPS428

Bayesian inference on power Lindley distribution based on different loss functions

Abbas Pak[a], M. E. Ghitany[b] and Mohammad Reza Mahmoudi[c]

[a]*Shahrekord University*
[b]*Kuwait University*
[c]*Fasa University*

Abstract. This paper focuses on Bayesian estimation of the parameters and reliability function of the power Lindley distribution by using various symmetric and asymmetric loss functions. Assuming suitable priors on the parameters, Bayes estimates are derived by using squared error, linear exponential (linex) and general entropy loss functions. Since, under these loss functions, Bayes estimates of the parameters do not have closed forms we use lindley's approximation technique to calculate the Bayes estimates. Moreover, we obtain the Bayes estimates of the parameters using a Markov Chain Monte Carlo (MCMC) method. Simulation studies are conducted in order to evaluate the performances of the proposed estimators under the considered loss functions. Finally, analysis of a real data set is presented for illustrative purposes.

References

Ahmed, E. A. (2014). Bayesian estimation based on progressive type II censoring from two-parameter bathtub-shaped lifetime model: An Markov chain Monte Carlo approach. *Journal of Applied Statistics* **41**, 752–768. MR3291784

Bader, M. G. and Priest, A. M. (1982). Statistical aspects of fibre and bundle strength in hyprid composites. In *International Conference on Composite Materials (ICCM-IV)* (T. Hayashi, K. Kawata and S. Umekawa, eds.), *Progress in Science and Engineering Composites* **4**, 1129–1136. *Tokyo, Japan*.

Banerjee, A. and Kundu, D. (2008). Inference based on type-II hybrid censored data from a Weibull distribution. *IEEE Transactions on Reliability* **57**, 369–378.

Gelman, A., Carlin, J. B., Stern, H. S. and Rubin, D. B. (2003). *Bayesian Data Analysis*, 2nd ed. London: Chapman and Hall. MR2027492

Ghitany, M. E., Al-Mutairi, D. K. and Aboukhamseen, S. M. (2015). Estimation of the reliability of a stress-strength system from power Lindley distributions. *Communications in Statistics Simulation and Computation* **44**, 118–136. MR3238554

Ghitany, M. E., Al-Mutairi, D. K., Balakrishnan, N. and Al-Enezi, L. J. (2013). Power lindley distribution and associated inference. *Computational Statistics & Data Analysis* **64**, 20–33. MR3061887

Lin, C. T., Chou, C. C. and Huang, Y. L. (2012). Inference for the Weibull distribution with progressive hybrid censoring. *Computational Statistics & Data Analysis* **56**, 451–467. MR2853746

Lindley, D. V. (1980). Approximate Bayesian method. *Trabajos de Estadistica* **31**, 223–237.

R Development Core Team (2011). *A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Key words and phrases. Power Lindley distribution, Bayesian estimation, maximum likelihood estimation, squared error loss function, asymmetric loss function.

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/journals-and-publications/brazilian-journal-of-probability-and-statistics/

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/journals-and-publications/brazilian-journal-of-probability-and-statistics/brazilian-journal-of-probability-and-statistics-manuscript-submission/

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/individual-membership/

Individual and General Subscriptions:
http://www.imstat.org/orders/

Electronic Access to IMS Journals:
http://www.imstat.org/journals-and-publications/electronic-access/

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