Contents

BJPS is an IMS supported journal

ISSN 0103-0752

Brazilian Statistical Association

Brazilian Journal of Probability and Statistics

Volume 37 • Number 1 • March 2023

Brazilian Journal of Probability and Statistics

Volume 37 • Number 1 • March 2023

www.imstat.org/bjps

ISSN 0103-0752 (Print) ISSN 2317-6199 (Online), Volume 37, Number 1, March 2023. 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, PO Box 729, Middletown, Maryland 21769, 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
2023, Vol. 37, No. 1, 1–25
https://doi.org/10.1214/22-BJPS558

Revisiting the Samejima–Bolfarine–Bazán IRT models: New features and extensions

Jorge Luis Bazán[1,a], Sandra Elizabeth Flores Ari[2], Caio L. N. Azevedo[3] and Dipak K. Dey[4]

[1]*Department of Applied Mathematics and Statistics, University of São Paulo, São Carlos, Brazil,* [a]*jlbazan@icmc.usp.br*
[2]*Department of Statistics, University of São Paulo, São Paulo, Brazil*
[3]*Department of Statistics, Campinas State University, Campinas, Brazil*
[4]*Department of Statistics, University of Connecticut, Storrs, Connecticut, USA*

Abstract. In 2010, the Samejima–Bolfarine–Bazán (SBB) Item Response Theory (IRT) models were introduced by (*Journal of Educational and Behavioral Statistics* **35** (2010) 693–713) under a Bayesian approach. These models extend the regular Bayesian One and Two Parameter Logistic IRT models by incorporating a parameter accounting for asymmetry of the Item Characteristic Curve (ICC) which is named the complexity of the item. It includes the Logistic Positive Exponent (LPE) IRT model formulated initially by (*Psychometrika* **65** (2000) 319–335) and the Reflection of the LPE (RLPE). In the present work, new properties of the SBB models are developed including a random effect for testlet structures with a Bayesian inference through a Markov chain Monte Carlo (MCMC) algorithm which includes the parameter estimation and model comparison. The asymmetric behavior of the Item Characteristic Curve (ICC) is detected using a marginal item information function. Two simulation studies are developed to analyze the sensitiveness of the penalized parameter in the asymmetric behavior of the ICC and to evaluate the parameter recovery of the proposed model. A real data set, with a testlet structure and empirical evidence of asymmetric behavior of the ICCs, is used to apply the models.

References

Albert, J. and Ghosh, M. (2000). Item response modeling. In *Generalized Linear Models: A Bayesian Perspective* (D. K. Dey, S. K. Ghosh and B. K. Mallick, eds.) Boca Raton: CRC Press. MR1893789

Albert, J. H. (1992). Bayesian estimation of normal ogive item response curves using Gibbs sampling. *Journal of Educational Statistics* **17**, 251–269.

Ames, A. J. and Samonte, K. (2020). Using SAS PROC MCMC for Item Response Theory Models. *Educational Psychological Measurement* **11**.

Arnold, B. C. and Groeneveld, R. A. (1995). Measuring skewness with respect to the mode. *American Statistician* **49**, 34–38. ISSN 00031305. http://www.jstor.org/stable/2684808. MR1341197 https://doi.org/10.2307/2684808

Azevedo, C. L., Bolfarine, H. and Andrade, D. F. (2011). Bayesian inference for a skew-normal IRT model under the centred parameterization. *Computational Statistics & Data Analysis* **55**, 353–365. MR2736560 https://doi.org/10.1016/j.csda.2010.05.003

Bazán, J. L., Branco, M. D., Bolfarine, H., et al (2006). A skew item response model. *Bayesian Analysis* **1**, 861–892. MR2282209 https://doi.org/10.1214/06-BA128

Bazán, J. L., Branco, M. D., Bolfarine, H., et al (2014). Extensions of the skew-normal ogive item response model. *Brazilian Journal of Probability and Statistics* **28**, 1–23. MR3165426 https://doi.org/10.1214/12-BJPS191

Bolfarine, H. and Bazán, J. L. (2010). Bayesian estimation of the logistic positive exponent IRT model. *Journal of Educational and Behavioral Statistics* **35**, 693–713.

Bolt, D. M. and Liao, X. (2022). Item Complexity: A Neglected Psychometric Feature of Test Items? *Psychometrika*. MR4504988 https://doi.org/10.1007/s11336-022-09842-0

Key words and phrases. Asymmetrical link, Bayesian estimation, Complexity, item response model, logistic positive exponent, SBB models, testlet-effect.

Bradlow, E. T., Wainer, H. and Wang, X. (1999). A Bayesian random effects model for testlets. *Psychometrika* **64**, 153–168.

Braeken, J. (2011). A boundary mixture approach to violations of conditional independence. *Psychometrika* **76**, 57–76. MR2783872 https://doi.org/10.1007/s11336-010-9190-4

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P. and Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software* **76**.

Curtis, S. M., et al (2010). BUGS code for item response theory. *Journal of Statistical Software* **36**, 1–34.

Depaoli, S., Winter, S. D. and Visser, M. (2007). The importance of prior sensitivity analysis in Bayesian statistics: Demonstrations using an interactive shiny app. *Frontiers in Psychology* **75**, 585–609.

Fox, J.-P. (2010). *Bayesian Item Response Modeling: Theory and Applications*. Berlin: Springer. MR2657265 https://doi.org/10.1007/978-1-4419-0742-4

Gelman, A., Carlin, J. B., Stern, H. S. and Rubin, D. B. (2014). *Bayesian Data Analysis*. London: Chapman & Hall/CRC. MR2027492

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

Henson, R. K. (2001). Understanding internal consistency reliability estimates: A conceptual primer on coefficient alpha. (Methods, plainly speaking). *Measurement and Evaluation in Counseling and Development* **34**, 177–190.

Huang, H.-Y. and Wang, W.-C. (2013). Higher order testlet response models for hierarchical latent traits and testlet-based items. *Educational and Psychological Measurement* **73**, 491–511.

Ip, E. H. (2010). Interpretation of the three-parameter testlet response model and information function. *Applied Psychological Measurement* **34**, 467–482.

Johnson, M. S., Sinharay, S. and Bradlow, E. T. (2006). 17 hierarchical item response theory models. In *Psychometrics* (C. R. Rao and S. Sinharay, eds.), *Handbook of Statistics* **26**, 587–606. Amsterdam: Elsevier.

Kim, J. S. and Bolt, D. M. (2007). Estimating item response theory models using Markov chain Monte Carlo methods. *Educational Measurement, Issues and Practice* **26**, 38–51.

Kim, S.-H., Cohen, A. S., Baker, F. B., Subkoviak, M. J. and Leonard, T. (1994). An investigation of hierarchical Bayes procedures in item response theory. *Psychometrika* **59**, 405–421.

Lee, S. and Bolt, D. M. (2018a). An alternative to the 3PL: Using asymmetric item characteristic curves to address guessing effects. *Journal of Educational Measurement* **55**, 90–111.

Lee, S. and Bolt, D. M. (2018b). Asymmetric item characteristic curves and item complexity: Insights from simulation and real data analyses. *Psychometrika* **83**, 453–475. MR3798136 https://doi.org/10.1007/s11336-017-9586-5

Patz, R. J. and Junker, B. W. (1999). A straightforward approach to Markov chain Monte Carlo methods for item response models. *Journal of Educational and Behavioral Statistics* **24**, 146–178.

Richards, F. J. (1959). A flexible growth function for empirical use. *Journal of Experimental Botany* **10**, 290–301.

Rijmen, F. (2010). Formal relations and an empirical comparison among the bi-factor, the testlet, and a second-order multidimensional IRT model. *Journal of Educational Measurement* **47**, 361–372.

Robitzsch, A. (2019) *sirt: Supplementary Item Response Theory Models*. R package version 3.6-21. https://CRAN.R-project.org/package=sirt.

Robitzsch, A. (2022). On the choice of the item response model for scaling PISA data: Model selection based on information criteria and quantifying model uncertainty. *Entropy* **24**, 1–26. MR4449633 https://doi.org/10.3390/e24060760

Roos, M., Martins, T. G., Held, L. and Rue, H. (2015). Sensitivity analysis for Bayesian hierarchical models. *Bayesian Analysis* **10**, 321–349. MR3420885 https://doi.org/10.1214/14-BA909

Rupp, A. A., Dey, D. K. and Zumbo, B. D. (2004). To Bayes or not to Bayes, from whether to when: Applications of Bayesian methodology to modeling. *Structural Equation Modeling* **11**, 424–451. MR2061898 https://doi.org/10.1207/s15328007sem1103_7

Sahu, S. K. (2002). Bayesian estimation and model choice in item response models. *Journal of Statistical Computation and Simulation* **72**, 217–232. MR1909259 https://doi.org/10.1080/00949650212387

Samejima, F. (1995). Acceleration model in the heterogeneous case of the general graded response model. *Psychometrika* **60**, 549–572. MR1369941 https://doi.org/10.1007/BF02294328

Samejima, F. (1997). Departure from normal assumptions: A promise for future psychometrics with substantive mathematical modeling. *Psychometrika* **62**, 471–493.

Samejima, F. (1999). Usefulness of the logistic positive exponent family of models in educational measurement. *Annual AERA Meeting*.

Samejima, F. (2000). Logistic positive exponent family of models: Virtue of asymmetric item characteristic curves. *Psychometrika* **65**, 319–335.

San Martín, E. (2018). Identifiability of structural characteristics: How relevant is it for the Bayesian approach? *Brazilian Journal of Probability and Statistics* **32**, 346–373. MR3787758 https://doi.org/10.1214/16-BJPS346

San Martín, E., González, J. and Tuerlinckx, F. (2015). On the unidentifiability of the fixed-effects 3PL model. *Psychometrika* **80**, 450–467. MR3353967 https://doi.org/10.1007/s11336-014-9404-2

Schroeders, U., Robitzsch, A. and Schipolowski, S. (2014). A comparison of different psychometric approaches to modeling testlet structures: An example with C-tests. *Journal of Educational Measurement* **51**, 400–418.

Sinharay, S. (2004). Experiences with Markov chain Monte Carlo convergence assessment in two psychometric examples. *Journal of Educational and Behavioral Statistics* **29**, 461–488.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P. and Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society, Series B, Statistical Methodology* **64**, 583–639. MR1979380 https://doi.org/10.1111/1467-9868.00353

Spiegelhalter, D. J., Best, N. G., Carlin, B. P. and Van Der Linde, A. (2014). The deviance information criterion: 12 years on. *Journal of the Royal Statistical Society, Series B, Statistical Methodology* **76**, 485–493. MR3210727 https://doi.org/10.1111/rssb.12062

Stone, C. A. and Zhu, X. (2015). *Bayesian Analysis of Item Response Theory Models Using SAS*. Cary: SAS Institute. MR3968058

Swaminathan, H. and Gifford, J. A. (1985). Bayesian estimation in the two-parameter logistic model. *Psychometrika* **50**, 349–364. MR0871012 https://doi.org/10.1007/BF02295598

Tuerlinckx, F. and De Boeck, P. (2001). The effect of ignoring item interactions on the estimated discrimination parameters in item response theory. *Psychological Methods* **6**, 181.

Wainer, H., Bradlow, E. T. and Wang, X. (2007). *Testlet Response Theory and Its Applications*. Cambridge: Cambridge University Press.

Wainer, H. and Wang, X. (2000). Using a new statistical model for testlets to score TOEFL. *Journal of Educational Measurement* **37**, 203–220.

Wang, W.-C. and Wilson, M. (2005). The rasch testlet model. *Applied Psychological Measurement* **29**, 126–149.

Wang, X., Baldwin, S., Wainer, H., Bradlow, E. T., Reeve, B. B., Smith, A. W., Bellizzi, K. M. and Baumgartner, K. B. (2010). Using testlet response theory to analyze data from a survey of attitude change among breast cancer survivors. *Statistics in Medicine* **29**, 2028–2044. MR2758445 https://doi.org/10.1002/sim.3945

Wang, X., Bradlow, E. T. and Wainer, H. (2002). A general Bayesian model for testlets: Theory and applications. *Applied Psychological Measurement* **26**, 109–128. MR1881784 https://doi.org/10.1177/014662160202600107

Brazilian Journal of Probability and Statistics
2023, Vol. 37, No. 1, 26–54
https://doi.org/10.1214/22-BJPS559

Multivariate Birnbaum–Saunders distribution based on a skewed distribution and associated EM-estimation

Filidor Vilca[1,a], Camila Borelli Zeller[2,b] and Narayanaswamy Balakrishnan[3,c]

[1]*Departamento de Estatística, Universidade Estadual de Campinas, São Paulo, Brazil,* [a]*fily@unicamp.br*
[2]*Departamento de Estatística, Universidade Federal de Juiz de Fora, Juiz de Fora, Brazil,* [b]*camila.zeller@ufjf.edu.br*
[3]*Department of Mathematics and Statistics, McMaster University, Hamilton, Ontario, Canada,* [c]*bala@mcmaster.ca*

Abstract. We develop here a multivariate generalization of Birnbaum–Saunders (BS) distribution based on the multivariate skew-normal distribution. Some distributional characteristics and properties are presented, as well as a simple and efficient EM algorithm for the iterative computation of the maximum likelihood (ML) estimates of model parameters, through the hierarchical representation of the proposed model. The standard errors of the maximum likelihood estimates are calculated from the observed Fisher information matrix. Moreover, by using the tools, we present a log-linear regression model, where the the ML estimates are once again obtained using an EM algorithm. Finally, simulation studies and two applications to real data sets are presented for illustrating the model and the inferential results developed here.

References

Arellano-Valle, R. B., Bolfarine, H. and Lachos, V. H. (2005). Skew-normal linear mixed models. *Journal of Data Science* **3**, 415–438.

Arnold, B. C., Castilho, E. and Sarabia, J. M. (2002). Conditionally specified multivariate skewed distributions. *Sankhya Series A* **64**, 206–226. MR1981754

Arnold, B. C., Castillo, E. and Sarabia, J. M. (2001). Conditionally specified distributions: An introduction. *Statistical Science* **16**, 249–274. MR1874154 https://doi.org/10.1214/ss/1009213728

Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics* **12**, 171–178. MR0808153

Azzalini, A. and Dalla Valle, A. (1996). The multivariate skew-normal distribution. *Biometrika* **83**, 715–726. MR1440039 https://doi.org/10.1093/biomet/83.4.715

Balakrishnan, N. and Kundu, D. (2019). Birnbaum–Saunders distribution: A review of models, analysis, and applications. *Applied Stochastic Models in Business and Industry* **35**, 4–132. (with discussions). MR3915800 https://doi.org/10.1002/asmb.2348

Balakrishnan, N. and Lai, C. D. (2009). *Continuous Bivariate Distributions*, 2nd ed. New York: Springer. MR2840643 https://doi.org/10.1007/b101765

Balakrishnan, N., Leiva, V., Sanhueza, A. and Vilca, F. (2009). Estimation in the Birnbaum–Saunders distribution based on scale-mixture of normals and the EM-algorithm. *Statistics and Operations Research Transactions* **33**, 171–192. MR2643505

Barros, M., Paula, G. A. and Leiva, V. (2008). A new class of survival regression models with heavy-tailed errors: Robustness and diagnostics. *Lifetime Data Analysis* **14**, 316–332. MR2516848 https://doi.org/10.1007/s10985-008-9085-1

Birnbaum, Z. W. and Saunders, S. C. (1969a). A new family of life distributions. *Journal of Applied Probability* **6**, 637–652. MR0253493 https://doi.org/10.2307/3212003

Birnbaum, Z. W. and Saunders, S. C. (1969b). Estimation for a family of life distributions with applications to fatigue. *Journal of Applied Probability* **6**, 328–347. MR0251807 https://doi.org/10.2307/3212004

Castillo, N. O., Gómez, H. W. and Bolfarine, H. (2011). Epsilon Birnbaum–Saunders distribution family: Properties and inference. *Statistical Papers* **52**, 871–883. MR2846690 https://doi.org/10.1007/s00362-009-0293-x

Key words and phrases. Birnbaum–Saunders distribution, skew-normal distribution, marginal and Conditional distributions, EM algorithm.

Desmond, A. (1985). Stochastic models of failure in random environments. *Canadian Journal of Statistics* **13**, 171–183. MR0818323 https://doi.org/10.2307/3315148

Guiraud, P., Leiva, V. and Fierro, R. (2009). A non-central version of the Birnbaum–Saunders distribution for reliability analysis. *IEEE Transactions on Reliability* **58**, 152–160.

Gupta, A. K., González-Farías, G. and Domínguez-Molina, J. A. (2004). A multivariate skew normal distribution. *Journal of Multivariate Analysis* **89**, 181–190. MR2041215 https://doi.org/10.1016/S0047-259X(03)00131-3

Jamalizadeh, A. and Kundu, D. (2015). A multivariate Birnbaum–Saunders distribution based on multivariate skew normal distribution. *Journal of Japan Statistical Society* **45**, 1–20. MR3444401 https://doi.org/10.14490/jjss.45.1

Johnson, R. A. and Wichern, D. W. (1999). *Applied Multivariate Statistical Analysis*. New Jersey: Prentice-Hall. MR0653327

Kotz, S., Balakrishnan, N. and Johnson, N. L. (2000). *Continuous Multivariate Distributions—Vol. 1*, 2nd ed. New York: Wiley. MR1788152 https://doi.org/10.1002/0471722065

Kundu, D., Balakrishnan, N. and Jamalizadeh, A. (2010). Bivariate Birnbaum–Saunders distribution and associated inference. *Journal of Multivariate Analysis* **101**, 113–125. MR2557622 https://doi.org/10.1016/j.jmva.2009.05.005

Kundu, D., Balakrishnan, N. and Jamalizadeh, A. (2013). Generalized multivariate Birnbaum–Saunders distributions and related inferential issues. *Journal of Multivariate Analysis* **116**, 230–244. MR3049902 https://doi.org/10.1016/j.jmva.2012.10.017

Lange, K. L. and Sinsheimer, J. S. (1993). Normal/independent distributions and their applications in robust regression. *Journal of Computational and Graphical Statistics* **2**, 175–198. MR1272391 https://doi.org/10.2307/1390698

Leiva, V. (2016). *The Birnbaum–Saunders Distribution*. New York, NY, USA: Academic Press. MR3430824 https://doi.org/10.1016/B978-0-12-803769-0.00001-7

Leiva, V., Riquelme, M., Balakrishnan, N. and Sanhueza, A. (2008). Lifetime analysis based on the generalized Birnbaum–Saunders distribution. *Computational Statistics & Data Analysis* **52**, 2079–2097. MR2418490 https://doi.org/10.1016/j.csda.2007.07.003

Lemonte, A. (2013). A new extended Birnbaum–Saunders regression model for lifetime modeling. *Computational Statistics & Data Analysis* **64**, 34–50. MR3061888 https://doi.org/10.1016/j.csda.2013.02.025

Maehara, R., Bolfarine, H., Vilca, F. and Balakrishnan, N. (2021). A robust Birnbaum–Saunders regression model based on asymmetric heavy-tailed distributions. *Metrika* **84**, 1049–1080. MR4305448 https://doi.org/10.1007/s00184-021-00815-4

Mann, N. R., Schafer, R. E. and Singpurwalla, N. D. (1974). *Methods for Statistical Analysis of Reliability and Life Data*. New York: Wiley. MR0365976

Marchant, C., Leiva, V., Cysneiros, F. and Vivanco, J., (2016). Diagnostics in multivariate generalized Birnbaum-Saunders regression models. *Journal of Applied Statistics* **43**, 2829–2849. MR3546117 https://doi.org/10.1080/02664763.2016.1148671

Marshall, A. W. and Olkin, I. (2007). *Life Distributions*. New York: Springer. MR2344835

Martínez-Flórez, G., Barranco-Chamorro, I., Bolfarine, H. and Gómez, H. W. (2019). Flexible Birnbaum–Saunders distribution. *Symmetry* **11**, 1305.

Martínez-Flórez, G., Bolfarine, H. and Gómez, H. W. (2014). An alpha-power extension for the Birnbaum–Saunders distribution. *Statistics* **48**, 896–912. MR3234069 https://doi.org/10.1080/02331888.2013.846910

Meng, X. L. and Rubin, D. B. (1993). Maximum likelihood estimation via the ECM algorithm: A general framework. *Biometrika* **80**, 267–278. MR1243503 https://doi.org/10.1093/biomet/80.2.267

Montenegro, L., Lachos, V. and Bolfarine, H. (2010). Inference for a skew extension of the Grubbs model. *Statistical Papers* **51**, 701–715. MR2679342 https://doi.org/10.1007/s00362-008-0157-9

Navarro, J. and Sarabia, J. M. (2013). Reliability properties of bivariate conditional proportional hazard rate models. *Journal of Multivariate Analysis* **113**, 116–127. MR2984360 https://doi.org/10.1016/j.jmva.2011.03.009

Rieck, J. R. and Nedelman, J. R. (1991). A log-linear model for the Birnbaum–Saunders distribution. *Technometrics* **33**, 51–60.

Romeiro, R. G., Vilca, F. and Balakrishnan, N. (2018). A robust multivariate Birnbaum–Saunders distribution: EM estimation. *Statistics* **52**, 321–344. MR3772184 https://doi.org/10.1080/02331888.2017.1398258

Romeiro, R. G., Vilca, F., Balakrishnan, N. and Zeller, C. B. (2020). A robust multivariate Birnbaum–Saunders regression model. *Statistics* **54**, 1094–1123. MR4178900 https://doi.org/10.1080/02331888.2020.1824231

Santana, L., Vilca, L. F. and Leiva, V. (2011). Influence analysis in skew-Birnbaum–Saunders regression models and applications. *Journal of Applied Statistics* **38**, 1633–1649. MR2819378 https://doi.org/10.1080/02664763.2010.515679

Vilca, F., Balakrishnan, N. and Zeller, C. B. (2014). A robust extension of the bivariate Birnbaum–Saunders distribution and associated inference. *Journal of Multivariate Analysis* **124**, 418–435. MR3147335 https://doi.org/10.1016/j.jmva.2013.11.005

Vilca, L. F. and Leiva, V. (2006). A new fatigue life model based on the family of skew-elliptical distributions. *Communications in Statistics—Theory and Methods* **35**, 229–244. MR2274046 https://doi.org/10.1080/03610920500440065

Vilca, L. F., Romeiro, R. and Balakrishnan, N. (2016). A bivariate Birnbaum-Saunders regression model. *Computational Statistics & Data Analysis* **97**, 169–183. MR3447043 https://doi.org/10.1016/j.csda.2015.12.003

Vilca, L. F., Santana, L., Leiva, V. and Balakrishnan, N. (2011). Estimation of extreme percentiles in Birnbaum–Saunders distributions. *Computational Statistics & Data Analysis* **55**, 1665–1678. MR2748670 https://doi.org/10.1016/j.csda.2010.10.023

Brazilian Journal of Probability and Statistics
2023, Vol. 37, No. 1, 55–72
https://doi.org/10.1214/22-BJPS560

A new class of bivariate Sushila distributions in presence of right-censored and cure fraction

Ricardo Puziol de Oliveira[a], Marcos Vinicius de Oliveira Peres, Jorge Alberto Achcar and Edson Z. Martinez

Ribeirão Preto Medical School, University of São Paulo, Ribeirão Preto, SP, Brazil,
[a]*rpuziol.oliveira@gmail.com*

Abstract. The present study introduces a new bivariate distribution based on the Sushila distribution to model bivariate lifetime data in presence of a cure fraction, right- censored data and covariates. The new bivariate probability distribution was obtained using a methodology used in the reliability theory based on fatal shocks, usually used to build new bivariate models. Additionally, the cure rate was introduced in the model based on a generalization of standard mixture models extensively used for the univariate lifetime case. The inferences of interest for the model parameters are obtained under a Bayesian approach using MCMC (Markov Chain Monte Carlo) simulation methods to generate samples of the joint posterior distribution for all parameters of the model. A simulation study was developed to study the inferential properties of the new methodology.The proposed methodology also was applied to analyze a set of real medical data obtained from a retrospective cohort study that aimed to assess specific clinical conditions that affect the lives of patients with diabetic retinopathy. For the discrimination of the proposed model with other usual models used in the analysis of bivariate survival data, some Bayesian techniques of model discrimination were used and the model validation was verified from usual Cox-Snell residuals, which allowed us to identify the adequacy of the proposed bivariate cure rate model.

References

Achcar, J. A. and Bolfarine, H. (1986a). Use of accurate approximations for posterior densities in regression models with censored data. *Revista de la Sociedad Chilena de Estadística* **3**, 84–104.

Achcar, J. A. and Bolfarine, H. (1986b). The log-linear model with a generalized gamma distribution for the error: A Bayesian approach. *Statistics & Probability Letters* **4**, 325–332. MR0858326 https://doi.org/10.1016/0167-7152(86)90053-2

Achcar, J. A. and Bolfarine, H. (1988). Predictive densities in survival analysis with a generalized gamma regression model. *Brazilian Journal of Probability and Statistics*, 23–31. MR0971218

Achcar, J. A. and Bolfarine, H. (1989). Constant hazard against a change-point alternative: A Bayesian approach with censored data. *Communications in Statistics Theory and Methods* **18**, 3801–3819. MR1040677 https://doi.org/10.1080/03610928908830124

Achcar, J. A., Brookmeyer, R. and Hunter, W. G. (1985). An application of Bayesian analysis to medical follow-up data. *Statistics in Medicine* **4**, 509–520.

Achcar, J. A., Coelho-Barros, E. A. and Mazucheli, J. (2012). Cure fraction models using mixture and non-mixture models. *Tatra Mountains Mathematical Publications* **51**, 1–9. MR3014926 https://doi.org/10.2478/v10127-012-0001-4

Achcar, J. A. and Leandro, R. A. (1998). Use of Markov Chain Monte Carlo methods in a Bayesian analysis of the Block and Basu bivariate exponential distribution. *Annals of the Institute of Statistical Mathematics* **50**, 403–416. MR1664583 https://doi.org/10.1023/A:1003582409664

Balakrishnan, N. and Ristić, M. M. (2016). Multivariate families of gamma-generated distributions with finite or infinite support above or below the diagonal. *Journal of Multivariate Analysis* **143**, 194–207. MR3431428 https://doi.org/10.1016/j.jmva.2015.09.012

Key words and phrases. Bayesian inference, correlation coefficient, Marshall and Olkin approach, Sushila distribution.

Block, H. W. and Basu, A. P. (1974). A continuous, bivariate exponential extension. *Journal of the American Statistical Association* **69**, 1031–1037. MR0433713

Brooks, S. P. (2002). Discussion on the paper by Spiegelhalter, Best, Carlin and van der Linde. WILEY-BLACKWELL 111 RIVER ST, HOBOKEN 07030-5774, NJ USA.

Cancho, V. G. and Bolfarine, H. (2001). Modeling the presence of immunes by using the exponentiated-Weibull model. *Journal of Applied Statistics* **28**, 659–671. MR1858829 https://doi.org/10.1080/02664760120059200

Cancho, V. G., Bolfarine, H. and Achcar, J. A. (1999). A Bayesian analysis for the exponentiated-Weibull distribution. *Journal of Applied Statistical Science* **8**, 227–242. MR1706227

Cancho, V. G., Dey, D. K. and Louzada, F. (2016). Unified multivariate survival model with a surviving fraction: An application to a Brazilian customer churn data. *Journal of Applied Statistics* **43**, 572–584. MR3441556 https://doi.org/10.1080/02664763.2015.1071341

Carlin, B. P. and Louis, T. A. (2000). Empirical Bayes: Past, present and future. *Journal of the American Statistical Association* **95**, 1286–1289. MR1825277 https://doi.org/10.2307/2669771

Chen, M.-H., Shao, Q.-M. and Ibrahim, J. G. (2012). *Monte Carlo Methods in Bayesian Computation*. Berlin: Springer. MR1742311 https://doi.org/10.1007/978-1-4612-1276-8

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

Cox, D. R. (1972). Regression models and life tables. *Journal of the Royal Statistical Society, Series B* **34**, 187–220. MR0341758

Cox, D. R. and Snell, E. J. (1968). A general definition of residuals. *Journal of the Royal Statistical Society, Series B, Methodological* **30**, 248–275. MR0237052

Crowley, J. and Hu, M. (1977). Covariance analysis of heart transplant survival data. *Journal of the American Statistical Association* **72**, 27–36.

De Angelis, R., Capocaccia, R., Hakulinen, T., Soderman, B. and Verdecchia, A. (1999). Mixture models for cancer survival analysis: Application to population-based data with covariates. *Statistics in Medicine* **18**, 441–454.

Dey, D. K., Chen, M.-H. and Chang, H. (1997). Bayesian approach for nonlinear random effects models. *Biometrics* **53**, 1239–1252.

Farewell, V. T. (1982). The use of mixture models for the analysis of survival data with long-term survivors. *Biometrics* **38**, 1041–1046.

Geisser, S. and Eddy, W. F. (1979). A predictive approach to model selection. *Journal of the American Statistical Association* **74**, 153–160.

Gelfand, A. E., Dey, D. K. and Chang, H. (1992). Model determination using predictive distributions with implementation via sampling-based methods. Technical report, DTIC Document. MR1380275

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

Ghitany, M. E., Atieh, B. and Nadarajah, S. (2008). Lindley distribution and its application. *Mathematics and Computers in Simulation* **78**, 493–506. MR2424558 https://doi.org/10.1016/j.matcom.2007.06.007

Huster, W. J., Brookmeyer, R. and Self, S. G. (1989). Modelling paired survival data with covariates. *Biometrics* **45**, 145–156. MR0999443 https://doi.org/10.2307/2532041

Johnson, N. L. and Kotz, S. (1975). A vector multivariate hazard rate. *Journal of Multivariate Analysis* **5**, 53–66. MR0365901 https://doi.org/10.1016/0047-259X(75)90055-X

Kaplan, E. L. and Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association* **53**, 457–481. MR0093867

Lambert, P. C. (2007). Modeling of the cure fraction in survival studies. *Stata Journal* **7**, 351.

Lambert, P. C., Thompson, J. R., Weston, C. L. and Dickman, P. W. (2006). Estimating and modeling the cure fraction in population-based cancer survival analysis. *Biostatistics* **8**, 576–594.

Lu, W. (2010). Efficient estimation for an accelerated failure time model with a cure fraction. *Statistica Sinica* **20**, 661. MR2682635

Marshall, A. W. and Olkin, I. (1967a). A multivariate exponential distribution. *Journal of the American Statistical Association* **62**, 30–44.

Marshall, A. W. and Olkin, I. (1967b). A generalized bivariate exponential distribution. *Journal of Applied Probability* **4**, 291–302.

Oliveira, R. P., Achcar, J. A., Mazucheli, J. and Bertoli, W. (2021). A new class of bivariate Lindley distributions based on stress and shock models and some of their reliability properties. *Reliability Engineering & Systems Safety* **211**, 107528.

Oliveira, R. P. d. (2019). Multivariate lifetime models to evaluate long-term survivors in medical studies. PhD thesis, Universidade de São Paulo.

Othus, M., Barlogie, B., LeBlanc, M. L. and Crowley, J. J. (2012). Cure models as a useful statistical tool for analyzing survival. *Clinical Cancer Research* **18**, 3731–3736.

Price, D. L. and Manatunga, A. K. (2001). Modelling survival data with a cured fraction using frailty models. *Statistics in Medicine* **20**, 1515–1527.

R Core Team (2015) R: A language and environment for statistical computing R Foundation for Statistical Computing. Vienna, Austria.

Ristić, M. M., Popović, B. V., Zografos, K. and Balakrishnan, N. (2018). Discrimination among bivariate beta-generated distributions. *Statistics* **52**, 303–320. MR3772183 https://doi.org/10.1080/02331888.2017.1397156

Shanker, R., Sharma, S., Shanker, U. and Shanker, R. (2013). Sushila distribution and its application to waiting times data. *International Journal of Business Management* **3**, 1–11.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P. and Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society, Series B, Statistical Methodology* **64**, 583–639. MR1979380 https://doi.org/10.1111/1467-9868.00353

Su, Y.-S. and Yajima, M. (2012). R2jags: A Package for Running jags from R. *R package version 0.03-08, URL http://CRAN.R-project.org/package=R2jags*.

The National Eye Institute (2018). *Facts About Diabetic Eye Disease*. Available at: https://nei.nih.gov/health/diabetic/retinopathy.

Tsodikov, A. D., Ibrahim, J. G. and Yakovlev, A. Y. (2003). Estimating cure rates from survival data: An alternative to two-component mixture models. *Journal of the American Statistical Association* **98**, 1063–1078. MR2055496 https://doi.org/10.1198/01622145030000001007

Vahidpour, M. (2016). Cure Rate Models. PhD thesis, École Polytechnique de Montréal.

Vaidyanathan, V. S., Sharon Varghese, A., et al (2016). Morgenstern type bivariate Lindley distribution. *Statistics, Optimization & Information Computing* **4**, 132–146. MR3519336 https://doi.org/10.19139/soic.v4i2.183

Wienke, A., Lichtenstein, P. and Yashin, A. I. (2003). A bivariate frailty model with a cure fraction for modeling familial correlations in diseases. *Biometrics* **59**, 1178–1183. MR2019823 https://doi.org/10.1111/j.0006-341X.2003.00135.x

Wienke, A., Locatelli, I. and Yashin, A. I. (2006). The modelling of a cure fraction in bivariate time-to-event data. *Austrian Journal of Statistics* **35**, 67–76.

Yin, G. and Ibrahim, J. G. (2005). Cure rate models: A unified approach. *Canadian Journal of Statistics* **33**, 559–570.

Yu, B., Tiwari, R. C., Cronin, K. A. and Feuer, E. J. (2004). Cure fraction estimation from the mixture cure models for grouped survival data. *Statistics in Medicine* **23**, 1733–1747.

Brazilian Journal of Probability and Statistics
2023, Vol. 37, No. 1, 73–100
https://doi.org/10.1214/22-BJPS561

Expansions for posterior distributions

Christopher Withers[1,a] and Saralees Nadarajah[2,b]

[1]*Callaghan Innovation, Lower Hutt, New Zealand,* [a]*kit.withers@gmail.com*
[2]*Howard University, Washington DC 20059, USA,* [b]*saraleesan.nadarajah@howard.edu*

Abstract. Suppose that X_n is a sample of size n with log likelihood $nl(\theta)$, where θ is an unknown parameter in R^p having a prior distribution $\xi(\theta)$. We need not assume that the sample values are independent or even stationary. Let $\widehat{\theta}$ be the maximum likelihood estimate (MLE). We show that $\theta|X_n$ is asymptotically normal with mean $\widehat{\theta}$ and covariance $-n^{-1}l_{\bullet,\bullet}(\widehat{\theta})^{-1}$, where $l_{\bullet,\bullet}(\theta) = \partial^2 l(\theta)/\partial\theta\partial\theta'$. In contrast, $\widehat{\theta}|\theta$ is asymptotically normal with mean θ and covariance $n^{-1}[I(\theta)]^{-1}$, where $I(\theta) = -E[l_{\bullet,\bullet}(\widehat{\theta})|\theta]$ is Fisher's information. So, frequentist inference conditional on θ *cannot* be used to approximate Bayesian inference, except for exponential families. However, under mild conditions $-l_{\bullet,\bullet}(\widehat{\theta})|\theta \to I(\theta)$ in probability. So, Bayesian inference (that is, conditional on X_n) *can* be used to approximate frequentist inference.

For $t(\theta)$ any smooth function, we obtain posterior cumulant expansions, posterior Edgeworth–Cornish–Fisher (ECF) expansions and posterior tilted Edgeworth expansions for $\mathcal{L}t(\theta)|X_n$, as well as confidence regions for $t(\theta)|X_n$ of high accuracy. We also give expansions for the Bayes estimate (estimator) of $t(\theta)$ about $t(\widehat{\theta})$, and for the maximum *a posteriori* estimate about $\widehat{\theta}$, as well as their relative efficiencies with respect to squared error loss.

References

Comtet, L. (1974). *Advanced Combinatorics*. Dordrecht: Reidel. MR0460128

Daniels, H. E. (1954). Saddlepoint approximations in statistics. *The Annals of Mathematical Statistics* **25**, 631–650. MR0066602 https://doi.org/10.1214/aoms/1177728652

Dasgupta, S., Khare, K. and Ghosh, M. (2014). Asymptotic expansion of the posterior density in high dimensional generalized linear models. *Journal of Multivariate Analysis* **131**, 126–148. MR3252640 https://doi.org/10.1016/j.jmva.2014.06.013

Johnson, R. A. (1970). Asymptotic expansions associated with posterior distributions. *The Annals of Mathematical Statistics* **41**, 851–864. MR0263198 https://doi.org/10.1214/aoms/1177696963

Kolassa, J. E. and Kuffner, T. A. (2020). On the validity of the formal Edgeworth expansion for posterior densities. *The Annals of Statistics* **48**, 1940–1958. MR4134781 https://doi.org/10.1214/19-AOS1871

Kotz, S., Balakrishnan, N. and Johnson, N. L. (2000). *Continuous Multivariate Distributions, 1*. New York: Wiley. MR1788152 https://doi.org/10.1002/0471722065

Kotz, S. and Nadarajah, S. (2004). *Multivariate t Distributions and Their Applications*. Cambridge: Cambridge University Press. MR2038227 https://doi.org/10.1017/CBO9780511550683

Miyata, Y. (2008). Higher order expansions for posterior distributions using posterior modes. *Journal of Japanese Statistical Society* **38**, 415–429. MR2510947 https://doi.org/10.14490/jjss.38.415

O'Hagan, A. (1994). *Kendall's Advanced Theory of Statistics, 2B, Bayesian Inference*. New York: Wiley. MR1285356

Shenton, L. R. and Bowman, K. O. (1977). *Maximum Likelihood Estimation in Small Samples. Griffin's Statistical Monograph No. 38*. London.

Sweeting, T. J. and Adekola, A. O. (1987). Asymptotic posterior normality for stochastic processes revisited. *Journal of the Royal Statistical Society, Series B* **49**, 215–222. MR0905193

Walker, A. M. (1969). On the asymptotic behaviour of posterior distributions. *Journal of the Royal Statistical Society, Series B* **31**, 80–88. MR0269000

Weiss, L. and Wolfowitz, J. (1974). *Maximum Probability Estimators and Related Topics*. Berlin: Springer. MR0359152

Key words and phrases. Cumulant expansions, Edgeworth–Cornish–Fisher expansions, Maximum likelihood estimates.

Weng, R. C. (2010). A Bayesian Edgeworth expansion by Stein's identity. *Bayesian Analysis* **5**, 741–764. MR2740155 https://doi.org/10.1214/10-BA526

Weng, R. C. and Hsu, C.-H. (2013). A study of expansions of posterior distributions. *Communications in Statistics Theory and Methods* **42**, 346–364. MR3004640 https://doi.org/10.1080/03610926.2011.579701

Weng, R. C. and Woodroofe, M. (2000). Integrable expansions for posterior distributions for multiparameter exponential families with applications to sequential confidence levels. *Statistica Sinica* **10**, 693–713. MR1787775

Withers, C. S. (1982). The distribution and quantiles of a function of parameter estimates. *Annals of the Institute of Statistical Mathematics* **34**, 55–68. MR0650324 https://doi.org/10.1007/BF02481007

Withers, C. S. (1983). Expansions for the distribution and quantiles of a regular functional of the empirical distribution with applications to nonparametric confidence intervals. *The Annals of Statistics* **11**, 577–587. MR0696069

Withers, C. S. (1987). Bias reduction by Taylor series. *Communications in Statistics Theory and Methods* **16**, 2369–2383. MR0915469 https://doi.org/10.1080/03610928708829512

Withers, C. S. (1989). Accurate confidence intervals when nuisance parameters are present. *Communications in Statistics Theory and Methods* **18**, 4229–4259. MR1058938 https://doi.org/10.1080/03610928908830152

Withers, C. S. and Nadarajah, S. (2008). Analytic bias reduction for k-sample functionals. *Sankhyā, A* **70**, 186–222. MR2551813

Withers, C. S. and Nadarajah, S. (2010). Tilted Edgeworth expansions for asymptotically normal vectors. *Annals of the Institute of Statistical Mathematics* **62**, 1113–1142. MR2729156 https://doi.org/10.1007/s10463-008-0206-0

Withers, C. S. and Nadarajah, S. (2013). Saddlepoint expansions in terms of Bell polynomials. *Integral Transforms and Special Functions* **24**, 410–423. MR3055529 https://doi.org/10.1080/10652469.2012.702345

Withers, C. S. and Nadarajah, S. (2022). Expansions for posterior distributions. Technical Report, Department of Mathematics, University of Manchester, UK.

Wu, Y. and Ghosh, M. (2017). Asymptotic expansion of the posterior based on pairwise likelihood. *Sankhyā, A* **79**, 39–75. MR3634818 https://doi.org/10.1007/s13171-016-0094-y

Zaikin, A. A. (2017). On asymptotic expansion of posterior distribution. *Lobachevskii Journal of Mathematics* **37**, 515–525. MR3528029 https://doi.org/10.1134/S1995080216040181

Brazilian Journal of Probability and Statistics
2023, Vol. 37, No. 1, 101–123
https://doi.org/10.1214/23-BJPS562
This research was funded, in whole or in part, by [EPSRC, EP/L015684/1].

Componentwise equivariant estimation of order restricted location and scale parameters in bivariate models: A unified study

Naresh Garg[a] **and Neeraj Misra**[b]

Department of Mathematics and Statistics, Indian Institute of Technology Kanpur, Kanpur-208016, Uttar Pradesh, India, [a]*nareshng@iitk.ac.in,* [b]*neeraj@iitk.ac.in*

Abstract. The problem of estimating location (scale) parameters θ_1 and θ_2 of two distributions when the ordering between them is known apriori (say, $\theta_1 \leq \theta_2$) has been extensively studied in the literature. Many of these studies are centered around deriving estimators that dominate the best location (scale) equivariant estimators, for the unrestricted case, by exploiting the prior information that $\theta_1 \leq \theta_2$. Several of these studies consider specific distributions such that the associated random variables are statistically independent. In this paper, we consider a general bivariate model and a general loss function, and unify various results proved in the literature. We also consider applications of these results to a bivariate normal and a Cheriyan and Ramabhadran's bivariate gamma model. A simulation study is also considered to compare the risk performances of various estimators under bivariate normal and Cheriyan and Ramabhadran's bivariate gamma models.

References

Barlow, R. E., Bartholomew, D. J., Bremner, J. M. and Brunk, H. D. (1972). *Statistical Inference Under Order Restrictions. The Theory and Application of Isotonic Regression*. London: John Wiley & Sons. MR0326887

Berger, J. (2013). *Statistical Decision Theory: Foundations, Concepts, and Methods*. New York: Springer. MR0580664

Blumenthal, S. and Cohen, A. (1968). Estimation of two ordered translation parameters. *The Annals of Mathematical Statistics* **39**, 517–530. MR0223007 https://doi.org/10.1214/aoms/1177698414

Bobotas, P. (2019a). Estimation of the smallest scale parameter of two-parameter exponential distributions. *Communications in Statistics Theory and Methods* **48**, 2748–2765. MR3963133 https://doi.org/10.1080/03610926.2018.1472792

Bobotas, P. (2019b). Improved estimation of the smallest scale parameter of gamma distributions. *Journal of the Korean Statistical Society* **48**, 97–105. MR3926974 https://doi.org/10.1016/j.jkss.2018.08.007

Brewster, J. F. and Zidek, J. V. (1974). Improving on equivariant estimators. *The Annals of Statistics* **2**, 21–38. MR0381098

Chang, Y.-T. and Shinozaki, N. (2015). Estimation of two ordered normal means under modified Pitman nearness criterion. *Annals of the Institute of Statistical Mathematics* **67**, 863–883. MR3390170 https://doi.org/10.1007/s10463-014-0479-4

Cohen, A. and Sackrowitz, H. B. (1970). Estimation of the last mean of a monotone sequence. *The Annals of Mathematical Statistics* **41**, 2021–2034. MR0270483 https://doi.org/10.1214/aoms/1177696702

Garg, N. and Misra, N. (2022). Supplement to "Componentwise equivariant estimation of order restricted location and scale parameters in bivariate models: A unified study." https://doi.org/10.1214/23-BJPS562SUPP

Gupta, R. D. and Singh, H. (1992). Pitman nearness comparisons of estimates of two ordered normal means. *Australian Journal of Statistics* **34**, 407–414. MR1210354 https://doi.org/10.1111/j.1467-842x.1992.tb01056.x

Hwang, J. T. G. and Peddada, S. D. (1994). Confidence interval estimation subject to order restrictions. *The Annals of Statistics* **22**, 67–93. MR1272076 https://doi.org/10.1214/aos/1176325358

Iliopoulos, G. (2000). A note on decision theoretic estimation of ordered parameters. *Statistics & Probability Letters* **50**, 33–38. MR1804623 https://doi.org/10.1016/S0167-7152(00)00078-X

Kelly, R. E. (1989). Stochastic reduction of loss in estimating normal means by isotonic regression. *The Annals of Statistics* **17**, 937–940. MR0994278 https://doi.org/10.1214/aos/1176347153

Key words and phrases. Best location equivariant estimator (BLEE), Best scale equivariant estimator (BSEE), Brewster–Zidek type estimator, Generalized Bayes estimators, Stein type estimator.

Kotz, S., Balakrishnan, N. and Johnson, N. L. (2000). *Continuous Multivariate Distributions: Models and Applications*, 2nd ed. *Wiley Series in Probability and Statistics: Applied Probability and Statistics* **1**. New York: Wiley-Interscience. MR0418337

Kubokawa, T. (1994). A unified approach to improving equivariant estimators. *The Annals of Statistics* **22**, 290–299. MR1272084 https://doi.org/10.1214/aos/1176325369

Kubokawa, T. and Saleh, A. K. M. E. (1994). Estimation of location and scale parameters under order restrictions. *Journal of Statistical Research* **28**, 41–51. MR1370413

Kumar, S. and Sharma, D. (1988). Simultaneous estimation of ordered parameters. *Communications in Statistics Theory and Methods* **17**, 4315–4336. MR0981031 https://doi.org/10.1080/03610928808829876

Kumar, S. and Sharma, D. (1989). On the Pitman estimator of ordered normal means. *Communications in Statistics Theory and Methods* **18**, 4163–4175. MR1058934 https://doi.org/10.1080/03610928908830148

Kushary, D. and Cohen, A. (1989). Estimating ordered location and scale parameters. *Statistics & Decisions* **7**, 201–213. MR1029476

Lee, C. I. C. (1981). The quadratic loss of isotonic regression under normality. *The Annals of Statistics* **9**, 686–688. MR0615447

Misra, N. and Dhariyal, I. D. (1995). Some inadmissibility results for estimating ordered uniform scale parameters. *Communications in Statistics Theory and Methods* **24**, 675–685. MR1326266 https://doi.org/10.1080/03610929508831516

Misra, N., Dhariyal, I. D. and Kundu, D. (2002). Natural estimators for the larger of two exponential location parameters with a common unknown scale parameter. *Statistics & Decisions* **20**, 67–80. MR1904424

Misra, N., Iyer, S. K. and Singh, H. (2004). The LINEX risk of maximum likelihood estimators of parameters of normal populations having order restricted means. *Sankhya The Indian Journal of Statistics* **66**, 652–677. MR2205815

Misra, N. and Singh, H. (1994). Estimation of ordered location parameters: The exponential distribution. *Statistics* **25**, 239–249. MR1366828 https://doi.org/10.1080/02331889408802448

Oono, Y. and Shinozaki, N. (2006). On a class of improved estimators of variance and estimation under order restriction. *Journal of Statistical Planning and Inference* **136**, 2584–2605. MR2279823 https://doi.org/10.1016/j.jspi.2004.10.023

Pal, N. and Kushary, D. (1992). On order restricted location parameters of two exponential distributions. *Statistics & Decisions* **10**, 133–152. MR1165709

Patra, L. K. and Kumar, S. (2017). Estimating ordered means of a bivariate normal distribution. *American Journal of Mathematical and Management Sciences* **36**, 118–136.

Patra, L. K., Kumar, S. and Petropoulos, C. (2021). Componentwise estimation of ordered scale parameters of two exponential distributions under a general class of loss function. *Statistics* **55**, 595–617. MR4313441 https://doi.org/10.1080/02331888.2021.1943395

Petropoulos, C. (2010). A class of improved estimators for the scale parameter of a mixture model of exponential distribution and unknown location. *Communications in Statistics Theory and Methods* **39**, 3153–3162. MR2755430 https://doi.org/10.1080/03610920903205198

Petropoulos, C. (2017). Estimation of the order restricted scale parameters for two populations from the Lomax distribution. *Metrika* **80**, 483–502. MR3635472 https://doi.org/10.1007/s00184-017-0615-2

Robertson, T., Wright, F. T. and Dykstra, R. L. (1988). *Order Restricted Statistical Inference*. Chichester: John Wiley & Sons. MR0961262

Sackrowitz, H. (1970). Estimation for monotone parameter sequences: The discrete case. *The Annals of Mathematical Statistics* **41**, 609–620. MR0254961 https://doi.org/10.1214/aoms/1177697101

Shaked, M. and Shanthikumar, J. G. (2007). *Stochastic Orders*. New York: Springer. MR2265633 https://doi.org/10.1007/978-0-387-34675-5

Stein, C. (1964). Inadmissibility of the usual estimator for the variance of a normal distribution with unknown mean. *Annals of the Institute of Statistical Mathematics* **16**, 155–160. MR0171344 https://doi.org/10.1007/BF02868569

van Eeden, C. (2006). *Restricted Parameter Space Estimation Problems. Admissibility and Minimaxity Properties*. *Lecture Notes in Statistics* **188**. New York: Springer. MR2265239 https://doi.org/10.1007/978-0-387-48809-7

Vijayasree, G., Misra, N. and Singh, H. (1995). Componentwise estimation of ordered parameters of k (≥ 2) exponential populations. *Annals of the Institute of Statistical Mathematics* **47**, 287–307. MR1345425 https://doi.org/10.1007/BF00773464

Brazilian Journal of Probability and Statistics
2023, Vol. 37, No. 1, 124–139
https://doi.org/10.1214/23-BJPS563

A new distance-based distribution: Detecting concentration in directional data

Saul A. Souza, Getúlio A. Amaral and Abraão D. C. Nascimento[a]

Department of Statistics, Federal University of Pernambuco, Recife, Brazil, [a]*abraao@de.ufpe.br*

Abstract. In this article, we propose a simple method to study high concentrations in spherical data, in particular in the directional perspective. To this end, we define a distance-based distribution in the interval (0, 1), called the T-statistic distance (*DT*) law, to describe the scattering of points on the unit sphere. Our model is derived from the von Mises–Fisher (vMF) distribution, which is one of the most known directional laws. We show that if the data are vMF distributed, their concentration can be modeled by our distribution. Some of its properties are derived and discussed: Moment generating function, kurtosis and skewness. Likelihood-based inference procedures are provided for both points and hypotheses concerning the *DT* concentration. Further, we propose a new test statistic in terms of the *DT* distribution to deal with scattering within spherical data and derive its exact density. Numerical studies show that maximum likelihood estimates behave asymptotically well even for small sample sizes and that the likelihood ratio test for the *DT* distribution often performs better than Wald tests. We apply our model to paleomagnetic data to illustrate how it is used to analyze spherical data concentration. Results show that the distance-based approach works well to identify high concentration on the unit sphere.

References

Bowman, A. and Azzalini, A. (1997). Applied smoothing techniques for data analysis: The kernel approach with S-plus illustrations. *Computational Statistics* **15**, 301–302. https://doi.org/10.1007/s001800000033

Brockwell, P. J. and Davis, R. A. (1991). *Time Series: Theory and Methods*. *Springer Series in Statistics*. New York: Springer. MR1093459 https://doi.org/10.1007/978-1-4419-0320-4

Chikuse, Y. (2003). Concentrated matrix Langevin distributions. *Journal of Multivariate Analysis* **85**, 375–394. MR1983803 https://doi.org/10.1016/s0047-259x(02)00065-9

Cortez, P. (2021). *Modern Optimization with R*. Berlin: Springer. https://doi.org/10.1007/978-3-319-08263-9

Cuevas-Pacheco, F. and Møller, J. (2018). Log Gaussian Cox processes on the sphere. *Spatial Statistics* **26**, 69–82. MR3846280 https://doi.org/10.48550/arXiv.1803.03051

Diggle, P. J. and Fisher, N. I. (1985). Sphere: A contouring program for spherical data. *Computers & Geosciences* **11**, 725–766. https://doi.org/10.1016/0098-3004(85)90015-9

Figueiredo, A. M. S. (2012). Goodness-of-fit for a concentrated von Mises–Fisher distribution. *Computational Statistics* **17**, 69–82. MR2877811 https://doi.org/10.1007/s00180-011-0238-4

Fisher, N., Lewis, T. and Embleton, B. (1993). *Statistical Analysis of Spherical Data*. Cambridge: Cambridge University Press. MR1247695

Fisher, R. (1953). Dispersion on a sphere. *Proceedings of the Royal Society A Mathematical, Physical and Engineering Sciences* **217**, 295–305. MR0056866 https://doi.org/10.1098/rspa.1953.0064

Gradshteyn, I. and Ryzhik, I. (2000). *Table of Integrals, Series, and Products*. Amsterdam: Elsevier. MR0669666

Heslop, D., Roberts, A. P. and Hawkins, R. (2014). A statistical simulation of magnetic particle alignment in sediments. *Geophysical Journal International* **197**, 828–837. https://doi.org/10.1093/gji/ggu038

Hoaglin, D. C. and Iglewicz, B. (1987). Fine-tuning some resistant rules for outlier labeling. *Journal of the American Statistical Association* **82**, 1147–1149. MR0867622 https://doi.org/10.2307/2289392

Ko, D. (1992). Robust estimation of the concentration parameter of the Von Mises–Fisher distribution. *The Annals of Statistics* **20**, 917–928. MR1165599 https://doi.org/10.1214/aos/1176348663

Lawrence, T., Baddeley, A., Milne, R. K. and Nair, G. (2016). Point pattern analysis on a region of a sphere. *Statistica* **5**, 144–157. MR3492966 https://doi.org/10.1002/sta4.108

Key words and phrases. Concentration criteria, directional data, hypothesis tests, stochastic distance.

Lehmann, E. L. (1999). *Elements of Large-Sample Theory*. Berlin: Springer. ISBN: 0-387-98595-6.

Lewin, L. (1981). *Polylogarithms and Associated Functions*. Amsterdam: Elsevier. MR0618278

Mardia, K. and Jupp, P. (1999). *Directional Statistics*. New York: Wiley. MR1828667

Nascimento, A. D. C., da Silva, R. C. and Amaral, G. J. A. (2020). Distance-based hypothesis tests on the Watson distribution. *Communications in Statistics Simulation and Computation* **49**, 2225–2238. MR4159573 https://doi.org/10.1080/03610918.2018.1515358

Neyman, J. and Pearson, E. S. (1928). On the use and interpretation of certain test criteria for purposes of statistical inference: Part I. *Biometrika* **20A**, 175–240. https://doi.org/10.2307/2331945

R Core Team (2013). *R: A Language and Environment ForStatistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. Available at http://www.R-project.org/.

Radhakrishna Rao, C. (1948). Large sample tests of statistical hypotheses concerning several parameters with applications to problems of estimation. *Mathematical Proceedings of the Cambridge Philosophical Society* **44**, 50–57. MR0024111 https://doi.org/10.1017/s0305004100023987

Rizzo, M. (2007). *Statistical Computing with R*. London: Chapman & Hall/CRC.

Robeson, S. M., Li, A. and Huang, C. (2014). Point-pattern analysis on the sphere. *Spatial Statistics* **10**, 76–86. MR3280091 https://doi.org/10.48550/arXiv.1803.03051

Ross, S. (2010). *A First Course in Probability*. London: Pearson Prentice Hall. MR0380910

Schou, G. (1978). Estimation of the concentration parameter in von Mises–Fisher distributions. *Biometrika* **65**, 369–977. MR0513935 https://doi.org/10.1093/biomet/65.2.369

Wald, A. (1943). Tests of statistical hypotheses concerning several parameters when the number of observations is large. *Transactions of the American Mathematical Society* **54**, 426–482. MR0012401 https://doi.org/10.2307/1990256

Watson, G. (1984). The theory of concentrated Langevin distributions. *Journal of Multivariate Analysis* **14**, 74–82. MR0734099 https://doi.org/10.1016/0047-259x(84)90047-2

Brazilian Journal of Probability and Statistics
2023, Vol. 37, No. 1, 140–154
https://doi.org/10.1214/23-BJPS564

Exact and asymptotic goodness-of-fit tests based on the maximum and its location of the empirical process

Dietmar Ferger[a]

Fakultät Mathematik, Technische Universität Dresden, D-01069 Dresden, Germany,
[a]*dietmar.ferger@tu-dresden.de*

Abstract. The supremum of the standardized empirical process is a promising statistic for testing whether the distribution function F of i.i.d. real random variables is either equal to a given distribution function F_0 (hypothesis) or $F \geq F_0$ (one-sided alternative). Since (*The Annals of Statistics* **7** (1979) 108–115) it is well-known that an affine-linear transformation of the suprema converge in distribution to the Gumbel law as the sample size tends to infinity. This enables the construction of an asymptotic level-α test. However, the rate of convergence is extremely slow. As a consequence the probability of the type I error is much larger than α even for sample sizes beyond 10.000. Now, the standardization consists of the weight-function $1/\sqrt{F_0(x)(1-F_0(x))}$. Substituting the weight-function by a suitable random constant leads to a new test-statistic, for which we can derive the exact distribution (and the limit distribution) under the hypothesis. A comparison via a Monte-Carlo simulation shows that the new test is uniformly better than the Smirnov-test and an appropriately modified test due to (*The Annals of Statistics* **11** (1983) 933–946). Our methodology also works for the two-sided alternative $F \neq F_0$.

References

Anderson, T. W. and Darling, D. A. (1952). Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes. *The Annals of Mathematical Statistics* **23**, 193–212. MR0076213

Billingsley, P. (1999). *Convergence of Probability Measures*, 2nd ed. New York: Wiley. MR0172324

Birnbaum, Z. W. and Pyke, R. (1958). On some distributions related to the statistic D_n^+. *The Annals of Mathematical Statistics* **29**, 179–187. MR1215046

Cahill, N. D., D'Errico, J. R., Narayan, D. A. and Narayan, J. Y. (2002). Fibonacci determinants. *The College Mathematics Journal* **33**, 221–225. MR0012388 https://doi.org/10.1098/rspa.1945.0011

Chang, L. C. (1955). On the ratio of the empirical distribution to the theoretical distribution function. *Acta Mathematica Sinica* **5**, 347–368. (English translation in Selected Transl. Math. Statist. Prob., 4 (1964), 17–38.). MR2174873 https://doi.org/10.1051/ps:2005014

Chibisov, M. D. (1966). Some theorems on the limiting behavior of the empirical distribution function. *Selected Translations in Mathematical Statistics and Probability* **6**, 147–156. MR3510226

Csörgő, M. and Horváth, L. (1993). *Weighted Approximations in Probability and Statistics*. Wiley, Chichester, England. MR3758582 https://doi.org/10.1016/j.spl.2017.10.008

Daniels, H. E. (1945). The statistical theory of the strength of bundles of threads. *Proceedings of the Royal Society of London Series A* **183**, 405–435.

Ferger, D. (2005). On the minimizing point of the incorrectly centered empirical process and its limit distribution in nonregular experiments. *ESAIM Probabilités Et Statistique* **9**, 307–322.

Ferger, D. (2015). Arginf-sets of multivariate cadlag processes and their convergence in hyperspace topologies. *Theory of Stochastic Processes* **20**, 13–41. MR0107924 https://doi.org/10.1007/bf01831720

Ferger, D. (2018). On the supremum of a Brownian bridge standardized by its maximizing point with applications to statistics. *Statistics & Probability Letters* **134**, 63–69. MR0515687

Gutjahr, M. (1988). *Zur Berechnung geschlossener Ausdrücke für die Verteilung von Statistiken, die auf einer empirischen Verteilungsfunktion basieren*. München: Phd-thesis, Ludwig-Maximilians-University. MR0300379 https://doi.org/10.1214/aoms/1177692700

Key words and phrases. Goodness of fit, empirical process, measurability and continuity of the argmax-functional.

Ishii, G. (1959). On the exact probabilities of Rényi's tests. *Ann. Inst. Statist. Math. Tokyo* **11**, 17–24. MR0061792 https://doi.org/10.1007/BF02127580

Jaeschke, D. (1979). The asymptotic distribution of the supremum of the standardized empirical distribution function on subintervals. *The Annals of Statistics* **7**, 108–115. MR0148088

Mason, D. and Schuenemeyer, J. H. (1983). A modified Kolmogorov–Smirnov test sensitive to tail alternatives. *The Annals of Statistics* **11**, 933–946. MR0838963

Noé, M. (1972). The calculations of distributions of two-sided Kolmogorov–Smirnov type statistics. *The Annals of Mathematical Statistics* **43**, 58–64. MR0012387

Rényi, A. (1953). On the theory of order statistics. *Acta Scientiarum Mathematica Hungaricae* **4**, 191–227. MR0277077 https://doi.org/10.1214/aoms/1177693490

Rényi, A. (1962). *Wahrscheinlichkeitsrechnung. Mit einem Anhang über Informationstheorie*. Berlin: VEB Deutscher Verlag der Wissenschaften. MR0217858

Shorack, G. R. and Wellner, J. A. (1986). *Empirical Processes with Applications to Statistics*. New York: Wiley.

Smirnov, N. V. (1944). Approximate laws of distribution of random variables from empirical data. *Uspehi Matematičeskih Nauk* **10**, 179–206. (in Russian).

Steck, G. P. (1971). Rectangle probabilities for uniform order statistics and the probability that the empirical distribution function lies between two distribution functions. *The Annals of Mathematical Statistics* **42**, 1–11.

Takács, L. (1967). *Combinatorial Methods in the Theory of Stochastic Processes*. New York: Wiley.

Brazilian Journal of Probability and Statistics
2023, Vol. 37, No. 1, 155–176
https://doi.org/10.1214/23-BJPS565

A two-step estimation procedure for locally stationary ARMA processes with tempered stable innovations

Shu Wei Chou-Chen[1,2,a] **and Pedro A. Morettin**[1]

[1]*Institute of Mathematics and Statistics, University of São Paulo, Brazil*
[2]*Centro de Investigación en Matemática Pura y Aplicada—Escuela de Estadística, Universidad de Costa Rica, Costa Rica,* [a]*shuwei.chou@ucr.ac.cr*

Abstract. The class of locally stationary processes assumes a time-varying (tv) spectral representation and finite second moment. Different areas have observed phenomena with heavy tail distributions or infinite variance. Using stable distribution as a heavy-tailed innovation is an attractive option. However, its estimation is difficult due to the absence of a closed expression for the density function and the non-existence of second moment. In this paper, we propose the tvARMA model with tempered stable innovations, which have lighter tails than the stable distribution and have finite moments. A two-step method is proposed to estimate this parametric model. In the first step, we use the blocked Whittle estimation to estimate the time-varying structure of the process. In the second step, we recover residuals from the first step and use the maximum likelihood method to estimate the rest of the parameters related to the standardized classical tempered stable (stdCTS) innovations. We perform simulation studies to evaluate the consistency of the maximum likelihood estimation of independent stdCTS samples. Then, we execute simulations to study the two-step estimation of our model. Finally, an empirical application is illustrated.

References

Akashi, F., Liu, Y. and Taniguchi, M. (2015). An empirical likelihood approach for symmetric α-stable processes. *Bernoulli* **21**, 2093–2119. ISSN 13507265, 15739759. Available at http://www.jstor.org/stable/43590524. MR1041388 https://doi.org/10.1214/aos/1176347495

Baeumer, B. and Meerschaert, M. M. (2010). Tempered stable Lévy motion and transient super-diffusion. *Journal of Computational and Applied Mathematics* **233**, 2438–2448. ISSN 0377-0427. MR2577834 https://doi.org/10.1016/j.cam.2009.10.027

Calzolari, G. and Halbleib, R. (2018). Estimating stable latent factor models by indirect inference. *Journal of Econometrics* **205**, 280–301. ISSN 0304-4076. MR0331469 https://doi.org/10.1016/0047-259X(72)90038-3

Calzolari, G., Halbleib, R. and Parrini, A. (2014). Estimating GARCH-type models with symmetric stable innovations: Indirect inference versus maximum likelihood. *Computational Statistics & Data Analysis* **76**, 158–171. ISSN 0167-9473. MR3209433 https://doi.org/10.1016/j.csda.2013.07.028

Chou-Chen, S. W. and Morettin, P. A. (2020). Indirect inference for locally stationary ARMA processes with stable innovations. *Journal of Statistical Computation and Simulation* **90**, 3106–3134. https://doi.org/10.1080/00949655.2020.1797030

Cont, R. and Tankov, P. (2015). *Financial Modelling with Jump Processes, Second Edition*. *Chapman and Hall/CRC Financial Mathematics Series*. London: Taylor & Francis. ISBN 9781420082197. MR2042661

Dahlhaus, R. (1996a). Maximum likelihood estimation and model selection for locally stationary processes. *Journal of Nonparametric Statistics* **6**, 171–191. MR1383050 https://doi.org/10.1080/10485259608832670

Dahlhaus, R. (1996b). On the Kullback-Leibler information divergence of locally stationary processes. *Stochastic Processes and Their Applications* **62**, 139–168. ISSN 0304-4149. Available at http://www.sciencedirect.com/science/article/pii/0304414995000909. MR1388767 https://doi.org/10.1016/0304-4149(95)00090-9

Dahlhaus, R. (1996c). Asymptotic statistical inference for nonstationary processes with evolutionary spectra. P. M. Robinson and M. Rosenblatt, eds. 145–159. New York, NY: Springer. MR1466743 https://doi.org/10.1007/978-1-4612-2412-9_11

Key words and phrases. Locally stationary process, tempered stable distribution, blocked Whittle estimation, two-step estimation.

Dahlhaus, R. (1997). Fitting time series models to nonstationary processes. *The Annals of Statistics* **25**, 1–37. ISSN 00905364. MR2634748

Dahlhaus, R. (2012). Locally stationary processes. In *Time Series Analysis: Methods and Applications* (S. S. R. Tata Subba Rao and C. R. Rao, eds.), *Handbook of Statistics* **30**, 351–413. Amsterdam: Elsevier. ISSN 0169-7161. MR3015127 https://doi.org/10.3726/978-3-653-01706-9

Dahlhaus, R. and Giraitis, L. (1998). On the optimal segment length for parameter estimates for locally stationary time series. *Journal of Time Series Analysis* **19**, 629–655. ISSN 1467-9892. MR1665941 https://doi.org/10.1111/1467-9892.00114

Dahlhaus, R. and Polonik, W. (2009). Empirical spectral processes for locally stationary time series. *Bernoulli* **15**, 1–39. ISSN 13507265. MR2546797 https://doi.org/10.3150/08-BEJ137

Feng, L. and Shi, Y. (2017). Fractionally integrated GARCH model with tempered stable distribution: A simulation study. *Journal of Applied Statistics* **44**, 2837–2857. ISSN 0266-4763. MR3721076 https://doi.org/10.1080/02664763.2016.1266310

Gonzales-Aparicio, I., Zucker, A., Carerri, F., Monforti, F., Huld, T. and Badger, J. (2016). EMHIRES dataset. Part I: Wind power generation European Meteorological derived HIgh resolution RES generation time series for present and future scenarios. Technical report, EUR 28171 EN; 10.2790/831549.

Grabchak, M. (2016a). *Tempered Stable Distributions: Stochastic Models for Multiscale Processes*. *SpringerBriefs in Mathematics*. Berlin: Springer. ISBN 9783319249278. MR3468484 https://doi.org/10.1007/978-3-319-24927-8

Grabchak, M. (2016b). On the consistency of the MLE for Ornstein-Uhlenbeck and other selfdecomposable processes. *Statistical Inference for Stochastic Processes* **19**, 29–50. ISSN 1572-9311. MR3472784 https://doi.org/10.1007/s11203-015-9118-9

Hitaj, A., Hubalek, F., Mercuri, L. and Rroji, E. (2018). On properties of the MixedTS distribution and its multivariate extension. *International Statistical Review* **86**, 512–540. MR3882129 https://doi.org/10.1111/insr.12265

Kawai, R. and Masuda, H. (2011). On simulation of tempered stable random variates. *Journal of Computational and Applied Mathematics* **235**, 2873–2887. ISSN 0377-0427. MR2763192 https://doi.org/10.1016/j.cam.2010.12.014

Kim, Y. S., Rachev, S., Dong, M. and Chung, D. (2006). The modified tempered stable distribution, GARCH models and option pricing. *Probability and Mathematical Statistics* **29**. MR2553002

Kim, Y. S., Rachev, S. T., Bianchi, M. L. and Fabozzi, F. (2008). Financial market models with Lévy processes and time-varying volatility. *Journal of Banking & Finance* **32**, 1363–1378.

Koponen, I. (1995). Analytic approach to the problem of convergence of truncated Lévy flights towards the Gaussian stochastic process. *Physical Review E* **52**, 1197–1199.

Küchler, U. and Tappe, S. (2013). Tempered stable distributions and processes. *Stochastic Processes and Their Applications* **123**, 4256–4293. ISSN 0304-4149. MR3096354 https://doi.org/10.1016/j.spa.2013.06.012

Lombardi, M. J. and Calzolari, G. (2008). Indirect estimation of α-stable distributions and processes. *Econometrics Journal* **11**, 193–208. ISSN 1368-423X. https://doi.org/10.1111/j.1368-423X.2008.00234.x

Olea, R., Palma, W., Rubio, P. and Vargas, M. (2021) LSTS: Locally Stationary Time Series. R package version 2.1. Available at https://CRAN.R-project.org/package=LSTS.

Rroji, E. and Mercuri, L. (2015). Mixed tempered stable distribution. *Quantitative Finance* **15**, 1559–1569. MR3378092 https://doi.org/10.1080/14697688.2014.969763

Sampaio, J. M. and Morettin, P. A. (2015). Indirect estimation of randomized generalized autoregressive conditional heteroskedastic models. *Journal of Statistical Computation and Simulation* **85**, 2702–2717. https://doi.org/10.1080/00949655.2014.934244

Sampaio, J. M. and Morettin, P. A. (2020). Stable randomized generalized autoregressive conditional heteroskedastic models. *Econometrics and Statistics* **15**, 67–83. ISSN 2452-3062. Available at https://www.sciencedirect.com/science/article/pii/S2452306218300947. MR4124735 https://doi.org/10.1016/j.ecosta.2018.11.002

She, R., Mi, Z. and Ling, S. (2022). Whittle parameter estimation for vector ARMA models with heavy-tailed noises. *Journal of Statistical Planning and Inference* **219**, 216–230. ISSN 0378-3758. Available at https://www.sciencedirect.com/science/article/pii/S0378375821001257. MR4357710 https://doi.org/10.1016/j.jspi.2021.12.003

Brazilian Journal of Probability and Statistics
2023, Vol. 37, No. 1, 177–194
https://doi.org/10.1214/23-BJPS566

An extension of the partially linear Rice regression model for bimodal and correlated data

Julio C. S. Vasconcelos[1,a], **Edwin M. M. Ortega**[1,b], **Roberto Vila**[2,3,c] and **Vicente G. Cancho**[4,d]

[1]*Department of Statistics, University of São Paulo, ESALQ, Piracicaba, Brazil,* [a]*juliocezarvasconcelos@hotmail.com,* [b]*edwin@usp.br*
[2]*Department of Statistics, University Brasilia, Brasilia, Brazil*
[3]*Department of Mathematics and Statistics, McMaster University, Hamilton, Ontario, Canada,* [c]*rovig161@gmail.com*
[4]*Department of Statistics, University of São Paulo, ICMC, São Carlos, Brazil,* [d]*garibay@icmc.usp.br*

Abstract. In this paper, we propose a new regression model based on an extension of the Rice distribution to model linear and nonlinear effects for correlated data in the presence of bimodality. The new model is referred to as the odd log-logistic Rice distribution and we provide general mathematical properties, including the event risk and moments. We discuss parameter estimation by the penalized maximum likelihood method. We also present several simulations with different parameter configurations and sample sizes to analyze the behavior of the maximum likelihood estimators, as well as to study the empirical distribution of the quantile residuals. The usefulness of the proposed regression model is proved empirically through analysis of a real dataset.

References

Dunn, P. K. and Smyth, G. K. (1996). Randomized quantile residuals. *Journal of Computational and Graphical Statistics* **5**, 236–244. https://doi.org/10.2307/1390802

Gleaton, J. U. and Lynch, J. D. (2006). Properties of generalized log-logistic families of lifetime distributions. *Journal of Probability and Statistical Science* **4**, 51–64. MR2240301

Green, P. J. and Silverman, B. W. (1993). *Nonparametric Regression and Generalized Linear Models: A Roughness Penalty Approach*. London: Chapman and Hall. Available at https://www.routledge.com/Nonparametric-Regression-and-Generalized-Linear-Models-A-roughness-penalty/Green-Silverman/p/book/9780412300400. MR1270012 https://doi.org/10.1007/978-1-4899-4473-3

Hastie, T. J. and Tibshirani, R. J. (1990). *Generalized Additive Models*. London: Chapman and Hall. Available at https://projecteuclid.org/journals/statistical-science/volume-1/issue-3/Generalized-Additive-Models/10.1214/ss/1177013604.full. MR1082147

Ibacache-Pulgar, G., Paula, G. A. and Galea, M. (2012). Influence diagnostics for elliptical semiparamteric mixed models. *Statistical Modelling* **12**, 165–193. MR2963795 https://doi.org/10.1177/1471082X1001200203

Ke, C. and Wang, Y. (2001). Semiparametric nonlinear mixed-effects models and their applications. *Journal of the American Statistical Association* **96**, 1272–1298. MR1946577 https://doi.org/10.1198/016214501753381913

Kenward, M. G. (1987). A method for comparing profiles of repeated measurements. *Journal of the Royal Statistical Society Series C Applied Statistics* **36**, 296–308. https://doi.org/10.2307/2347788

Lauwers, L., Barbé, K., Van Moer, W. and Pintelon, R. (2009). Estimating the parameters of a Rice distribution: A Bayesian approach. In *2009 IEEE Instrumentation and Measurement Technology Conference*, 114–117. New York: IEEE. Available at https://ieeexplore.ieee.org/document/5168426.

Li, W. and Xue, L. (2015). Efficient inference in a generalized partially linear model with random effect for longitudinal data. *Communications in Statistics – Theory and Methods* **44**, 241–260. MR3292591 https://doi.org/10.1080/03610926.2012.740126

Li, X., Bilen-Green, C., Farahmand, K. and Langley, L. (2018). A semiparametric method for estimating the progression of cognitive decline in dementia. *IISE Transactions on Healthcare Systems Engineering* **8**, 303–314. https://doi.org/10.1080/24725579.2018.1455247

Key words and phrases. P-splines, partially linear regression, residual analysis, simulation study.

Mattos, T. B., Matos, L. A. and Lachos, V. H. (2021). A semiparametric mixed-effects model for censored longitudinal data. *Statistical Methods in Medical Research* **30**, 2582–2603. MR4345214 https://doi.org/10.1177/09622802211046387

Prataviera, F., Ortega, E. M. M. and Cordeiro, G. M. (2020). A new bimodal Maxwell regression model with engineering applications. *Applied Mathematics & Information Sciences* **14**, 817–831. MR4158299 https://doi.org/10.18576/amis

Rice, S. O. (1945). Mathematical analysis of random noise. *The Bell System Technical Journal* **24**, 46–156. MR0011918 https://doi.org/10.1002/j.1538-7305.1945.tb00453.x

Rigby, R. A. and Stasinopoulos, D. M. (2004). Smooth centile curves for skew and kurtotic data modelled using the Box Cox power exponential distribution. *Statistics in Medicine* **23**, 3053–3076. MR2252360 https://doi.org/10.1191/1471082X06st122oa

Stasinopoulos, D. M. and Rigby, R. A. (2007). Generalized additive models for location scale and shape (GAMLSS) in R. *Journal of Statistical Software* **23**, 1–46. https://doi.org/10.18637/jss.v023.i07

Stasinopoulos, M. D., Rigby, R. A., Heller, G. Z., Voudouris, V. and De Bastiani, F. (2017). *Flexible Regression and Smoothing: Using Gamlss in R*. New York: CRC Press. MR3799717 https://doi.org/10.1177/1471082X18759144

Szczesniak, R. D., Li, D. and Raouf, S. A. (2015). Semiparametric mixed models for medical monitoring data: An overview. *Journal of Biometrics and Biostatistics* **6**, 1–10. Available at https://pubmed.ncbi.nlm.nih.gov/29593934/.

Wood, S. N. (2006). *Generalized Additive Models. An Introduction with R*. London: Chapman & Hall. MR2206355

Wright, K. (2014). Agridat: Agricultural datasets. R package version, 1. Available at http://CRAN.R-project.org/package=agridat.

Yakovleva, T. (2019). Nonlinear properties of the Rice statistical distribution: Theory and applications in stochastic data analysis. *Journal of Applied Mathematics and Physics* **7**, 2767–2779. Available at https://www.scirp.org/journal/paperinformation.aspx?paperid=96330.

Yakovleva, T. V. and Kulberg, N. S. (2013). Noise and signal estimation in MRI: Two-parametric analysis of rice-distributed data by means of the maximum likelihood approach. *American Journal of Theoretical and Applied Statistics* **2**, 67–79.

Yee, T. W. (2008). The VGAM package. *R News* **8**, 28–39.

Zhang, W., Fan, J. and Sun, Y. (2009). A semiparamtric model for cluster data. *The Annals of Statistics* **37**, 2377–2408. MR2543696 https://doi.org/10.1214/08-AOS662

Zhang, W., Leng, C. and Tang, C. Y. (2015). A joint modelling approach for longitudinal studies. *Journal of the Royal Statistical Society, Series B, Statistical Methodology* **77**, 219–238. MR3299406 https://doi.org/10.1111/rssb.12065

Brazilian Journal of Probability and Statistics
2023, Vol. 37, No. 1, 195–218
https://doi.org/10.1214/23-BJPS567

Scaling limits and fluctuations of a family of N-urn branching processes

Lirong Ren[a] **and Xiaofeng Xue**[b]

School of Mathematics and Statistics, Beijing Jiaotong University, Beijing, China, [a]*20121634@bjtu.edu.cn,* [b]*xfxue@bjtu.edu.cn*

Abstract. In this paper, we are concerned with a family of N-urn branching processes, where some particles are initially placed in N urns, and then each particle gives birth to several new particles in some urn when dies. This model includes the N-urn Ehrenfest model and N-urn branching random walk as special cases. We show that the scaling limit of the process is driven by a $C(\mathbb{T})$-valued linear ordinary differential equation and the fluctuation of the process is driven by a generalized Ornstein–Uhlenbeck process in the dual of $C^{\infty}(\mathbb{T})$, where $\mathbb{T} = (0, 1]$ is the one-dimensional torus. A crucial step for the proofs of the above main results is to show that numbers of particles in different urns are approximately independent. As applications of our main results, the limit theorems of the hitting times of the process are also discussed.

References

Ethier, S. N. and Kurtz, T. G. (1985). *Markov Processes: Characterization and Convergence*. New York: Wiley.

Holley, R. A. and Stroock, D. W. (1978). Generalized Ornstein–Uhlenbeck processes and infinite particle branching Brownian motions. *Publications of the Research Institute for Mathematical Sciences* **14**, 741–788.

Kurtz, T. (1978). Strong approximation theorems for density dependent Markov chains. *Stochastic Processes and Their Applications* **6**, 223–240.

Lang, S. (1983). *Undergraduate Analysis*, 1st ed. New York: Springer.

Mitoma, I. (1983). Tightness of probabilities on $C([0, 1], S')$ and $D([0, 1], S')$. *Annals of Probability* **11**, 989–999.

Whitt, W. (2007). Proofs of the martingale FCLT. *Probability Surveys* **4**, 268–302.

Xin, C., Zhao, M., Yao, Q. and Cui, E. (2020). On the distribution of the hitting time for the N-urn Ehrenfest model. *Statistics & Probability Letters* **157**, 108625, 11 pages.

Xue, X. F. (2022). Hydrodynamics of the generalized N-urn Ehrenfest model. *Potential Analysis*. Available at https://rdcu.be/cFzrj.

Key words and phrases. Branching process, scaling limit, fluctuation.

Brazilian Journal of Probability and Statistics
2023, Vol. 37, No. 1, 219–245
https://doi.org/10.1214/23-BJPS568

On the finiteness of the moments of the measure of level sets of random fields

Diego Armentano[1,a], **Jean Marc Azaïs**[2,d], **Federico Dalmao**[3,e], **José Rafael León**[1,4,b] **and Ernesto Mordecki**[1,c]

[1] *Universidad de la República, Montevideo, Uruguay,* [a]*diego@cmat.edu.uy,* [b]*rlramos@fing.edu.uy,* [c]*mordecki@cmat.edu.uy*
[2] *IMT, Université de Toulouse, Toulouse, France,* [d]*jean-marc.azais@math.univ-toulouse.fr*
[3] *Universidad de la República, Salto, Uruguay,* [e]*fdalmao@unorte.edu.uy*
[4] *Universidad Central de Venezuela, Caracas, Venezuela*

Abstract. General conditions on smooth real valued random processes and fields are given to ensure the finiteness of the moments of the measure of their level sets. As a by product, a new generalized Kac–Rice formula for the expectation of the measure of these level sets in the one-dimensional case is obtained when the second moment is uniformly bounded. The conditions involve: (i) the differentiability of the trajectories up to a certain order k; (ii) the finiteness of the moments of the k-th partial derivatives of the field up to another order; and (iii) the boundedness of the field's joint density and some of its derivatives. Particular attention is given to the shot noise processes and fields. Other applications include stationary Gaussian processes, Chi-square processes and regularized diffusion processes. We also sketch the application of these tools to study critical points of random fields.

References

Adler, R. (1981). *The Geometry of Random Fields*. Chichester: John Wiley & Sons, Ltd. MR0611857

Adler, R. and Taylor, J. (2007). *Random Fields and Geometry*. New York: Springer. MR2319516

Ancona, M. and Letendre, T. (2021). Zeros of smooth stationary Gaussian processes. *Electronic Journal of Probability* **26**, Paper No. 68, 81 pp. MR4262341 https://doi.org/10.1214/21-ejp637

Angst, J. and Poly, G. (2018). On the absolute continuity of the random nodal volumes. *Annals of Probability* **48**(5), 2145–2175. MR4152638 https://doi.org/10.1214/19-AOP1418

Armentano, D., Azais, J.-M., Ginsbourger, D. and León, J. R. (2019). Conditions for the finiteness of the moments of the volume of level sets. *Electronic Communications in Probability* **24**, Paper No. 17, 8 pp. MR3933041 https://doi.org/10.1214/19-ECP214

Azaïs, J.-M. and Wschebor, M. (2009). *Level Sets and Extrema of Random Processes and Fields*. Hoboken, NJ: John Wiley & Sons, Inc. MR2478201 https://doi.org/10.1002/9780470434642

Azaïs, J. M. and León, J. R. (2020). Necessary and sufficient conditions for the finiteness of the second moment of the measure of level sets. *Electronic Journal of Probability* **26**, Paper No. 107, 15 pp. MR4147520

Baxevani, A., Podgorki, K. and Rychlik, I. (2014). Sample path asymmetries in non-Gaussian random seas. *Scandinavian Journal of Statistics* **4**, 1102–1123. MR3277040 https://doi.org/10.1111/sjos.12086

Belyaev, Y. K. (1966). On the number of intersections of a level by a Gaussian stochastic process. *Theory of Probability and Its Applications* **11**, 106–113. MR0195178

Berzin, C., Latour, A. and León, J. R. (2022). Kac–Rice formula: A contemporary overview of the main results and applications. arXiv:2205.08742.

Biermé, H. and Desolneux, A. (2011). Regularity and crossings of shot noise processes. MAP5 2010-20. hal-00484118v2. MR3024968 https://doi.org/10.1214/11-AAP807

Biermé, H. and Desolneux, A. (2012a). A Fourier approach for the level crossings of shot noise processes with jumps. *Journal of Applied Probability* **49**, 100–113. MR2952884 https://doi.org/10.1239/jap/1331216836

Biermé, H. and Desolneux, A. (2012b). Crossings of smooth shot noise processes. *Annals of Applied Probability* **22**, 2240–2281. MR3024968 https://doi.org/10.1214/11-AAP807

Key words and phrases. Moments of measure of level sets, Kac–Rice formula, Crofton formula, shot noise process.

Biermé, H. and Desolneux, A. (2016). On the perimeter of excursion sets of shot noise random fields. *Annals of Probability* **44**, 521–543. MR3457393 https://doi.org/10.1214/14-AOP980

Borodin, A. N. (2017). *Stochastic Processes*. Cham: Springer. MR3727115 https://doi.org/10.1007/978-3-319-62310-8

Borovkov, K. and Last, G. (2012). On Rice formula for stationary multivariate piecewise smooth process. *Journal of Applied Probability* **49**, 351–363. MR2977800 https://doi.org/10.1239/jap/1339878791

Dalmao, F. and Mordecki, E. (2015). Rice formula for processes with jumps and applications. *Extremes* **18**, 15–35. MR3317851 https://doi.org/10.1007/s10687-014-0200-2

Gass, L. (2021). Cumulants asymptotics for the zeros counting measure of real Gaussian processes. arXiv:2112.08247.

Geman, D. (1972). On the variance of the number of zeros of a stationary Gaussian process. *Annals of Mathematical Statistics* **43**, 977–982. MR0301791 https://doi.org/10.1214/aoms/1177692560

Howard, R. (1993). The kinematic formula in Riemannian homogeneous spaces. *Memoirs of the American Mathematical Society* **106**(509). MR1169230 https://doi.org/10.1090/memo/0509

Kac, M. (1944). On the average number of real roots of a random algebraic equation. *Bulletin of the American Mathematical Society* **49**, 282–332. MR0007812 https://doi.org/10.1090/S0002-9904-1943-07912-8

Knapp, A. W. (1996). *Lie groups: Beyond an Introduction*. Boston, Mass: Birkhäuser. MR1399083 https://doi.org/10.1007/978-1-4757-2453-0

Kratz, M. and León, J. R. (2006). On the second moment of the number of crossings by a stationary Gaussian process. *Annals of Probability* **34**, 1601–1607. MR2257657 https://doi.org/10.1214/009117906000000142

Leonenko, M. M. (1975). The central limit theorem for homogeneous random fields, and the asymptotic normality of estimators of the regression coefficients. *Ukrainian Mathematical Journal* **27**, 556–559. MR0353431

Machado, U. and Rychlik, I. (2003). Wave statistics in non-linear random seas. *Extremes* **6**, 125–146. MR2076640 https://doi.org/10.1023/B:EXTR.0000025663.45811.9b

Marcus, M. (1977). Level crossings of a stochastic process with absolutely continuous sample paths. *Annals of Probability* **5**, 52–71. MR0426124 https://doi.org/10.1214/aop/1176995890

Morgan, F. (2016). *Geometric Measure Theory, a Beginner's Guide*, 5th ed. Amsterdam: Elsevier/Academic Press. MR3497381

Nualart, D. and Wschebor, M. (1991). Intégration par parties dans l'espace de Wiener et approximation du temps local. *Probability Theory and Related Fields* **90**, 83–109. MR1124830 https://doi.org/10.1007/BF01321135

Orsingher, E. and Battaglia, F. (1982). Probability distributions and level crossings of shot noise models. *Stochastics* **8**, 45–61. MR0687045 https://doi.org/10.1080/17442508208833227

Podgorki, K. and Rychlik, I. (2008). Envelope crossing distribution for Gaussian fields. *Probalistic Engineering Mechanics* **23**, 364–371.

Podgorki, K., Rychlik, I. and Wallin, J. (2015). Slepian noise approach for Gaussian and Laplace moving average processes. *Extremes* **4**, 665–695. MR3418772 https://doi.org/10.1007/s10687-015-0227-z

Protter, P. (2004). *Stochastic Integration and Differential Equations*, 2nd ed. Berlin: Springer. MR2020294

Rice, S. O. (1944). Mathematical analysis of random noise. *The Bell System Technical Journal* **23**, 282–332. MR0010932 https://doi.org/10.1002/j.1538-7305.1944.tb00874.x

Rice, S. O. (1945). Mathematical analysis of random noise II. *The Bell System Technical Journal* **24**, 46–156. MR0011918 https://doi.org/10.1002/j.1538-7305.1945.tb00453.x

Santaló, L. A. (1976). *Integral Geometry and Geometric Probability*. Reading, MA: Addison-Wesley. MR0433364

Worsley, K. J. (1995a). Boundary corrections for the expected Euler characteristic of excursion sets of random fields, with an application to astrophysics. *Advances in Applied Probability* **27**, 943–959. MR1358902 https://doi.org/10.2307/1427930

Worsley, K. J. (1995b). Estimating the number of peaks in a random field using the Hadwiger characteristic of excursion sets, with applications to medical images. *Annals of Statistics* **23**, 640–669. MR1332586 https://doi.org/10.1214/aos/1176324540

Wschebor, M. (1985). *Surfaces aléatoires, Mesure géométrique des ensembles de niveau. Lecture Notes in Mathematics* **1147**. Berlin: Springer. MR0871689 https://doi.org/10.1007/BFb0075073

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: https://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 https://www.e-publications.org/ims/submission/.

Preparation of Manuscripts: https://imstat.org/journals-and-publications/brazilian-journal-of-probability-and-statistics/brazilian-journal-of-probability-and-statistics-manuscript-submission/

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. https://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, PO Box 729, Middletown, Maryland 21769, USA. 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:
https://www.imstat.org/individual-membership/

Individual and General Subscriptions:
https://imstat.org/journals-and-publications/subscriptions-and-orders/

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

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