Contents

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

Brazilian Journal of Probability and Statistics

Volume 38 • Number 3 • September 2024

Brazilian Journal of Probability and Statistics

Volume 38 • Number 3 • September 2024

www.imstat.org/bjps

ISSN 0103-0752 (Print) ISSN 2317-6199 (Online), Volume 38, Number 3, September 2024. 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
2024, Vol. 38, No. 3, 339–346
https://doi.org/10.1214/24-BJPS608

Asymptotics of two-point correlations in the multi-species q-TAZRP

Jeffrey Kuan[a] **and Zhengye Zhou**

Texas A&M University, College Station, United States, [a]*jkuan@tamu.edu*

Abstract. A previous paper by the authors found explicit contour integral formulas for certain joint moments of the multi-species q-TAZRP (totally asymmetric zero range process), using algebraic methods. These contour integral formulas have a "pseudo–factorized" form which makes asymptotic analysis simpler. In this brief note, we use those contour integral formulas to find the asymptotics of the two-point correlations. As expected, the term arising from the "shift–invariance" makes a non-trivial asymptotic contribution.

References

Ayala, M., Carinci, G. and Redig, F. (2018). Quantitative Boltzmann-Gibbs principles via orthogonal polynomial duality. *Journal of Statistical Physics* **171**, 980–999. MR3805584 https://doi.org/10.1007/s10955-018-2060-7

Ayala, M., Carinci, G. and Redig, F. (2021). Higher order fluctuation fields and orthogonal duality polynomials. *Electronic Journal of Probability* **26**, Paper No. 27, 35 pp. MR4235478 https://doi.org/10.1214/21-EJP586

Baldasso, R., Menezes, O., Neumann, A. and Souza, R. R. (2017). Exclusion process with slow boundary. *Journal of Statistical Physics* **167**, 1112–1142. MR3647054 https://doi.org/10.1007/s10955-017-1763-5

Belitsky, V. and Schütz, G. M. (2015a). Self-duality for the two-component asymmetric simple exclusion process. *Journal of Mathematics and Physics* **56**, 083302, 20 pp. MR3391361. MR3391361 https://doi.org/10.1063/1.4929663

Belitsky, V. and Schütz, G. M. (2015b). Quantum algebra symmetry of the ASEP with second-class particles. *Journal of Statistical Physics* **161**, 821–842. MR3413634 https://doi.org/10.1007/s10955-015-1363-1

Belitsky, V. and Schütz, G. M. (2018). Self-duality and shock dynamics in the n-species priority ASEP. *Stochastic Processes and Their Applications* **128**, 1165–1207. MR3769659 https://doi.org/10.1016/j.spa.2017.07.003

Borodin, A., Gorin, V. and Wheeler, M. (2022). Shift-invariance for vertex models and polymers. *Proceedings of the London Mathematical Society* **3**, 182–299. MR4400934 https://doi.org/10.1112/plms.12427

Carinci, G., Franceschini, C., Giardinà, C., Groenevelt, W. and Redig, F. (2019). Orthogonal dualities of Markov processes and unitary symmetries. *SIGMA* **15**, Paper No. 053, 27 pp. MR3980548 https://doi.org/10.3842/SIGMA.2019.053

Carinci, G., Franceschini, C. and Groenevelt, W. (2021). q-Orthogonal dualities for asymmetric particle systems. *Electronic Journal of Probability* **26**, Paper No. 108, 38 pp. MR4290510 https://doi.org/10.1214/21-EJP663

Franceschini, C. and Giardinà, C. (2019). Stochastic duality and orthogonal polynomials. In *Sojourns in Probability Theory and Statistical Physics. III. Interacting Particle Systems and Random Walks, a Festschrift for Charles M. Newman*. *Springer Proc. Math. Stat.* **300**, 187–214. Singapore: Springer. MR4044393 https://doi.org/10.1007/978-981-15-0302-3_7

Franceschini, C., Gonçalves, P. and Sau, F. (2022). Symmetric inclusion process with slow boundary: Hydrodynamics and hydrostatics. *Bernoulli* **28**, 1340–1381. MR4388941 https://doi.org/10.3150/21-bej1390

Franceschini, C., Kuan, J. and Zhou, Z. (2022). Orthogonal polynomial duality and unitary symmetries of multi-species ASEP(q,θ) and higher–spin vertex models via $*$–bialgebra structure of higher rank quantum groups. ArXiv e-prints. MR4729864 https://doi.org/10.1007/s00220-024-04979-8

Galashin, P. (2021). Symmetries of stochastic colored vertex models. *Annals of Probability* **49**, 2175–2219. MR4317703 https://doi.org/10.1214/20-aop1502

Groenevelt, W. (2019). Orthogonal stochastic duality functions from Lie algebra representations. *Journal of Statistical Physics* **174**, 97–119. MR3904511 https://doi.org/10.1007/s10955-018-2178-7

He, J. Shift invariance of half space integrable models. ArXiv e-prints: arXiv:2205.13029.

Kuan, J. (2018a). A multi-species ASEP(q,j) and q-TAZRP with stochastic duality. *International Mathematics Research Notices* **17**, 5378–5416. MR3862123 https://doi.org/10.1093/imrn/rnx034

Kuan, J. (2018b). An algebraic construction of duality functions for the stochastic $\mathcal{U}_q(A_n^{(1)})$ vertex model and its degenerations. *Communications in Mathematical Physics* **359**, 121–187. MR3781448 https://doi.org/10.1007/s00220-018-3108-x

Key words and phrases. Two-point correlations, multi-species q-TAZRP.

Kuan, J. (2021). Joint q-moments and shift invariance for the multi-species q-TAZRP on the infinite line. *Annales de l'Institut Henri Poincare (D) Combinatorics, Physics and their Interactions.* to appear in.

Kuniba, A., Mangazeev, V. V., Maruyama, S. and Okado, M. (2016). Stochastic R matrix for $U_q(A_n^{(1)})$. *Nuclear Physics B* **913**, 248–277. MR3575300 https://doi.org/10.1016/j.nuclphysb.2016.09.016

Nemes, G. and Olde, D. (2019). Asymptotic expansions for the incomplete gamma function in the transition regions. *Mathematics Competitions* **88**, 1805–1827. MR3925486 https://doi.org/10.1090/mcom/3391

Tsunoda, K. (2020). Hydrostatic limit for exclusion process with slow boundary revisited. In *Stochastic Analysis on Large Scale Interacting Systems*. *RIMS Kôkyûroku Bessatsu, B79, Res. Inst. Math. Sci.*, 149–162. Kyoto: RIMS. MR4279016

Uchiyama, M. and Wadati, M. (2005). Correlation function of asymmetric simple exclusion process with open boundaries. *Journal of Nonlinear Mathematical Physics* **12**, 676–688. MR2118894 https://doi.org/10.2991/jnmp.2005.12.s1.52

Yoshihiro, T. (2015). Algebraic construction of multi-species q-Boson system. Available at arXiv:1507.02033.

Zhou, Z. (2021). Orthogonal polynomial stochastic duality functions for multi-species SEP($2j$) and multi-species IRW. *SIGMA* **17**, Paper No. 113, 11 pp. MR4355490 https://doi.org/10.3842/SIGMA.2021.113

Brazilian Journal of Probability and Statistics
2024, Vol. 38, No. 3, 347–369
https://doi.org/10.1214/24-BJPS609

A new regression model for the analysis of bimodal censored data: A comparison with random survival forest

Gabriela M. Rodrigues[1,a], **Edwin M. M. Ortega**[1,b], **Roberto Vila**[2,c] **and Rafael De Andrade Moral**[3,d]

[1]*Department of Exact Science, University of São Paulo, São Paulo, Brazil,* [a]*gabrielar@usp.br,* [b]*edwin@usp.br*
[2]*Department of Statistics, University of Brasília, Brasília, Brazil,* [c]*rovig161@unb.br*
[3]*Department of Mathematics and Statistics, Maynooth University, Maynooth, Irlanda,* [d]*rafael.deandrademoral@mu.ie*

Abstract. We propose a new regression model based on the extension of the generalized Rayleigh distribution. This regression model can be used to model survival data where the hazard rate function is increasing, decreasing, bathtub or unimodal-shaped. We employ a frequentist estimation procedure for the parameters of the proposed model. The predictive performance of the new model is compared with a machine learning approach based on random survival forests, where we explore model validation and computational aspects. We carry out simulation studies for different parameter settings, sample sizes, and censoring percentages, considering the new regression model and the machine learning methodology. Finally, we showcase our methods using real data from a medical case study.

References

Adham, D., Abbasgholizadeh, N. and Abazari, M. (2017). Prognostic factors for survival in patients with gastric cancer using a random survival forest. Asian Pacific journal of cancer prevention. *APJCP* **18**, 129.

Alizadeh, M., Tahmasebi, S. and Haghbin, H. (2020). The exponentiated odd log-logistic family of distributions: Properties and applications. *Journal of Statistical Modelling: Theory and Applications* **1**, 29–52. MR3945455 https://doi.org/10.1556/012.2019.56.1.1420

Alonso, L. and Renard, F. (2020). A new approach for understanding urban microclimate by integrating complementary predictors at different scales in regression and machine learning models. *Remote Sensing* **12**, 2434.

Alzaatreh, A., Lee, C. and Famoye, F. (2013). A new method for generating families of continuous distributions. *Metron* **71**, 63–79. MR3080620 https://doi.org/10.1007/s40300-013-0007-y

Babaei Rikan, S., Sorayaie Azar, A., Naemi, A., Bagherzadeh Mohasefi, J., Pirnejad, H. and Wiil, U. K. (2024). Survival prediction of glioblastoma patients using modern deep learning and machine learning techniques. *Scientific Reports* **14**, 2371.

Borra, S. and Di Ciaccio, A. (2010). Measuring the prediction error. A comparison of cross-validation, bootstrap and covariance penalty methods. *Computational Statistics & Data Analysis* **54**, 2976–2989. MR2727727 https://doi.org/10.1016/j.csda.2010.03.004

Breiman, L. (1984). Classification and regression trees. Routledge.

Breiman, L. (2001). Random forests. *Machine Learning* **45**, 5–32. MR3874153

Burman, P. (1989). A comparative study of ordinary cross-validation, v-fold cross-validation and the repeated learning-testing methods. *Biometrika* **76**, 503–514. MR1040644 https://doi.org/10.1093/biomet/76.3.503

Cavalcante, T., Ospina, R., Leiva, V., Cabezas, X. and Martin-Barreiro, C. (2023). Weibull regression and machine learning survival models: Methodology, comparison, and application to biomedical data related to cardiac surgery. *Biology* **12**, 442.

Cerono, G. and Chicco, D. (2024). Ensemble machine learning reveals key features for diabetes duration from electronic health records. *PeerJ Computer Science* **10**, e1896.

Choudhury, A. (2005). A simple derivation of moments of the exponentiated Weibull distribution. *Metrika* **62**, 17–22. MR2236293 https://doi.org/10.1007/s001840400351

Cordeiro, G. M. and de Castro, M. (2011). A new family of generalized distributions. *Journal of Statistical Computation and Simulation* **81**, 883–898. MR2806932 https://doi.org/10.1080/00949650903530745

Key words and phrases. Survival analysis, cross validation, epidemiological data, random forests, decision tree.

Dagum, C. (1975). A model of income distribution and the conditions of existence of moments of finite order. *Bulletin of the International Statistical Institute* **46**, 199–205.

Díaz, I., Hoffman, K. L. and Hejazi, N. S. (2024). Causal survival analysis under competing risks using longitudinal modified treatment policies. *Lifetime Data Analysis* **30**, 213–236. MR4684790 https://doi.org/10.1007/s10985-023-09606-7

Eugene, N., Lee, C. and Famoye, F. (2002). Beta-normal distribution and its applications. *Communications in Statistics - Theory and Methods* **31**, 497–512. MR1902307 https://doi.org/10.1081/STA-120003130

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

Harrell, F. E., Califf, R. M., Pryor, D. B., Lee, K. L. and Rosati, R. A. (1982). Evaluating the yield of medical tests. *JAMA* **247**, 2543–2546.

Ishwaran, H., Kogalur, U. B., Blackstone, E. H. and Lauer, M. S. (2008). Random survival forests. *Annals of Applied Statistics* **2**, 841–860. MR2516796 https://doi.org/10.1214/08-AOAS169

Ishwaran, H., Kogalur, U. B. and Kogalur, M. U. B. (2023). Package 'randomForestSRC'. Breast 6 854.

Jung, S. Y., Papp, J. C., Sobel, E. M., Yu, H. and Zhang, Z.-F. (2019). Breast cancer risk and insulin resistance: Post genome-wide gene–environment interaction study using a random survival forest. *Cancer Research* **79**, 2784–2794.

Khan, M. A., Shah, M. I., Javed, M. F., Khan, M. I., Rasheed, S., El-Shorbagy, M., El-Zahar, E. R. and Malik, M. (2022). Application of random forest for modelling of surface water salinity. *Ain Shams Engineering Journal* **13**, 101635.

Kundu, D. and Raqab, M. Z. (2005). Generalized Rayleigh distribution: Different methods of estimations. *Computational Statistics & Data Analysis* **49**, 187–200. MR2129172 https://doi.org/10.1016/j.csda.2004.05.008

Lanjoni, B. R., Ortega, E. M. and Cordeiro, G. M. (2016). Extended Burr XII regression models: Theory and applications. *Journal of Agricultural, Biological, and Environmental Statistics* **21**, 203–224. MR3459302 https://doi.org/10.1007/s13253-015-0236-z

LeBlanc, M. and Crowley, J. (1993). Survival trees by goodness of split. *Journal of the American Statistical Association* **88**, 457–467. MR1224370

Lux, T. and Papapantoleon, A. (2017). Improved Fréchet–Hoeffding bounds on d-copulas and applications in model-free finance. *The Annals of Applied Probability* **27**, 3633–3671. MR3737934 https://doi.org/10.1214/17-AAP1292

Maia, M., Pimentel, J., Ospina, R. and Ara, A. (2023). Wavelet support vector censored regression. *Analytics* **2023**, 410–425.

Marshall, A. W. and Olkin, I. (1997). A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families. *Biometrika* **84**, 641–652. MR1603936 https://doi.org/10.1093/biomet/84.3.641

Mudholkar, G. S. and Srivastava, D. K. (1993). Exponentiated Weibull family for analyzing bathtub failure-rate data. *IEEE Transactions on Reliability* **42**, 299–302.

Mudholkar, G. S., Srivastava, D. K. and Kollia, G. D. (1996). A generalization of the Weibull distribution with application to the analysis of survival data. *Journal of the American Statistical Association* **91**, 1575–1583. MR1439097 https://doi.org/10.2307/2291583

Ortega, E. M., Cordeiro, G. M., Campelo, A. K., Kattan, M. W. and Cancho, V. G. (2015). A power series beta Weibull regression model for predicting breast carcinoma. *Statistics in Medicine* **34**, 1366–1388. MR3322774 https://doi.org/10.1002/sim.6416

Oukawa, G. Y., Krecl, P. and Targino, A. C. (2022). Fine-scale modeling of the urban heat island: A comparison of multiple linear regression and random forest approaches. *Science of the Total Environment* **815**, 152836.

Qiu, X., Gao, J., Yang, J., Hu, J., Hu, W., Kong, L. and Lu, J. J. (2020). A comparison study of machine learning (random survival forest) and classic statistic (Cox proportional hazards) for predicting progression in high-grade glioma after proton and carbon ion radiotherapy. *Frontiers in Oncology* **10**, 551420.

R Core Team (2021), R: A language and environment for statistical computing. Vienna, Austria.

Ramires, T. G., Cordeiro, G. M., Kattan, M. W., Hens, N. and Ortega, E. M. (2018). Predicting the cure rate of breast cancer using a new regression model with four regression structures. *Statistical Methods in Medical Research* **27**, 3207–3223. MR3865269 https://doi.org/10.1177/0962280217695344

Rigby, R. A. and Stasinopoulos, D. M. (2005). Generalized additive models for location, scale and shape. *Journal of the Royal Statistical Society Series C Applied Statistics* **54**, 507–554. MR2137253 https://doi.org/10.1111/j.1467-9876.2005.00510.x

Rodrigues, G. M., Ortega, E. M. M., Vila, V. and De Andrade Moral, R. (2024). Supplement to "A new regression model for the analysis of bimodal censored data: A comparison with random survival forest." https://doi.org/10.1214/24-BJPS609SUPP

Shaked, M. and Shanthikumar, J. G. (1994). Stochastic orders and their applications. (No Title). MR1278322

Sklar, A. (1973). Random variables, joint distribution functions, and copulas. *Kybernetika* **9**, 449–460. MR0345164

Stasinopoulos, M., Rigby, B., Voudouris, V., Akantziliotou, C., Enea, M. and Kiose, D. (2023). *Package 'Gamlss'. Dist'2020*. Available at http://www.gamlss.org.

Subeesh, A., Bhole, S., Singh, K., Chandel, N. S., Rajwade, Y. A., Rao, K., Kumar, S. and Jat, D. (2022). Deep convolutional neural network models for weed detection in polyhouse grown bell peppers. *Artificial Intelligence in Agriculture* **6**, 47–54.

Surles, J. and Padgett, W. (1998). Inference for $P(Y < X)$ in the Burr type X model. *Journal of Applied Statistical Sciences* **7**, 225–238.

Surles, J. and Padgett, W. (2001). Inference for reliability and stress-strength for a scaled Burr type X distribution. *Lifetime Data Analysis* **7**, 187–200. MR1842327 https://doi.org/10.1023/A:1011352923990

Surles, J. and Padgett, W. (2005). Some properties of a scaled Burr type X distribution. *Journal of Statistical Planning and Inference* **128**, 271–280. MR2110189 https://doi.org/10.1016/j.jspi.2003.10.003

Zografos, K. and Balakrishnan, N. (2009). On families of beta-and generalized gamma-generated distributions and associated inference. *Statistical Methodology* **6**, 344–362. MR2751078 https://doi.org/10.1016/j.stamet.2008.12.003

Brazilian Journal of Probability and Statistics
2024, Vol. 38, No. 3, 370–410
https://doi.org/10.1214/24-BJPS610

Estimation of tail risk using extreme expectiles in linear GARCH models with heavy-tailed error

Fuming Lin[1], **Saralees Nadarajah**[2,a] **and Daimin Shi**[3]

[1]*School of Mathematics and Statistics, Sichuan University of Science and Engineering, Sichuan Zigong, China*
[2]*Department of Mathematics, University of Manchester, Manchester, United Kingdom,*
[a]*mbbsssn2@manchester.ac.uk*
[3]*School of Statistics, Southwestern University of Finance and Economics, Sichuan Chengdu, China*

Abstract. Extreme expectiles recently found their high performance in estimation of financial tail risk as expectile is a unique risk measure of both coherence and elicitability. In view of the conditional heteroscedasticity phenomenon in financial data, this paper is devoted to estimation of extreme conditional and unconditional expectiles in linear GARCH models with heavy-tailed error. Estimates are used to assess VaR and MES in extreme cases. Precisely, we employ a mature sieve approximation method to fit linear GARCH models to the data and obtain estimators of conditional variances and residuals. Extreme expectiles of these residuals are used as estimators of counterparts of the errors, and then combining the estimators of conditional variances, estimators are given for the extreme conditional and unconditional expectiles of the whole model. Moreover, based on these expectiles estimators, conditional marginal expected shortfall is also estimated. Asymptotic theory for estimators is derived. A simulation study shows merits of our methods. Finally, the methods are applied to two real data sets.

References

Acharya, V. V., Pedersen, L. H., Philippon, T. and Richardson, M. (2012). Measuring systemic risk. London. Discussion Paper DP8824.

Adrian, T. and Brunnermeier, M. K. (2011). CoVaR. National Bureau of Economic Research (No. w17454).

Artzner, P., Delbaen, F., Eber, J. M. and Heath, D. (1999). Coherent measures of risk. *Mathematical Finance* **9**, 203–228. MR1850791 https://doi.org/10.1111/1467-9965.00068

Bellini, F. and Bignozzi, V. (2015). On elicitable risk measures. *Quantitative Finance* **15**, 725–733. MR3334566 https://doi.org/10.1080/14697688.2014.946955

Bellini, F. and Di Bernardino, E. (2017). Risk management with expectiles. *European Journal of Finance* **23**, 487–506.

Bellini, F., Klar, B., Muller, A. and Gianina, E. R. (2014). Generalized quantiles as risk measures. *Insurance: Mathematics and Economics* **54**, 41–48. MR3145849 https://doi.org/10.1016/j.insmatheco.2013.10.015

Ben Tal, A. and Teboulle, M. (2007). An old new concept of convex risk measures: The optimized certainty equivalent. *Mathematical Finance* **17**, 449–476. MR2332265 https://doi.org/10.1111/j.1467-9965.2007.00311.x

Bingham, N. H., Goldie, C. M. and Teugels, J. L. (1987). *Regular Variation*. Cambridge: Cambridge University Press. MR1015093

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* **31**, 307–327. MR0853051 https://doi.org/10.1016/0304-4076(86)90063-1

Bradley, R. C. (2005). Basic properties of strong mixing conditions. A survey and some open questions. *Probability Surveys* **2**, 107–144. MR2178042 https://doi.org/10.1214/154957805100000104

Brownlees, C. T. and Engle, R. (2017). SRISK: A conditional capital shortfall measure of systemic risk. *The Review of Financial Studies* **30**, 48–79.

Cai, J.-J., Einmahl, J. H. J., de Haan, L. and Zhou, C. (2015). Estimation of the marginal expected shortfall: The mean when a related variable is extreme. *Journal of the Royal Statistical Society, Series B* **77**, 417–442. MR3310533 https://doi.org/10.1111/rssb.12069

Key words and phrases. Extreme expectiles, extreme value theory, GARCH models, tail risk, φ-mixing sequences.

Chen, Z. (1996). Conditional lp-quantiles and their application to testing of symmetry in non-parametric regression. *Statistics & Probability Letters* **29**, 107–115. MR1411408 https://doi.org/10.1016/0167-7152(95)00163-8
Chernozhukov, V. (2005). Extremal quantile regression. *The Annals of Statistics* **33**, 806–839. MR2163160 https://doi.org/10.1214/009053604000001165
Daouia, A., Girard, S. and Stupfler, G. (2018). Estimation of tail risk based on extreme expectiles. *Journal of the Royal Statistical Society, Series B* **80**, 263–292. MR3763692 https://doi.org/10.1111/rssb.12254
Davison, A. C., Padoan, S. A. and Stupfler, G. (2022). Tail risk inference via expectiles in heavy-tailed time series. *Journal of Business and Economic Statistics*, 1–14. MR4600855 https://doi.org/10.1080/07350015.2022.2078332
de Haan, L. and Ferreira, A. (2006). *Extreme Value Theory: An Introduction*. New York: Springer. MR2234156 https://doi.org/10.1007/0-387-34471-3
Delbaen, F., Bellini, F., Bignozzi, V. and Ziegel, J. (2016). Risk measures with the CxLS property. *Finance and Stochastics* **20**, 433–453. MR3479327 https://doi.org/10.1007/s00780-015-0279-6
Dowd, K. (1998). *Beyond Value at Risk, the New Science of Risk Management*. Chichester: Wiley.
Efron, B. (1991). Regression percentiles using asymmetric squared error loss. *Statistica Sinica* **1**, 93–125. MR1101317
Engle, R. F., Jondeau, E. and Rockinger, M. (2015). Systemic risk in Europe. *Review of Finance* **19**, 145–190.
Engle, R. F. and Manganelli, S. (2004). CAViaR: Conditional autoregressive value at risk by regression quantiles. *Journal of Business and Economic Statistics* **22**, 367–381. MR2091566 https://doi.org/10.1198/073500104000000370
Francq, C. and Zakoïan, J. M. (2006). Mixing properties of a general class of GARCH$(1,1)$ models without moment assumptions on the observed process. *Econometric Theory* **22**, 815–834. MR2291218 https://doi.org/10.1017/S0266466606060373
Girard, S., Stupfler, G. and Usseglio-Carleve, A. (2021). Extreme conditional expectile estimation in heavy-tailed heteroscedastic regression models. *The Annals of Statistics* **49**, 3358–3382. MR4352533 https://doi.org/10.1214/21-aos2087
Glosten, L. R., Jagannathan, R. and Runkle, D. E. (1993). On the relation between the expected value and the volatility of nominal excess return on stocks. *The Journal of Finance* **48**, 1779–1801.
Hill, B. M. (1975). A simple general approach to inference about the tail of a distribution. *The Annals of Statistics* **3**, 1163–1174. MR0378204
Jiang, Y., Lin, F. and Zhou, Y. (2021). The kth power expectile regression. *Annals of the Institute of Statistical Mathematics* **73**, 83–113. MR4205243 https://doi.org/10.1007/s10463-019-00738-y
Knight, K. (1999). Epi-convergence in distribution and stochastic equi-semicontinuity. Technical Report, University of Toronto, Toronto.
Koenker, R. and Bassett, G. Jr. (1978). Regression quantiles. *Econometrica* **46**, 33–50. MR0474644 https://doi.org/10.2307/1913643
Kuan, C. M., Yeh, J. H. and Hsu, Y. C. (2009). Assessing value at risk with care, the conditional autoregressive expectile models. *Journal of Econometrics* **150**, 261–270. MR2535521 https://doi.org/10.1016/j.jeconom.2008.12.002
Lin, F., Jiang, Y. and Zhou, Y. (2022). The kth power expectile estimation and testing. *Communications in Mathematics and Statistics*, 1–43. MR4205243 https://doi.org/10.1007/s10463-019-00738-y
Lin, Z. and Lu, C. (1996). *Limit Theory for Mixing Dependent Random Variables*. Science Press/Kluwer Academic Publishers. MR1486580
McNeil, A. J. and Frey, R. (2000). Estimation of tail-related risk measures for heteroscedastic financial time series: An extreme value approach. *Journal of Empirical Finance* **7**, 271–300.
Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica* **59**, 347–370. MR1097532 https://doi.org/10.2307/2938260
Newey, W. K. and Powell, J. L. (1987). Asymmetric least squares estimation and testing. *Econometrica* **55**, 819–847. MR0906565 https://doi.org/10.2307/1911031
Pan, J. and Duffie, D. (1997). An overview of value at risk. *The Journal of Derivatives* **4**, 7–49.
Risk magazine (1996). Value at risk. Risk Magazine Special Supplement, June, 68–71.
Riskmetrics (1995). *Technical Document. Morgan Guarantee Trust Company*. New York: Global Research.
Schnabel, S. K. and Eilers, P. H. C. (2009). Optimal expectile smoothing. *Computational Statistics & Data Analysis* **53**, 4168–4177. MR2744314 https://doi.org/10.1016/j.csda.2009.05.002
Taylor, S. (1986). *Modelling Financial Time Series*. New York: Wiley.
Wang, H. J., Li, D. and He, X. (2012). Estimation of high conditional quantiles for heavy-tailed distributions. *Journal of the American Statistical Association* **107**, 1453–1464. MR3036407 https://doi.org/10.1080/01621459.2012.716382
Weber, S. (2006). Distribution-invariant risk measures, information, and dynamic consistency. *Mathematical Finance* **16**, 419–441. MR2212272 https://doi.org/10.1111/j.1467-9965.2006.00277.x

Xiao, Z. and Koenker, R. (2009). Conditional quantile estimation for generalized autoregressive conditional heteroscedasticity models. *Journal of the American Statistical Association* **104**, 1696–1712. MR2750586 https://doi.org/10.1198/jasa.2009.tm09170

Zakoian, J. M. (1994). Threshold heteroscedastic models. *Journal of Economic Dynamics and Control* **18**, 931–955. MR1462740 https://doi.org/10.2307/20075968

Ziegel, J. F. (2016). Coherency and elicitability. *Mathematical Finance* **26**, 901–918. MR3551510 https://doi.org/10.1111/mafi.12080

Brazilian Journal of Probability and Statistics
2024, Vol. 38, No. 3, 411–434
https://doi.org/10.1214/24-BJPS611

Improved log-Birnbaum-Saunders inference under type II censoring

Thalytta Evilly Cavalcante Silva[a], **Audrey Helen Mariz de Aquino Cysneiros**[b] and **Aline Barbosa Tsuyuguchi**[c]

Departamento de Estatística, Universidade Federal do Pernambuco, Avenida Professor Moraes Rego, 1235 – Cidade Universitária, Recife, 50670-901, Pernambuco, Brazil, [a]*thalytta.cavalcante@ufpe.br,* [b]*audrey.cysneiros@ufpe.br,* [c]*aline.tsuyuguchi@ufpe.br*

Abstract. In this article, we report the development of adjustments to the profile likelihood function based on the proposals of Severini [Severini, T. A., 2000. Likelihood Methods in Statistics] and Cox and Reid [Cox, D.R., Reid, N. 1987. Parameter orthogonality and approximate conditional inference 49, 1-39] in the log Birnbaum-Saunders regression model under both type II censoring and noncensoring. We obtained the profile and modified profile maximum likelihood estimators, calculated the confidence intervals of the asymptotic type and bootstrap percentile and conducted tests based on profile and modified profile likelihood ratio statistics. Through Monte Carlo simulations, we numerically evaluated the behaviors of point estimators under different scenarios, as well as the performance of the likelihood ratio tests based on the profile and modified profile likelihoods, in terms of size and power. The results showed that both the tests and the estimators based on modified versions of the profile likelihood developed had superior performance in small samples in comparison with the estimators based on the unmodified version. Finally, an application to real data is presented.

References

Araújo, M. C., Cysneiros, A. H. M. A. and Montenegro, L. C. (2020). Improved heteroskedasticity likelihood ratio tests in symmetric nonlinear regression models. *Statistical Papers* **61**, 167–188. MR4056797 https://doi.org/10.1007/s00362-017-0933-5

Balakrishnan, N. and Kundu, D. (2018). Birnbaum-Saunders distribution: A review of models, analysis, and applications. *Applied Stochastic Models in Business and Industry* **35**. MR3915800 https://doi.org/10.1002/asmb.2348

Barndorff-Nielsen, O. (1983). On a formula for the distribution of the maximum likelihood estimator. *Biometrika* **70**, 343–365. MR0712023 https://doi.org/10.1093/biomet/70.2.343

Barreto, L. S., Cysneiros, A. H. M. A. and Cribari-Neto, F. (2013). Improved Birnbaum–Saunders inference under type II censoring. *Computational Statistics & Data Analysis* **57**, 68–81. MR2981074 https://doi.org/10.1016/j.csda.2012.06.005

Birnbaum, Z. W. and Saunders, S. C. (1969a). A new family of life distribution. *Journal of Applied Probability* **6**, 319–327. MR0253493 https://doi.org/10.2307/3212003

Cox, D. R. and Reid, N. (1987). Parameter orthogonality and approximate conditional inference. *Journal of the Royal Statistical Society, Series B, Methodological* **49**, 1–39. MR0893334

Cribari-Neto, F. and Fonseca, R. V. A. (2019). A new log-linear bimodal Birnbaum–Saunders regression model with application to survival data. *Brazilian Journal of Probability and Statistics* **33**, 329–355. MR3919026 https://doi.org/10.1214/17-bjps390

Cysneiros, A. H. M. A., Cribari-Neto, F. and Araújo-Júnior, C. A. G. (2008). On Birnbaum–Saunders inference. *Computational Statistics & Data Analysis* **52**, 4939–4950. MR2526204 https://doi.org/10.1016/j.csda.2008.04.003

Da Silva, M. F., Ferrari, S. L. P. and Cribari-Neto, F. (2008). Improved likelihood inference for the shape parameter in Weibull regression. *Journal of Statistical Computation and Simulation* **78**, 789–811. MR2458497 https://doi.org/10.1080/00949650701421964

Key words and phrases. Maximum likelihood estimator, interval estimation, log-Birnbaum-Saunders, likelihood ratio test, profile likelihood.

Doornik, J. A. (2009). *An Object-Oriented Matrix Language Ox 6*. London: Timberlake Consultants Press.
Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *The Annals of Statistics* **7**, 1–26. MR0515681
Farias, R. B. A. and Lemonte, A. J. (2011). Bayesian inference for the Birnbaum–Saunders nonlinear regression model. *Statistical Methods & Applications* **20**, 423–438. MR2859778 https://doi.org/10.1007/s10260-011-0165-0
Fraser, D. A. S. and Reid, N. (1995). Ancillaries and third order significance. *Utilitas Mathematica* **47**, 33–53. MR1330888
Galea, M., Leiva-Sánchez, V. and Paula, G. (2004). Influence diagnostics in log-Birnbaum-Saunders regression models. *Journal of Applied Statistics* **31**, 1049–1064. MR2100116 https://doi.org/10.1080/0266476042000280409
Hayashi, F. (2000). *Econometrics*. Princeton: Princeton University Press. MR1881537
Leiva, V., Barros, M., Paula, G. A. and Galea, M. (2007). Influence diagnostics in log-Birnbaum–Saunders regression models with censored data. *Computational Statistics & Data Analysis* **51**, 5694–5707. MR2407671 https://doi.org/10.1016/j.csda.2006.09.020
Leiva, V., Santos-Neto, M., Cysneiros, F. J. A. and Barros, M. (2014). Birnbaum–Saunders statistical modelling: A new approach. *Statistical Modelling* **14**, 21–48. MR3179546 https://doi.org/10.1177/1471082X13494532
Leiva, V., Vilca, F., Balakrishnan, N. and Sanhueza, A. (2010). A skewed sinh-normal distribution and its properties and application to air pollution. *Communications in Statistics - Theory and Methods* **39**, 426–443. MR2745286 https://doi.org/10.1080/03610920903140171
Lemonte, A. J. (2012). A log-Birnbaum–Saunders regression model with asymmetric errors. *Journal of Statistical Computation and Simulation* **82**, 1775–1787. MR2998066 https://doi.org/10.1080/00949655.2011.595715
Lemonte, A. J. and Cordeiro, G. M. (2009). Birnbaum–Saunders nonlinear regression models. *Computational Statistics & Data Analysis* **53**, 4441–4452. MR2744337 https://doi.org/10.1016/j.csda.2009.06.015
Marchant, C., Leiva, V. and Cysneiros, F. J. A. (2016). A multivariate log-linear model for Birnbaum-Saunders distributions. *IEEE Transactions on Reliability* **65**, 816–827. MR3430824 https://doi.org/10.1016/B978-0-12-803769-0.00001-7
Nocedal, J. and Wright, S. J. (1999). *Numerical Optimization*. New York: Springer. MR1713114 https://doi.org/10.1007/b98874
Oliveira, M. I., Barros, M., Campos, J. and Cysneiros, F. J. A. (2020). Bivariate Birnbaum-Saunders accelerated lifetime model: estimation and diagnostic analysis. *Journal of Applied Statistics*, 1–25. MR4399040 https://doi.org/10.1080/02664763.2020.1859466
Pace, L. and Salvan, A. (1997). *Principles of Statistical Inference from a Neo-Fisherian Perspective*. Singapore: World Scientific. MR1476674
Rieck, J. R. and Nedelman, J. R. (1991). A log-linear model for the Birnbaum-Saunders distribution. *Technometrics* **33**, 51–60.
Severini, T. A. (1999). An empirical adjustment to the likelihood ratio statistic. *Biometrika* **86**, 235–247. MR1705371 https://doi.org/10.1093/biomet/86.2.235
Severini, T. A. (2000). *Likelihood Methods in Statistics*. Oxford: Oxford University Press. MR1854870
Simonoff, J. S. and Tsai, C.-L. (1994). Use of modified profile likelihood for improved tests of constancy of variance in regression. *Journal of the Royal Statistical Society Series C Applied Statistics* **43**, 357–370. MR1281539 https://doi.org/10.2307/2986026
Smyth, G. K. and Verbyla, A. P. (1999). Adjusted likelihood methods for modelling dispersion in generalized linear models. *EnvironMetrics* **10**, 695–709.
Vilca, F., Romeiro, R. G. 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
White, H. (1982). Maximum likelihood estimation of misspecified models. *Econometrica* **50**, 1–25. MR0640163 https://doi.org/10.2307/1912526
Xie, F. C. and Wei, B. C. (2007). Diagnostics analysis for log-Birnbaum–Saunders regression models. *Computational Statistics & Data Analysis* **51**, 4692–4706. MR2361364 https://doi.org/10.1016/j.csda.2006.08.030

Brazilian Journal of Probability and Statistics
2024, Vol. 38, No. 3, 435–443
https://doi.org/10.1214/24-BJPS612

An explicit multiple case-deletion formula for a linear regression model with correlated errors and a resulting property of the BLUP of a multivariate predictand

N. Balakrishnan[1,a] **and Rahul Mukerjee**[2,b]

[1]*Department of Mathematics and Statistics, McMaster University, Hamilton, Ontario, Canada L8S 4K1,*
[a]*bala@mcmaster.ca*
[2]*Indian Institute of Management Calcutta, Joka, Diamond Harbour Road, Kolkata 700104, India,*
[b]*rmuk0902@gmail.com*

Abstract. With reference to a linear model with correlated errors, we obtain an updation formula for the best linear unbiased estimator (BLUE) of the regression coefficients under multiple case-deletion. The generality and clarity of this formula, compared to its existing counterparts, facilitate its use. Specifically, the established formula leads to an attractive property of the best linear unbiased predictor (BLUP) of a multivariate predictand in terms of the invariance of the aforementioned BLUE as well as the residual sum of squares when the BLUP is substituted for actual observations. This property is illustrated with a numerical example on order statistics from a location-scale model.

References

Balakrishnan, N. and Bhattacharya, R. (2022). On simultaneous best linear unbiased prediction of future order statistics and associated properties. *Journal of Multivariate Analysis* **188**, 104854. MR4353863 https://doi.org/10.1016/j.jmva.2021.104854

Balakrishnan, N. and Cohen, A. C. (1991). *Order Statistics and Inference: Estimation Methods*. San Diego: Academic Press. MR1084812

Christensen, R., Pearson, L. M. and Johnson, W. (1992). Case-deletion diagnostics for mixed models. *Technometrics* **34**, 38–45. MR1157792 https://doi.org/10.2307/1269550

Cook, R. D. (1977). Detection of influential observation in linear regression. *Technometrics* **19**, 15–18. MR0436478 https://doi.org/10.2307/1268249

Doganaksoy, N. and Balakrishnan, N. (1997). A useful property of best linear unbiased predictors with applications to life-testing. *American Statistician* **51**, 22–28. MR1440936 https://doi.org/10.2307/2684687

Fung, W. K., Zhu, Z. Y., Wei, B. C. and He, X. (2002). Influence diagnostics and outliers tests for semiparametric mixed models. *Journal of the Royal Statistical Society, Series B* **64**, 565–579. MR1924307 https://doi.org/10.1111/1467-9868.00351

Harter, H. L. and Balakrishnan, N. (1996). *CRC Handbook of Tables for the Use of Order Statistics in Estimation*. Boca Raton, Florida: CRC Press. MR1383819

Hilden-Minton, J. A. (1995). Multilevel Diagnostics for Mixed and Hierarchical Linear Models. PhD Thesis, University of California. Los Angeles. MR2693014

Leone, F. C. and Moussa-Hamouda, E. (1973). Relative efficiencies of 'O-BLUE' estimators in simple linear regression. *Journal of the American Statistical Association* **68**, 953–959. MR0365923

Moussa-Hamouda, E. (1988). Inference in regression problems based on order statistics. *Communications in Statistics—Theory and Methods* **17**, 2343–2367. MR0955116 https://doi.org/10.1080/03610928808829750

Moussa-Hamouda, E. and Leone, F. C. (1974). The O-Blue estimators for complete and censored samples in linear regression. *Technometrics* **16**, 441–446.

Rao, C. R. (1973). *Linear Statistical Inference and Its Applications*, 2nd ed. New York: Wiley. MR0346957

Tan, W. Y. and Tabatabai, M. A. (1988). A modified Winsorized regression procedure for linear models. *Journal of Statistical Computation and Simulation* **30**, 299–313. MR1005890 https://doi.org/10.1080/00949658808811103

Key words and phrases. Best linear unbiased estimator, best linear unbiased predictor, invariance, location-scale model, order statistics.

Yale, C. and Forsythe, A. B. (1976). Winsorized regression. *Technometrics* **18**, 291–300.
Zewotir, T. (2008). Multiple cases deletion diagnostics for linear mixed models. *Communications in Statistics—Theory and Methods* **37**, 1071–1084. MR2712300

Brazilian Journal of Probability and Statistics
2024, Vol. 38, No. 3, 444–462
https://doi.org/10.1214/24-BJPS615

Asymptotic distribution of the friendship paradox of a random geometric graph

Mingao Yuan[a]

Department of Statistics, North Dakota State University, Fargo, North Dakota, USA, [a]*mingao.yuan@ndsu.edu*

Abstract. In social networks, there is a paradox that an individual's friends tend to have more friends than the individual. The friendship paradox is a summary statistic of networks that measures the strength of the paradox. It has many applications in network data analysis. As far as we know, theoretical properties of the friendship paradox are not well studied. In this paper, we derive the asymptotic distribution of the friendship paradox in a random geometric graph. After centering and scaling, the friendship paradox converges in law to a series of independent standard normal random variables. It is interesting that the limiting distribution is not the standard normal distribution.

References

Cantwell, G., Kirkley, A. and Newman, M. E. J. (2021). The friendship paradox in real and model networks. *Journal of Complex Networks* **00**, 1–15. MR4266153 https://doi.org/10.1093/comnet/cnab011

Charitou, T., Bryan, K. and Lynn, D. (2016). Using biological networks to integrate, visualize and analyze genomics data. *Genetics Selection Evolution* **48**, 27.

Chiasserini, C. F., Garetto, M. and Leonardi, E. (2016). Social network de-anonymization under scale-free user relations. *IEEE/ACM Transactions on Networking* **24**, 3756–3769.

Duchemin, Q. and De Castro, Y. (2023). *Random Geometric Graph: Some Recent Developments and Perspectives. High Dimensional Probability IX. Progress in Probability*. Cham: Birkhauser. MR4633270 https://doi.org/10.1007/978-3-031-26979-0_14

Field, S. (1991). Why your friends have more friends than you do. *American Journal of Sociology* **96**, 1464–1477.

Galhotra, S., Mazumdar, A., Pal, S. and Saha, B. (2023). Community recovery in the geometric block model. *Journal of Machine Learning Research* **24**, 1–53. MR4690287

Jackson, M. (2019). The friendship paradox and systematic biases in perceptions and social norms. *Journal of Political Economy* **127**, 777–818.

Nettasinghe, B. and Krishnamurthy, V. (2019). "What do your friends think?": Efficient polling methods for networks using friendship paradox. *IEEE Transactions on Knowledge and Data Engineering* **33**, 1291–1305.

Nettasinghe, B. and Krishnamurthy, V. (2021). Maximum likelihood estimation of power-law degree distributions via friendship paradox-based sampling. *ACM Transactions on Knowledge Discovery from Data* **15**, 106.

Newman, M. (2003). The structure and function of complex networks. *SIAM Review* **45**, 167–256. MR2010377 https://doi.org/10.1137/S003614450342480

O'Malley, A. J. and Marsden, P. V. (2008). The analysis of social networks. *Health Services and Outcomes Research Methodology* **8**, 222–269.

Pal, S., Yu, F., Novick, Y., Swami, A. and Bar-Noy, A. (2019). A study on the friendship paradox–quantitative analysis and relationship with assortative mixing. *Applied Network Science* **4**, 1–26.

Serfling, R. J. (1980). *Approximation Theorems of Mathematical Statistics*. New York: John Wiley& Sons. MR0595165

Simpson, S., Bowman, F. and Laurienti, P. (2013). Analyzing complex functional brain networks: Fusing statistics and network science to understand the brain. *Statistics Surveys* **7**, 1–36. MR3161730 https://doi.org/10.1214/13-SS103

Key words and phrases. The friendship paradox, random geometric graph, asymptotic distribution.

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/

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/