Contents

BJPS is an IMS supported journal

ISSN 0103-0752

Brazilian Statistical Association

Brazilian Journal of Probability and Statistics

Volume 36 • Number 4 • December 2022

Brazilian Journal of Probability and Statistics

Volume 36 • Number 4 • December 2022

www.imstat.org/bjps

ISSN 0103-0752 (Print) ISSN 2317-6199 (Online), Volume 36, Number 4, December 2022. 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).

Copyright © 2022 by the Brazilian Statistical Association.

Printed in the United States of America

Partial financial support:
CNPq and CAPES (Brazil).

Brazilian Journal of Probability and Statistics
2022, Vol. 36, No. 4, 641–674
https://doi.org/10.1214/22-BJPS544
© Brazilian Statistical Association, 2022

New zero-inflated regression models with a variant of censoring

Yasin Altinisik[a] **and Emel Cankaya**[b]

Department of Statistics, Faculty of Science and Literature, Sinop University, Sinop, Turkey,
[a]*yaltinisik@@sinop.edu.tr,* [b]*ecankaya@sinop.edu.tr*

Abstract. There is ever growing demand of modeling overdispersed count data generated by various disiplines. Excessive number of zeros and heterogeneity in the population are two main sources of the overdispersion problem. Development of new count models that are more flexible than conventional Poisson model is thus necessary in order to address such sources. This study fullfils this need by proposing a new heterogeneous Poisson model with a capture of excess zeros, namely zero-inflated Poisson–Ailamujia (ZIPA) model. In line with the aim of curing overdispersion, a censored variant of this newly suggested model is also here developed. An extensive simulation study is conducted to assess the performances of both forms of new models in terms of bias, precision and accuracy measures. Additionally, two real world applications are presented to illustrate practical implications of zero-inflated (censored) Poisson–Ailamujia models in comparison to some alternatives.

References

Altinisik, Y. and Cankaya, E. (2022). Supplement to "New zero-inflated regression models with a variant of censoring." https://doi.org/10.1214/22-BJPS544SUPP

Altun, E. (2018). A new zero-inflated regression model with application. *Journal of Statisticians: Statistics and Actuarial Sciences* **2**, 73–80.

Berk, R. A. and MacDonald, J. M. (2008). Overdispersion and Poisson regression. *Journal of Quantitative Criminology* **24**, 269–284.

Cameron, A. C. and Trivedi, P. K. (1998). *Regression Analysis of Count Data*. Cambridge: University Press. MR1648274 https://doi.org/10.1017/CBO9780511814365

Caudill, S. B. and Mixon, F. G. (1995). Modeling household fertility decisions: Estimation and testing of censored regression models for count data. *Empirical Economics* **20**, 183–196.

Chatla, S. B. and Shmueli, G. (2018). Efficient estimation of COM-Poisson regression and a generalized additive model. *Computational Statistics & Data Analysis* **121**, 71–89. MR3759199 https://doi.org/10.1016/j.csda.2017.11.011

Chen, X. and Lai, Y. (2020). A censored-Poisson model based approach to the analysis of RNA-seq data. *Quantitative Biology* **8**, 155–171.

Daly, F. and Robert, E. G. (2016). The Conway–Maxwell–Poisson distribution: Distributional theory and approximation. *ALEA-Latin American Journal of Probability and Mathematical Statistics* **13**, 635–658. MR3522822 https://doi.org/10.30757/alea.v13-25

Deb, P. and Trivedi, P. K. (1997). Demand for medical care by the elderly: A finite mixture approach. *Journal of Applied Econometrics* **12**, 313–336.

Dunn, P. K. and Smyth, G. K. (1996). Randomized quantile residuals. *Journal of Computational and Graphical Statistics* **5**, 236–244.

Efrymson, M. A. (1960). Multiple regression analysis. In *Mathematical Methods for Digital Computers* (A. Raston and H. S. Wilf, eds.) New York: John Wiley. MR0117923

Feng, C., Li, L. and Sadeghpour, A. (2020). A comparison of residual diagnosis tools for diagnosing regression models for count data. *BMC Medical Research Methodology* **20**, 175.

Ghitany, M. E. and Mutairi, D. K. (2009). Estimation methods for the discrete Poisson–Lindley distribution. *Journal of Statistical Computation and Simulation* **79**, 1–9. MR2655672 https://doi.org/10.1080/00949650701550259

Key words and phrases. Lifetime datasets, zero-inflation, count regression, compound Poisson, overdispersion, censoring, Poisson–Ailamujia model.

Greene, W. H. (1994). Accounting for excess zeros and sample selection in Poisson and negative binomial regression models. NYU Working Paper No. EC-94-10.

Greenwood, M. and Yule, G. U. (1920). An inquiry into the nature of frequency distribution representative of multiple happenings with particular reference to the occurence of multiple attacks of disease or of repeated accidents. *Journal of Royal Statistical Society* **83**, 255–279.

Hartig, F. (2021). DHARMa: Residual diagnostics for hierarchical (multi-level/Mixed) regression models. R package version 0.4.1.

Hassan, A., Shalbaf, G. A., Bilal, S. and Rashid, A. (2020). A new flexible discrete distribution with applications to count data. *Journal of Statistical Theory and Applications* **19**, 102–108.

Henningsen, A. and Toomet, O. (2011). maxLik: A package for maximum likelihood estimation in R. *Computational Statistics* **26**, 443–458. MR2833141 https://doi.org/10.1007/s00180-010-0217-1

Hilbe, J. (2011). *Negative Binomial Regression*. Cambridge: Cambridge University Press. MR2797563 https://doi.org/10.1017/CBO9780511973420

Hocking, R. R. (1976). The analysis and selection of variables in linear regression. *Biometrics* **32**, 1–50. MR0398008 https://doi.org/10.2307/2529336

Hörmann, W., Leydold, J. and Derflinger, G. (2004). *Automatic Nonuniform Random Variate Generation*. Berlin Heidelberg: Springer. MR2143197 https://doi.org/10.1007/978-3-662-05946-3

Huang, A. (2017). Mean-parameterized Conway–Maxwell–Poisson regression models for dispersed counts. *Statistical Modelling* **17**, 1–22. MR3720937 https://doi.org/10.1177/1471082X17697749

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

Kleiber, C. and Zeileis, A. (2008). *Applied Econometrics with R*. New York: Springer.

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

Lee, J. H., Han, G., Fulp, W. J. and Giuliano, A. R. (2012). Analysis of overdispersed count data: Application to the human popillomavirus infection in men (HIM) study. *Epidemiology and Infection* **140**, 1087–1094.

Leydold, J. and Hörmann, W. (2019). Runuran: R interface to the 'UNU.RA' random variate generators. R package version 0.27.

Lv, H. Q., Gao, L. H. and Chen, C. L. (2002). Ailamujia distribution and its application in supportability data analysis. *Journal of Academy of Armored Force Engineering* **16**, 48–52.

Mood, C. (2010). Logistic regression: Why we cannot do what we think we can do, and what we can do about it. *European Sociological Review* **26**, 67–82.

Mouatassim, Y. and Ezzahid, E. H. (2012). Poisson regression and zero-inflated Poisson regression: Application to private health insurance data. *European Actuarial Journal* **2**, 187–204. MR3039549 https://doi.org/10.1007/s13385-012-0056-2

Mullahy, J. (1986). Specification and testing of some modified count data models. *Journal of Econometrics* **33**, 341–365. MR0867980 https://doi.org/10.1016/0304-4076(86)90002-3

Ribeiro, E. E., Zeviani, W. M., Bonato, W. H., Demétrio, C. G. B. and Hind, J. (2020). Reparameterization of COM-Poisson regression models with application in the analysis of experimental data. *Statistical Modelling* **20**, 443–466. MR4150265 https://doi.org/10.1177/1471082X19838651

Saffari, S. E. and Adnan, R. (2011). Zero-inflated Poisson regression models with right censored count data. *Matematika* **27**, 21–29. MR2842673

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

Terza, J. V. (1985). A Tobit-type estimator for the censored Poisson regression model. *Economics Letters* **18**, 361–365.

Walther, B. A. and Moore, J. L. (2015). The concepts of bias, precision and accuracy, and their use in testing the performance of species richness estimators, with a literature review of estimator performance. *Ecography* **28**, 815–829.

Xavier, D., Santos-Neto, M., Bourguignon, M. and Tomazella, V. (2017). Zero-Modified Poisson–Lindley distribution with applications in zero-inflated and zero-deflated count data. arXiv: Methodology.

Yang, J., Xie, M. and Goh, T. N. (2011). Outlier identification and robust parameter estimation in a zero-inflated Poisson model. *Journal of Applied Statistics* **38**, 421–430. MR2749595 https://doi.org/10.1080/02664760903456426

Zeileis, A., Kleiber, C. and Jackman, S. (2008). Regression models for count data in R. *Journal of Statistical Software* **27**.

Brazilian Journal of Probability and Statistics
2022, Vol. 36, No. 4, 675–684
https://doi.org/10.1214/22-BJPS547
© Brazilian Statistical Association, 2022

The law of the iterated logarithm for solutions of stochastic differential equations with random coefficients

Artem Logachov[1,a] **and Olga Logachova**[2,b]

[1]*Laboratory of Probability Theory and Mathematical Statistics, Sobolev Institute of Mathematics, Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia,* [a]*avlogachov@mail.ru*
[2]*Department of High Mathematics, Siberian State University of Geosystems and Technologies, Novosibirsk, Russia,* [b]*omboldovskaya@mail.ru*

Abstract. The functional law of the iterated logarithm is obtained for solutions of stochastic differential equations with random coefficients.

References

Baldi, P. (1986). Large deviations and functional iterated logarithm law for diffusion processes. *Probability Theory and Related Fields* **71**, 435–453. MR0824713 https://doi.org/10.1007/BF01000215
Budkov, D. S. and Makhno, S. Y. (2011). Law of the iterated logarithm for solutions of stochastic equations. *Theory of Probability and Mathematical Statistics* **83**, 47–57. MR2768847 https://doi.org/10.1090/S0094-9000-2012-00840-5
Bulinskii, A. V. (1981). A new variant of the functional law of the iterated logarithm. *Theory of Probability and Its Applications* **25**, 493–503. MR0582580
Caramellino, L. (1998). Strassen's law of the iterated logarithm for diffusion processes for small time. *Stochastic Processes and Their Applications* **74**, 1–19. MR1624072 https://doi.org/10.1016/S0304-4149(97)00100-2
Carfagnini, M., Földes, J. and Herzog, D. P. (2022). A functional law of the iterated logarithm for weakly hypoelliptic diffusions at time zero. *Stochastic Processes and Their Applications* **149**, 188–223. MR4410039 https://doi.org/10.1016/j.spa.2022.03.012
Dembo, A. and Zeitouni, O. (1998). *Large Deviations Techniques and Applications*. New York: Springer. MR1619036 https://doi.org/10.1007/978-1-4612-5320-4
Feng, J. and Kurtz, T. (2006). *Large Deviations for Stochastic Processes*. Providence: American Mathematical Society. MR2260560 https://doi.org/10.1090/surv/131
Freidlin, M. I. and Wentzell, A. D. (1984). *Random Perturbations of Dynamical Systems*. New York: Springer. MR0722136 https://doi.org/10.1007/978-1-4684-0176-9
Gantert, N. (1993). An inversion of Strassen's law of the iterated logarithm for small time. *The Annals of Probability* **21**, 1045–1049. MR1217579
Gikhman, I. I. and Skorokhod, A. V. (2007). *The Theory of Stochastic Processes III*. New York: Springer. MR2304859 https://doi.org/10.1007/978-3-540-49941-1
Kouritzin, M. A. and Heunis, A. J. (1994). An inversion of Strassen's law of the iterated logarithm for small time. *The Annals of Probability* **22**, 659–679. MR1288141
Logachov, A. (2015). The functional law of iterated logarithm for Itô stochastic integrals. *Journal of Mathematical Sciences* **207**, 47–58. MR3373838 https://doi.org/10.1007/s10958-015-2354-0
Makhno, S. Y. (1996). The law of iterated logarithm for solutions of stochastic equations with periodic coefficients. *Mathematical Notes* **59**, 557–559. MR1445459 https://doi.org/10.1007/BF02308825
Mueller, C. (1981). A unification of Strassen's law and Lévy's modulus of continuity. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* **56**, 163–179. MR0618270 https://doi.org/10.1007/BF00535739
Strassen, V. (1964). An invariance principle for the law of the iterated logarithm. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* **3**, 211–226. MR0175194 https://doi.org/10.1007/BF00534910

Key words and phrases. Large deviations principle, law of the iterated logarithm, stochastic differential equations.

Brazilian Journal of Probability and Statistics
2022, Vol. 36, No. 4, 685–691
https://doi.org/10.1214/22-BJPS548
© Brazilian Statistical Association, 2022

Probability solutions of the Sincov's functional equation on the set of nonnegative integers

Nikolai Kolev[1,a] **and Sabrina Mulinacci**[2,b]

[1]*Institute of Mathematics and Statistics, University of São Paulo, São Paulo, Brazil,* [a]*kolev.ime@gmail.com*
[2]*Department of Statistical Sciences, University of Bologna, Bologna, Italy,* [b]*sabrina.mulinacci@unibo.it*

Abstract. In this note, we establish when the bivariate discrete Schur-constant models possess the Sibuya-type aging property. It happens that the corresponding class is large, solving the counterpart of classical Sincov's functional equation on the set of nonnegative integers.

References

Aczél, J. (1966). *Lectures on Functional Equations and Their Applications*. New York: Academic Press, Elsevier. MR0208210

Castañer, A. and Claramunt, M. (2019). Equilibrium distributions and discrete Schur-constant models. *Methodology and Computing in Applied Probability* **21**, 449–459. MR3947876 https://doi.org/10.1007/s11009-018-9632-5

Castañer, A., Claramunt, M., Lefèvre, C. and Loisel, S. (2015). Discrete Schur-constant models. *Journal of Multivariate Analysis* **140**, 343–362. MR3372573 https://doi.org/10.1016/j.jmva.2015.06.003

Kolev, N. (2020). Discrete line integral on uniform grids: Probabilistic interpretation and applications. *Brazilian Journal of Probability and Statistics* **34**, 821–843. MR4153644 https://doi.org/10.1214/19-BJPS454

Kolev, N. and Mulinacci, S. (2022). New characterizations of bivariate discrete Schur-constant models. *Statistics & Probability Letters* **180**, 109233. MR4316066 https://doi.org/10.1016/spl.2021.109233

Lefèvre, C., Loisel, S. and Utev, S. (2018). Markov property in discrete Schur-constant models. *Methodology and Computing in Applied Probability* **20**, 1003–1012. MR3841631 https://doi.org/10.1007/s11009-017-9564-5

Nair, N., Sankaran, P. and Balakrishnan, N. (2018). *Reliability Modelling and Analysis in Discrete Time*. London: Academic Press. MR3793634

Sincov, D. (1903). Notes sur la calcul functionnel. *Bull. Soc. Phys. Math.* **13**, 48–72 (in Russian).

Key words and phrases. Characterization, discrete Schur-constant models, Sibuya-type aging property, Sincov's functional equation.

Brazilian Journal of Probability and Statistics
2022, Vol. 36, No. 4, 692–703
https://doi.org/10.1214/22-BJPS549
© Brazilian Statistical Association, 2022

A survival model for lifetime with long-term survivors and unobserved heterogeneity

Ana Paula J. do Espirito Santo[1,a], **Vicente G. Cancho**[1,b], **Francisco Louzada**[1,c] **and Edwin M. M. Ortega**[2]

[1]*Institute for Mathematical Science and Computing, University of São Paulo, São Carlos, São Paulo, Brazil,* [a]*fanapjesanto@gmail.com,* [b]*louzada@icmc.usp.br,* [c]*garibay@icmc.usp.br*
[2]*ESALQ, University of São Paulo, Piracicaba, São Paulo, Brazil*

Abstract. In this paper, we develop a new survival model induced by discrete frailty with Katz distribution. The new model encompasses as particular cases the mixture cure rate model and promotion cure rate model and has a proportional hazards structure when the covariates are modeled through mean frailty. Furthermore, we construct a regression model to evaluate the effects of covariates on both the cured fraction and risk of the event of interest. We discuss inference aspects of the proposed model in a classical approach, where an expectation maximization algorithm is developed to determine the maximum likelihood estimates of the models parameters. Finally, the model is fully illustrated with a dataset on cervical cancer.

References

Ata, N. and Özel, G. (2013). Survival functions for the frailty models based on the discrete compound Poisson process. *Journal of Statistical Computation and Simulation* **83**, 2105–2116. MR3169290 https://doi.org/10.1080/00949655.2012.679943

Barriga, G. D. C., Cancho, V. G., Garibay, D. V., Cordeiro, G. M. and Ortega, E. M. M. (2018). A new survival model with surviving fraction: An application to colorectal cancer data. *Statistical Methods in Medical Research* 1–16. MR4000187 https://doi.org/10.1177/0962280218786053

Berkson, J. and Gage, R. P. (1952). Survival curve for cancer patients following treatment. *Journal of the American Statistical Association* **47**, 501–515.

Boag, J. W. (1949). Maximum likelihood estimates of the proportion of patients cured by cancer therapy. *Journal of the Royal Statistical Society, Series B, Statistical Methodology* **11**, 15–53.

Cancho, V., Rodrigues, J. and de Castro, M. (2011). A flexible model for survival data with a cure rate: A Bayesian approach. *Journal of Applied Statistics* **38**, 57–70. MR2747683 https://doi.org/10.1080/02664760903254052

Cancho, V. G., Barriga, D. C., Cordeiro, G. M., Ortega, E. M. and Suzuki, A. K. (2021). A flexible model for survival data with a cure rate: A Bayesian approach. *Statistica Neerlandica* **75**, 299–323. MR4018462 https://doi.org/10.1142/S2424922X19500074

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

Caroni, C., Crowder, M. and Kimber, A. (2010). Proportional hazards models with discrete frailty. *Lifetime Data Analysis* **16**, 374–384. MR2657896 https://doi.org/10.1007/s10985-010-9151-3

Cordeiro, G. M., Cancho, V. G., Ortega, E. M. M. and Barriga, G. D. C. (2016). A model with long-term survivors: Negative binomial Birnbaum–Saunders. *Communications in Statistics Theory and Methods* **45**, 1370–1387. MR3462152 https://doi.org/10.1080/03610926.2013.863929

de Souza, D., Cancho, V. G., Rodrigues, J. and Balakrishnan, N. (2017). Bayesian cure rate models induced by frailty in survival analysis. *Statistical Methods in Medical Research* **26**, 2011–2028. MR3712217 https://doi.org/10.1177/0962280217708671

Dempster, A. P., Laird, N. M. and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the royal statistical society. Series B (methodological)*, 1–38. MR0501537

Dunn, P. K. and Smyth, G. K. (1996). Randomized quantile residuals. *Journal of Computational and Graphical Statistics* **5**, 236–244.

Fang, Y. (2003). GMM tests for the Katz family of distributions. *Journal of Statistical Planning and Inference* **110**, 55–73. MR1944633 https://doi.org/10.1016/S0378-3758(01)00284-1

Key words and phrases. Discrete frailty model, cure rate model, proportional hazards, EM algorithm.

Gallardo, D. I., Gomez, Y. M., Gomez, H. W. and de Castro, M. (2020). On the use of the modified power series family of distributions in a cure rate model context. *Statistical Methods in Medical Research* **29**, 1831–1845. MR4114816 https://doi.org/10.1177/0962280219876962

Hougaard, P. (1984). Life table methods for heterogeneous populations: Distributions describing the heterogeneity. *Biometrika* **71**, 75–83. MR0738328 https://doi.org/10.1093/biomet/71.1.75

Ibrahim, J. G., Chen, M. H. and Sinha, D. (2005). *Bayesian Survival Analysis*. New York: Springer. MR1876598 https://doi.org/10.1007/978-1-4757-3447-8

Katz, L. (1965). Unified treatment of a broad class of discrete probability distributions. *Classical and Contagious Discrete Distributions* **1**, 175–182. MR0212914

Lambert, P. C., Thompson, J. R., Weston, C. L. and Dickman, P. W. (2007). Estimating and modeling the cure fraction in population-based cancer survival analysis. *Biostatistics* **8**, 576–594.

McLachlan, G. and Krishnan, T. (2007). *The EM Algorithm and Extensions 382*. New York: John Wiley & Sons. MR2392878 https://doi.org/10.1002/9780470191613

Ortega, E. M. 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

Pal, S. and Balakrishnan, N. (2017). Expectation maximization algorithm for Box–Cox transformation cure rate model and assessment of model misspecification under Weibull lifetimes. *IEEE Journal of Biomedical and Health Informatics* **22**, 926–934.

Press, W. H., Teukolsky, S. A., Vetterling, W. T. and Flannery, B. P. (2007). *Numerical Recipes*, 3rd ed. Cambridge: Cambridge University Press. MR2371990

Rigby, R. A. and Stasinopoulos, D. M. (2005). Generalized additive models for location, scale and shape (with discussion). *Applied Statistics* **54**, 507–554. MR2137253 https://doi.org/10.1111/j.1467-9876.2005.00510.x

Rodrigues, J., Cancho, V. G., de Castro, M. and Louzada-Neto, F. (2009a). On the unification of the long-term survival models. *Statistics & Probability Letters* **79**, 753–759. MR2662300 https://doi.org/10.1016/j.spl.2008.10.029

Rodrigues, J., de Castro, M., Cancho, V. G. and Balakrishnan, N. (2009b). COM–Poisson cure rate survival models and an application to a cutaneous melanoma data. *Journal of Statistical Planning and Inference* **139**, 3605–3611. MR2549108 https://doi.org/10.1016/j.jspi.2009.04.014

Self, S. G. and Liang, K.-Y. (1987). Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. *Journal of the American Statistical Association* **82**, 605–610. MR0898365

Tsodikov, A. D., Ibrahim, J. G. and Yakovlev, A. Y. (2003a). 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

Tsodikov, A. D., Ibrahim, J. G. and Yakovlev, A. Y. (2003b). 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

Vaupel, J. W., Manton, K. G. and Stallard, E. (1979). The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography* **16**, 439–454.

Yakovlev, A., Yu, A. B., Bardou, V. J., Fourquet, A., Hoang, T., Rochefodiere, A. and Tsodikov, A. D. (1993). A simple stochastics model of tumor recurrence an its aplications to data on premenopausal breast cancer. In *Biometrie et Analyse de Donnes Spatio-Temporelles No, Vol. 12, B*, 33–82.

Yakovlev, A. Y. and Tsodikov, A. D. (1996). *Stochastic Models of Tumor Latency and Their Biostatistical Applications*. Singapore: World Scientific.

Brazilian Journal of Probability and Statistics
2022, Vol. 36, No. 4, 704–724
https://doi.org/10.1214/22-BJPS550
© Brazilian Statistical Association, 2022

Single-stage sampling procedure for heteroscedasticity analysis of means

Wei-Ming Wang[1,a], **Chun-Che Wen**[2,b], **Tsai-Hsiang Hung**[3,c], **Junjiang Zhong**[4,d] **and Miin-Jye Wen**[1,5,e]

[1]*Department of Statistics and Institute of Data Science, College of Management, National Cheng Kung University, Tainan, Taiwan,* [a]*weiming3524@gmail.com*

[2]*Department of Public Health Sciences, College of Graduate Studies, Medical University of South Carolina, Charleston, South Carolina, USA,* [b]*chechewen18@gmail.com*

[3]*Digital, Data & Technology, Cathay Financial Holdings Co., Ltd., Taipei, Taiwan,* [c]*hungth18@gmail.com*

[4]*School of Mathematics and Statistics, Xiamen University of Technology, Xiamen, China,* [d]*jjzhong@xmut.edu.cn*

[5]*Institute of International Management, College of Management, National Cheng Kung University, Tainan, Taiwan,* [e]*mjwen@ncku.edu.tw*

Abstract. The analysis of means (ANOM) is a method that can compare the mean of each treatment to the overall mean. According to the graphical result of a statistical data analysis, we can specify which one is different from another. One of the assumptions of the classical ANOM model is that the variances are equal. However, it is not always true for the practice. To solve unknown and unequal population variances, Nelson and Dudewicz (2002) proposed a two-stage sampling procedure. However, additional samples need to be added in the second stage of the two-stage sampling procedure, so it is not practical all the time due to limited time and insufficient budget. Thus, under heteroscedasticity, we applied Chen and Lam's (1989) single-stage sampling procedure to solve the drawback of the two-stage sampling procedure. In addition, we also provided an illustrative example and critical values for practical uses. In order to make the procedure user-friendly, we built an interface by using *R* Shiny.

References

Bishop, T. A. (1976). Heteroscedastic anova, manova, and multiple-comparisons. Ph.D. thesis, The Ohio State University. MR2626711

Brown, M. and Forsythe, A. (1974). Small sample behavior of some statistics which test the equality of several means. *Technometrics* **16**, 129–132. MR0334368 https://doi.org/10.2307/1267501

Chen, H. J. and Lam, K. (1989). Single-stage interval estimation of the largest normal mean under heteroscedasnoty. *Communications in Statistics Theory and Methods* **18**, 3703–3718. MR1040672 https://doi.org/10.1080/03610928908830118

Chen, H. J. and Wen, M. J. (2006). Optimal confidence interval for the largest normal mean under heteroscedasticity. *Computational Statistics & Data Analysis* **51**, 982–1001. MR2297501 https://doi.org/10.1016/j.csda.2005.10.004

Chen, S. Y. (2001). One-stage and two-stage statistical inference under heteroscedasticity. *Communications in Statistics Simulation and Computation* **30**, 991–1009. MR1878465 https://doi.org/10.1081/SAC-100107792

Chen, S. Y. and Chen, H. J. (1998). Single-stage analysis of variance under heteroscedasticity. *Communications in Statistics Simulation and Computation* **27**, 641–666.

Chen, S. Y., Chen, H. J. and Chang, H. F. (2004). A one-stage procedure for testing homogeneity of means against an ordered alternative under unequal variances. *Communications in Statistics Simulation and Computation* **33**, 49–67. MR2044861 https://doi.org/10.1081/SAC-120028433

Dag, O., Dolgun, A. and Konar, N. M. (2021). Onewaytests: An *R* Package for One-Way Tests in Independent Groups Designs. *R*. package version, 2.6.

Dudewicz, E. J. and Bishop, T. A. (1981). Analysis of variance with unequal variances. *Journal of Quality Technology* **13**, 111–114.

Key words and phrases. Heteroscedasticity, single-stage sampling procedure, analysis of means.

Dudewicz, E. J. and Nelson, P. R. (2003). Heteroscedastic analysis of means (HANOM). *American Journal of Mathematical and Management Sciences* **23**, 143–181. MR2017431 https://doi.org/10.1080/01966324.2003.10737608

Games, P. A. and Howell, J. F. (1976). Pairwise multiple comparison procedures with unequal n's and/or variances: A Monte Carlo study. *Journal of Educational Statistics* **1**, 113–125.

Hasler, M. (2016). Heteroscedasticity: Multiple degrees of freedom vs. sandwich estimation. *Statistical Papers* **57**, 55–68. MR3461948 https://doi.org/10.1007/s00362-014-0640-4

Hasler, M. (2019). SimComp: Simultaneous comparisons for multiple endpoints. http://CRAN.R-project.org/package=SimComp, R package version 3.3.

Juneau, P. (2003). Using SAS to perform a single-stage multiple comparison procedure for all pair-wise comparisons in a one-way layout with unequal variances. In *Proceedings of the PharmaSUG 2003 Annual Conference Miami, Florida, USA*, *May 4–7, 2003*.

Krutchkoff, R. G. (1988). One-way fixed effects analysis of variance when the error varianes may be unequal. *Journal of Statistical Computation and Simulation* **30**, 259–271.

Nelson, P. R. (1982). Exact critical points for the analysis of means. *Communications in Statistics Theory and Methods* **11**, 699–709.

Nelson, P. R. (1983). A comparison of sample sizes for the analysis of means and the analysis of variance. *Journal of Quality Technology* **15**, 33–39.

Nelson, P. R. and Dudewicz, E. J. (2002). Exact analysis of means with unequal variances. *Technometrics* **44**, 152–160. MR1951723 https://doi.org/10.1198/004017002317375109

Nelson, P. R., Wludyka, P. S. and Copeland, K. A. (2005). The analysis of means: A graphical method for comparing means, rates, and proportions. *Society for Industrial and Applied Mathematics*, 25. MR2162661 https://doi.org/10.1137/1.9780898718362

Ott, E. R. (1967). Analysis of means—a graphical procedure. *Industrial Quality Control* **24**, 101–109.

Pallmann, P. (2017). ANOM: Analysis of means. http://CRAN.R-project.org/package=ANOM, R package version 0.4.

Rao, C. V. (2005). Analysis of means—a review. *Journal of Quality Technology* **37**, 308–315. MR2065771 https://doi.org/10.1016/S0169-7161(03)23014-5

Stein, C. (1945). A two-sample test for a linear hypothesis whose power is independent of the variance. *The Annals of Mathematical Statistics* **16**, 243–258. MR0013885 https://doi.org/10.1214/aoms/1177731088

Wang, W.-M., Wen, C.-C., Hung, T.-H., Zhong, J. and Wen, M.-J. (2022). Supplement to "Single-stage sampling procedure for heteroscedasticity analysis of means." https://doi.org/10.1214/22-BJPS550SUPP

Weerahandi, S. (1987). Testing regression equality with unequal variances. *Econometrica: Journal of the Econometric Society*, 1211–1215. MR0913368 https://doi.org/10.2307/1911268

Wen, M. J. and Chen, H. J. (1994). Single-stage multiple comparison procedures under heteroscedasticity. *American Journal of Mathematical and Management Sciences* **14**, 1–48. MR1287444 https://doi.org/10.1080/01966324.1994.10737368

Wen, M. J., Huang, L. C. and Zhong, J. (2017). Single-stage sampling procedure of the t best populations under heteroscedasticity. *Communications in Statistics Theory and Methods* **46**, 9265–9273. MR3684567 https://doi.org/10.1080/03610926.2016.1206935

Wen, M. J., Wen, C. C. and Wang, W. M. (2022). Single-stage sampling procedure for heteroscedasticity in multiple comparisons with a control. *Communications in Statistics-Simulation and Computation.* To appear. https://doi.org/10.1080/03610918.2022.2053863

Wludyka, P. S. and Nelson, P. R. (1997). An analysis-of-means-type test for variances from normal populations. *Technometrics* **39**, 274–285. MR1470274 https://doi.org/10.1080/01966324.1997.10737429

Wu, S. F. and Chen, H. J. (1998). Multiple comparisons with the average for normal distributions. *American Journal of Mathematical and Management Sciences* **18**, 193–218. MR1655488 https://doi.org/10.1080/01966324.1998.10737459

Wu, S. F. and Chen, H. J. (2016). An optimal confidence region for the largest and the smallest means from a multivariate normal distribution. *Communications in Statistics Simulation and Computation* **45**, 952–967. MR3473923 https://doi.org/10.1080/03610918.2014.889160

Brazilian Journal of Probability and Statistics
2022, Vol. 36, No. 4, 725–750
https://doi.org/10.1214/22-BJPS552
© Brazilian Statistical Association, 2022

Robust estimation in functional comparative calibration models via maximum Lq-likelihood

Patricia Giménez[1,a], **Lucas Guarracino**[1,b] **and Manuel Galea**[2,c]

[1]*Centro Marplatense de Investigaciones Matemáticas, Universidad Nacional de Mar del Plata, Mar del Plata, Argentina,* [a]*pcgimene@mdp.edu.ar,* [b]*lucasguarracino@gmail.com*
[2]*Departamento de Estadística, Pontificia Universidad Católica de Chile, Santiago, Chile,* [c]*mgalea@mat.uc.cl*

Abstract. A fully parametric estimation procedure is proposed for robust estimation of the structural parameter in a functional comparative calibration model, under normality. The proposed estimator is obtained, based on maximum Lq-likelihood approach, first replacing the incidental parameters by estimators depending on the structural parameter. The estimator, called MLqE, depends on a single distortion parameter q, which controls the balance between robustness and efficiency. If q tends to 1, the maximum likelihood estimator (MLE) is obtained. The estimation procedure can be implemented easily by a simple and fast reweighting algorithm. For applying the method to practical and real-data scenarios, a data-based choice of an appropriate value of q is proposed. Consistency and asymptotic normality is established and the covariance matrix is given. The influence function is derived, to show the local robustness properties. Theoretical properties, ease of implementabiliy and empirical results on simulated and real data show the satisfactory behavior of the MLqE and its vantages over the MLE in presence of observations discordant with the assumed model.

References

Barnett, V. D. (1969). Simultaneous pairwise linear structural relationships. *Biometrics* **25**, 129–142.
Bland, J. M. and Altman, D. G. (1999). Measuring agreement in method comparison studies. *Statistical Methods in Medical Research* **8**, 135–160.
Bolfarine, H. and Galea-Rojas, M. (1995). Comments on "Functional comparative calibration" (by D. Kimura). *Biometrics* **51**, 1579–1580.
Cressie, N. and Read, T. R. C. (1984). Multinomial goodness-of-fit tests. *Journal of the Royal Statistical Society, Series B* **46**, 440–464.
Ferrari, D. and La Vecchia, D. (2012). On robust estimation via pseudo-additive information. *Biometrika* **99**, 238–244.
Ferrari, D. and Yang, Y. (2010). Maximum Lq-likelihood estimation. *The Annals of Statistics* **38**, 753–783.
Fuller, W. A. (1987). *Measurement Error Models*. NY: Wiley.
Galea, M. and de Castro, M. (2017). Robust inference in a linear functional model with replications using the t distribution. *Journal of Multivariate Analysis* **160**, 134–145.
Galea-Rojas, M. (1995). *Calibração Comparativa Estrutural e Functional*. *Tese de Doutorado, IME*. Brasil: Universidade de São Paulo.
Ghosh, A. and Basu, A. (2015). Robust estimation for non-homogeneous data and the selection of the optimal tuning parameter: The density power divergence approach. *Journal of Applied Statistics* **42**, 2056–2072.
Giménez, P. and Bolfarine, H. (1997). Corrected score functions in classical error-in-variables and incidental parameter models. *Australian Journal of Statistics* **39**, 325–344.
Giménez, P. and Bolfarine, H. (2000). Comparing consistent estimators in comparative calibration models. *Journal of Statistical Planning and Inference* **86**, 143–155.
Giménez, P., Guarracino, L. and Galea, M. (2022). Maximum Lq-likelihood estimation in functional measurement error models. *Statistica Sinica* **32**, 1723–1743.
Gleser, L. J. (1981). Estimation in a multivariate "error in variables" regression model: Large sample results. *The Annals of Statistics* **9**, 24–44.
Graybill, F. A. (1983). *Matrices with Applications in Statistics*. *Wadsworth*, 2nd ed.

Key words and phrases. Comparative calibration models, maximum Lq-likelihood, robustness.

Hampel, F. R. (1974). The influence curve and its role in robust estimation. *Journal of the American Statistical Association* **69**, 383–393.

Harville, D. A. (1997). *Matrix Algebra from a Statistician's Perspective*. NY: Springer.

Havrda, J. and Charvat, F. (1967). Quantification method of classification processes: Concept of structural entropy. *Kibernetika* **3**, 30–35.

Hong, C. and Kim, Y. (2001). Automatic selection of the tuning parameter in the minimum density power divergence estimation. *Journal of the Korean Statistical Society* **30**, 453–465.

Huber, P. J. (1983). Minimax aspects of bounded-influence regression (with discussion). *Journal of the American Statistical Association* **78**, 66–80.

Jaech, J. L. (1985). *Statistical Analysis of Measurement Errors*. *Exxon Monographs*. NY: Wiley.

Kimura, D. K. (1992). Functional comparative calibration using an EM algorithm. *Biometrics* **48**, 1263–1271.

La Vecchia, D., Camponovo, L. and Ferrari, D. (2015). Robust heart rate variability analysis by generalized entropy minimization. *Computational Statistics & Data Analysis* **82**, 137–151.

Luong, A. and Mak, T. K. (1991). Robust estimation in a linear relationship model. *Communications in Statistics Theory and Methods* **20**, 721–733.

Mak, T. K. (1982). Estimation in the presence of incidental parameters. *Canadian Journal of Statistics* **10**, 121–132.

Neyman, J. and Scott, E. (1948). Consistent estimates based on partially consistent observations. *Econometrica* **16**, 1–32.

Patefield, W. M. (1977). On the information matrix in the linear functional relationship problem. *Applied Statistics* **26**, 69–70.

Ribeiro, T. K. A. and Ferrari, S. L. P. (2022). Robust estimation in beta regression via maximum Lq-likelihood. Statistical Papers. https://doi.org/10.1007/s00362-022-01320-0

Tsallis, C. (1988). Possible generalization of Boltzmann-Gibbs statistics. *Journal of Statistical Physics* **52**, 479–487.

Vilca-Labra, F. E., Bolfarine, H. and Arellano-Valle, R. B. (1998). Elliptical functional models. *Journal of Multivariate Analysis* **65**, 36–57.

Warwick, J. and Jones, M. C. (2005). Choosing a robustness tuning parameter. *Journal of Statistical Computation and Simulation* **75**, 581–588.

Xu, L., Xiang, S. and Yao, W. (2019). Robust maximum Lq-likelihood estimation of joint mean-covariance models for longitudinal data. *Journal of Multivariate Analysis* **171**, 397–411.

Zamar, R. H. (1985). Orthogonal Regression M-estimators. Ph.D. Thesis, University of Washington, USA.

Brazilian Journal of Probability and Statistics
2022, Vol. 36, No. 4, 751–770
https://doi.org/10.1214/22-BJPS553
© Brazilian Statistical Association, 2022

Three mixed-effects regression models using an extended Weibull with applications on games in differential and integral calculus

Gauss M. Cordeiro[1,a], **Julio Cezar Souza Vasconcelos**[2,b], **Denize Palmito dos Santos**[2,c], **Edwin M. M. Ortega**[2,d] **and Renata Alcarde Sermarini**[2,e]

[1] *Universidade Federal de Pernambuco, Recife, Brazil,* [a]*gausscordeiro@gmail.com*
[2] *Universidade de São Paulo, Piracicaba, Brazil,* [b]*juliocezarvasconcelos@hotmail.com,* [c]*denizepalmito@gmail.com,* [d]*edwin@usp.br,* [e]*ralcarde@usp.br*

Abstract. Three new mixed-effects regressions models using an extended Weibull distribution are defined for repeated measures, and their parameters are estimated by maximum likelihood. Monte Carlo simulations report the accuracy of the maximum likelihood estimators and the distribution of the quantile residuals in these regressions. The usefulness of the proposed regressions is illustrated in differential and integral class from the Exact Sciences Department at the University of São Paulo (Brazil) with the objective of showing a pedagogical alternative of learning diagnostic methodology as a game approach. The results indicate that the questions correctly answered by the students took less time to be solved than those incorrectly answered. In addition, the algebraic application and multiple representation questions has the lowest percentages of correct answers and, in general, the longest time to be solved. So, it is possible to note that the used game approach enables the identification of possible difficult points in a class and provides the teacher with the opportunity of search for different strategies to reduce these difficulties faced by differential and integral calculus students when entering higher education, which often result from basic education.

References

Agustin, M. Z. N. and Agustin, M. A. (2009). Algebra and precalculus skills and performance in first-semester calculus. *International Journal of Case Method Research and Application* **21**, 232–236.

Atkinson, A. C. (1985). *Plots, Transformations and Regression: An Introduction to Graphical Methods of Diagnostic Regression Analysis*. Oxford: Clarendon Press.

Bourguignon, M., Silva, R. B. and Cordeiro, G. M. (2014). The Weibull-G family of probability distributions. *Journal of Data Science* **12**, 53–68. MR3099497

BRASIL. Conselho Nacional de Educação/Ministério da Educação (CNE/MEC). Diretrizes para escolas durante a Pandemia. Available at http://portal.mec.gov.br/component/content/index.php?option=com_content&view=article&id=89051:cne-aprova-diretrizes-para-escolas-durante-a-pandemia. Acessed on August, 29, 2021.

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

Cordeiro, G. M., Nadarajah, S. and Ortega, E. M. M. (2013). General results the beta Weibull distribution. *Journal of Statistical Computation and Simulation* **83**, 1082–1114. MR3169222 https://doi.org/10.1080/00949655.2011.649756

da Fonseca, M. O. D. S., Ramos, N. S., Trevisan, A. L. and Mendes, M. T. (2018). Elementos do Conceito de Limite de Sequências Numéricas Mobilizados em uma Tarefa de Cálculo Diferencial e Integral. *Perspectivas da Educação Matemática* **11**.

Dele-Ajayi, O., Strachan, R., Pickard, A. J. and Sanderson, J. J. (2019). Games for teaching mathematics in Nigeria: What happens to pupils' engagement and traditional classroom dynamics? *IEEE Access* **7**, 53248–53261.

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

Key words and phrases. Leaning diagnostic methodology, Monte Carlo simulations, new Weibull-G family, problems in teaching calculus, repeated measures.

Ferrer, F. P. (2017). The impact of algebra and trigonometry to calculus performance. *Asian Journal of Multidisciplinary Studies* **5**, 1–7.

Fokides, E. (2018). Digital educational games and mathematics. Results of a case study in primary school settings. *Asian Journal of Multidisciplinary Studies* **23**, 851–867.

Gupta, R. C., Gupta, R. D. and Gupta, P. L. (1998). Modeling failure time data by Lehman alternatives. *Communications in Statistics—Theory and Methods* **27**, 887–904. MR1613497 https://doi.org/10.1080/03610929808832134

Henrice, P. (1974). *Discrete Fourier Analysis—Cauchy Integrals—Construction of Conformal Maps—Univalent Functions*. New York: John Wiley and Sons, INC.

Kahoot! Learning games. Make learning awesome. Available at https://kahoot.com/schools-u/. Acessed on August 29, 2021.

Lee, C., Famoye, F. and Olumolade, O. (2007). Beta-Weibull distribution: Some properties and applications to censored data. *Journal of Modern Applied Statistical Methods* **6**, 17. MR2210672

Macêdo, J. A. and Gregor, I. C. S. (2020). Dificuldades nos processos de ensino e de aprendizagem de Cálculo Diferencial e Integral. *Educação Matemática Debate* **4**, 8.

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

Mudholkar, G. S., Srivastava, D. K. and Freimer, M. (1995). The exponentiated Weibull family. *Technometrics* **37**, 436–445.

Nadarajah, S., Cordeiro, G. M. and Ortega, E. M. M. (2012). General results for the Kumaraswamy-G distribution. *Journal of Statistical Computation and Simulation* **82**, 951–979. MR2949501 https://doi.org/10.1080/00949655.2011.562504

Prataviera, F., Loibel, S. M. C., Greco, K. F., Ortega, E. M. M. and Cordeiro, G. M. (2020). Modelling nonproportional hazard for survival data with different systematic components. *Environmental and Ecological Statistics* **27**, 467–489.

Prataviera, F., Vasconcelos, J. C. S., Cordeiro, G. M., Hashimoto, E. M. and Ortega, E. M. M. (2019). The exponentiated power exponential regression model with different regression structures: Application in nursing data. *Journal of Applied Statistics* **46**, 1792–1821. MR3953596 https://doi.org/10.1080/02664763.2019.1572719

Prieto, C. M., Orcos Palma, L., Blázquez Tobías, P. J. and León, F. J. M. (2019). Student assessment of the use of Kahoot in the learning process of science and mathematics. *Education Sciences* **9**, 55.

R Core Team (2021). *R: A Language and Environment for Statistical Computing. Version 4.1.1*. Vienna, Austria: The R Foundation for Statistical Computing. Available at https://www.R-project.org/.

Rasmussen, C., Marrongelle, K. and Borba, M. C. (2014). Research on calculus: What do we know and where do we need to go? *ZDM* **46**, 507–515.

Rodríguez-Fernández, L. (2017). Smartphones y aprendizaje: el uso de Kahoot en el aula universitaria. *Revista Mediterránea de Comunicación* **8**, 181–189.

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. MR2137253 https://doi.org/10.1111/j.1467-9876.2005.00510.x

Stewart, J. (2013). *Cálculo*, volume I. 7a. ediçã. São Paulo: Cengage Learning.

Tahir, M. H. and Nadarajah, S. (2015). Parameter induction in continuous univariate distributions: Well-established G families. *Anais da Academia Brasileira de Ciências* **87**, 539–568. MR3408221 https://doi.org/10.1590/0001-3765201520140299

Tahir, M. H., Zubairz, M., Mansoorx, M., Cordeiro, G. M., Alizadehk, M. and Hamedani, G. G. (2016). A new Weibull-G family of distributions. *Hacettepe Journal of Mathematics and Statistics* **45**, 629–647. MR3526719 https://doi.org/10.15672/hjms.2015579686

UNA-SUS (2020). Organização Mundial de Saúde declara pandemia do novo Coronavírus. Ascom SE/UNA-SUS. Available at https://www.unasus.gov.br/noticia/organizacao-mundial-de-saude-declara-pandemia-de-coronavirus. Acessed on August, 29, 2021.

Vandenbussche, J., Ritter, L. and Scherrer, C. (2018). An incentivized early remediation program in Calculus I. *International Journal of Mathematical Education in Science and Technology* **49**, 1235–1249.

Vasconcelos, J. C. S., Cordeiro, G. M., Ortega, E. M. M. and Rezende, E. M. (2021a). A new regression model for bimodal data and applications in agriculture. *Journal of Applied Statistics* **48**, 349–372. MR4194974 https://doi.org/10.1080/02664763.2020.1723503

Vasconcelos, J. C. S., Cordeiro, G. M., Ortega, E. M. M. and Ribeiro, J. G. (2021b). A regression model for extreme events and the presence of bimodality with application to energy generation data. *IET Renewable Power Generation* **15**, 452–461.

Vila, R., Ferreura, L., Saulo, H., Prataviera, F. and Ortega, E. M. M. (2020). A bimodal gamma distribution: Properties, regression model and applications. *Statistics* **54**, 469–493. MR4100720 https://doi.org/10.1080/02331888.2020.1764560

Wang, A. I. and Tahir, R. (2020). The effect of using Kahoot! for learning—A literature review. *Computers and Education* **149**, 103818.

Ward, M. (1934). The representation of Stirling's numbers and Stirling's polynomials as sums of factorial. *American Journal of Mathematics* **56**, 87–95. MR1507004 https://doi.org/10.2307/2370916

Yuxuan, C., Souza, R. C., Contessoto, A. G. and Amorim, A. R. (2021). Guidelines for the development of educational games to motivate the learning of theoretical concepts in engineering and computing courses. *Computer Applications in Engineering Education* **29**, 1312–1323.

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

Brazilian Journal of Probability and Statistics
2022, Vol. 36, No. 4, 771–776
https://doi.org/10.1214/22-BJPS554
© Brazilian Statistical Association, 2022

Evolution with mass extinction on $\mathbb{T}_d^+$

Carolina Grejo[1,a], **Fábio Lopes**[2,b], **Fábio Machado**[3,c] **and Alejandro Roldán-Correa**[4,d]

[1]*Institute of Mathematics and Computer Science, Universidade de São Paulo, São Carlos, Brazil,* [a]*carolina@ime.usp.br*
[2]*Departamento de Matemática, Universidad Tecnológica Metropolitana, Santiago, Chile,* [b]*f.marcellus@utem.cl*
[3]*Institute of Mathematics and Statistics, Universidade de São Paulo, São Paulo, Brazil,* [c]*fmachado@ime.usp.br*
[4]*Institute of Mathematics, Universidad de Antioquia, Medellín, Colombia,* [d]*alejandro.roldan@udea.edu.co*

Abstract. We propose a stochastic model for evolution through mutation and natural selection of a population that evolves on a $\mathbb{T}_d^+$ tree. We think of this model as a way of describing the evolution fitness landscape of a population. We obtain sharp and distinct conditions on the set of parameters for extinction and survival.

References

Aldous, D. and Krebs, W. B. (1990). The birth-and-assassination process. *Statistics & Probability Letters* **10**, 427–430.

Bordenave, C. (2008). On the birth-and-assassination process, with an application to scotching a rumor in a network. *Electronic Journal of Probabability* **13**, 2014–2030.

Guiol, H., Machado, F. and Schinazi, R. (2011). A stochastic model of evolution. *Markov Processes and Related Fields* **17**, 253–258.

Junior, V., Machado, F. and Roldán-Correa, A. (2016). Dispersion as a survival strategy. *Journal of Statistical Physics* **159**, 937–951.

Kortchemski, I. (2016). Predator-prey dynamics on infinite trees: A branching random walk approach. *Journal of Theoretical Probability* **29**, 1027–1046.

Lanchier, N. (2011). Contact process with destruction of cubes and hyperplanes: Forest fires versus tornadoes. *Journal of Applied Probability* **48**, 352–365.

Machado, F., Roldán-Correa, A. and Schinazi, R. (2017). Colonization and collapse. *ALEA, Lat. Am. J. Probab. Math. Stat.* **14**, 719–731.

Schinazi, R. and Schweinsberg, J. (2008). Spatial and non-spatial stochastic models for immune response. *Markov Processes and Related Fields* **14**, 255–276.

Tsitsiklis, J., Papadimitreou, C. and Humblet, P. (1986). The performance of a precedence based queuing discipline. *Journal of the Association for Computing Machinery* **33**, 593–602.

Key words and phrases. Branching processes, evolution, mutation, fitness.

Brazilian Journal of Probability and Statistics
2022, Vol. 36, No. 4, 777–791
https://doi.org/10.1214/22-BJPS555
© Brazilian Statistical Association, 2022

Orthogonal uniform composite designs for the third-order models

Brenda Mbouamba Yankam[a] **and Abimibola Victoria Oladugba**

Department of Statistics, University of Nigeria, Nsukka, Nigeria, [a]*brenda.yankam.pg03685@unn.edu.ng*

Abstract. The third-order response surface designs are designs used to estimate the parameters of the third-order polynomial model. In this paper, the orthogonal uniform composite designs that combine two-level orthogonal array and four-level uniform design denoted by $OUCD_4$ for estimating the parameters of the third-order models are proposed. The $OUCD_4$ are good space-filling and efficient designs, that uses resolution IV as the initial design, provide more flexible run sizes, have the ability to execute in-depth analysis and perform multiple analysis with different parts of the data for cross-validation.

References

Ahmad, T., Akhtar, M. and Gilmour, S. G. (2012). Multilevel augmented pairs second-order response surface designs and their robustness to missing data. *Communications in Statistics Theory and Methods* **41**, 437–452. MR2874385 https://doi.org/10.1080/03610926.2010.513783

Arshad, H. M., Akhtar, M. and Gilmour, S. G. (2012). Augmented Box-Behnken designs for fitting third-order response surfaces. *Communications in Statistics Theory and Methods* **41**, 4225–4239. MR2989307 https://doi.org/10.1080/03610926.2011.568154

Box, G. E. P. and Wilson, K. B. (1951). On the experimental attainment of optimum conditions. *International Journal of Forecasting* **13**, 1–45. MR0046009 https://doi.org/10.1016/0169-2070(93)90088-5

Chen, X. P., Guo, B., Liu, M. Q. and Wang, X. L. (2017). Robustness of orthogonal-array based composite designs to missing data. *Journal of Statistical Planning and Inference* **194**, 15–24. MR3725785 https://doi.org/10.1016/j.jspi.2017.10.004

Cheng, C. S., Deng, L. Y. and Tang, B. (2002). Generalized minimum aberration and design efficiency for nonregular fractional factorial designs. *Statistica Sinica* **12**, 991–1000. MR1947057

Cornell, J. A. and Montgomery, D. C. (1996). Interaction models as alternatives to low-order polynomials. *Journal of Quality Technology* **28**, 163–176. https://doi.org/10.1080/00224065.1996.11979657

Draper, N. R. (1982). Center points in second-order response surface designs. *Technometrics* **24**, 127–133. MR0655576 https://doi.org/10.2307/1268490

Draper, N. R. and Lin, D. K. (1990). Small response surface designs. *Technometrics* **32**, 187–194. MR1067121 https://doi.org/10.2307/1268862

Dutka, M., Ditaranto, M. and Løvås, T. (2015). Application of a central composite design for the study of NOx emission performance of a low NOx burner. *Energies* **8**, 3606–3627. https://doi.org/10.3390/en8053606

Fang, K. T., Lin, D. K. J., Winker, P. and Zhang, Y. (2000). Uniform design: Theory and application. *Technometrics* **42**, 237–248. MR1801031 https://doi.org/10.2307/1271079

Gichuki, K. T., Joseph, K. and John, M. (2020). The D-, A-, E- and T-optimal values of a second order rotatable design in four dimension constructed using balanced incomplete block designs. *American Journal of Applied Mathematics* **8**, 83–88. https://doi.org/10.11648/j.ajam.20200803.12

Iwundu, M. (2017). Missing observations: The loss in relative A-, D- and G-efficiency. *International Journal of Advanced Mathematical Sciences* **5**, 43. https://doi.org/10.14419/ijams.v5i2.7786

Mee, R. W. (2007). Optimal three-level designs for response surfaces in spherical experimental regions. *Journal of Quality Technology* **39**, 340–354. https://doi.org/10.1080/00224065.2007.11917700

Morris, M. D. (2000). A class of three-level experimental designs for response surface modeling. *Technometrics* **42**, 111–121. https://doi.org/10.1080/00401706.2000.10485990

Myers, R. H., Montgomery, D. C. and Anderson-Cook, C. M. (2009). *Response Surface Methodology: Process and Product Optimization Using Designed Experiments*. New York: Wiley. MR2464113

Key words and phrases. Central composite designs, orthogonal uniform composite designs, D-efficiency, third-order model, uniform design.

Oladugba, A. V. and Yankam, B. M. (2022). Robustness of orthogonal uniform composite designs against outlier. *Communications in Statistics—Simulation and Computation*, 1–16. MR4097121 https://doi.org/10.1177/1471082X18821213

Rady, E. A., Abd El-Monsef, M. M. E. and Seyam, M. M. (2009). Relationships among several optimality criteria. *Interstat Journals* **247**, 1–11. http://interstat.statjournals.net/YEAR/2009/articles/0906001.pdf.

Rashid, F., Akram, M., Akbar, A. and Javed, A. (2016). Some new augmented fractional Box–Behnken designs. *Communications in Statistics Theory and Methods* **46**, 2007–2012. MR3574743 https://doi.org/10.1080/03610926.2015.1032423

Rotich, J. C., Kosgei, M. K. and Kerich, G. K. (2017). Optimal third order rotatable designs constructed from Balanced Incomplete Block Design (BIBD). *Current Journal of Applied Science and Technology* **22**, 1–5. https://doi.org/10.9734/CJAST/2017/34937

Tang, Y., Xu, H. and Lin, D. K. (2012). Uniform fractional factorial designs. *The Annals of Statistics* **40**, 891–907. MR2985937 https://doi.org/10.1214/12-AOS987

Westlake, W. J. (1965). Composite designs based on irregular fractions of factorials. *Biometrics* **21**, 324–336.

Xu, H., Cheng, S. W. and Wu, C. F. J. (2004). Optimal projective three-level designs for factor screening and interaction detection. *Technometrics* **46**, 280–292. MR2082498 https://doi.org/10.1198/004017004000000310

Xu, H., Jaynes, J. and Ding, X. (2014). Combining two-level and three-level orthogonal arrays for factor screening and response surface exploration. *Statistica Sinica* **24**, 269–289. MR3183684 https://doi.org/10.5705/ss.2012.210

Xu, H. and Wu, C. F. J. (2001). Generalized minimum aberration for asymmetrical fractional factorial designs. *The Annals of Statistics* **29**, 1066–1077. MR1869240 https://doi.org/10.1214/aos/1013699993

Yang, Y. (2008). Multiple criteria third-order response surface design and comparison. Master's thesis, Florida State University, Famu-Fsu College of Engineering.

Yankam, B. M. and Oladugba, A. V. (2021). Robustness of orthogonal uniform composite designs against missing data. *Communications in Statistics Theory and Methods* 1–16. https://doi.org/10.1080/03610926.2021.1927095

Yankam, B. M. and Oladugba, A. V. (2022). Augmented orthogonal uniform composite designs for fitting third-order model. *Journal of Statistical Theory and Practice* **16**, 1–14. MR4412560 https://doi.org/10.1007/s42519-022-00260-0

Zhang, A., Li, H., Quan, S. and Yang, Z. (2018). UniDOE: Uniform design of experiments. *R package version* **1**, 1–14.

Zhang, X., Liu, M. and Zhou, Y. (2020). Orthogonal uniform composite designs. *Journal of Statistical Planning and Inference* **206**, 100–110. MR4036698 https://doi.org/10.1016/j.jspi.2019.08.007

Zhang, X. R., Qi, Z. F., Zhou, Y. D. and Yang, F. (2017). Orthogonal-array composite design for the third-order models. *Communications in Statistics Theory and Methods* **47**, 3488–3507. MR3803416 https://doi.org/10.1080/03610926.2017.1359297

Zhou, Y., Fang, K. and Ning, J. (2013). Mixture discrepancy for quasi-random point sets. *Journal of Complexity* **29**, 283–301. MR3062468 https://doi.org/10.1016/j.jco.2012.11.006

Zhou, Y. and Xu, H. (2015). Space-filling properties of good lattice point sets. *Biometrika* **102**, 959–966. MR3431565 https://doi.org/10.1093/biomet/asv044

Zhou, Y. D. and Xu, H. (2016). Composite designs based on orthogonal arrays and definitive screening designs. *Journal of the American Statistical Association* **112**, 1675–1683. MR3750890 https://doi.org/10.1080/01621459.2016.1228535

Brazilian Journal of Probability and Statistics
2022, Vol. 36, No. 4, 792–806
https://doi.org/10.1214/22-BJPS556
© Brazilian Statistical Association, 2022

Partially linear models with *p*-order autoregressive skew-normal errors

Clécio da Silva Ferreira[1,a], **Michel H. Montoril**[2,b] **and Gilberto A. Paula**[3,c]

[1]*Department of Statistics, Federal University of Juiz de Fora, Juiz de Fora, MG, Brazil,* [a]*clecio.ferreira@ufjf.edu.br*

[2]*Department of Statistics, Federal University of São Carlos, São Carlos, SP, Brazil,* [b]*michel@ufscar.br*

[3]*Institute of Mathematics and Statistics, University of São Paulo, São Paulo, SP, Brazil,* [c]*giapaula@ime.usp.br*

Abstract. This paper proposes partially linear models with random errors following p-order autoregressive (AR) with skew-normal errors. The maximum likelihood estimators are derived from the Expectation-Maximization algorithm, which have analytic expressions for the M and E-steps. The estimation of the effective degrees of freedom concerning the nonparametric component are obtained based on a linear smoother. The conditional quantile residuals are used for the construction of simulated confidence bands to assess departures from the error assumptions, as well as autocorrelation and partial autocorrelation graphs are considered to check adequacy of the AR error structure. A simulation study is carried out to evaluate the efficiency of the EM algorithm. Finally, the methodology is illustrated by a real data set on cardiovascular mortality.

References

Arellano-Valle, R. B., Ozan, S., Bolfarine, H. and Lachos, V. H. (2005). Skew-normal measurement error models. *Journal of Multivariate Analysis* **96**, 265–281. MR2204978 https://doi.org/10.1016/j.jmva.2004.11.002

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

Bayes, C. L. and Branco, M. D. (2007). Bayesian inference for the skewness parameter of the scalar skew-normal distribution. *Brazilian Journal of Probability and Statistics* **21**, 141–163. MR2397043

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

Bazrafkan, M., Zare, K., Maleki, M. and Khodadi, Z. (2021). Partially linear models based on heavy-tailed and asymmetrical distributions. *Stochastic Environmental Research and Risk Assessment* **36**, 1436–3259.

Ferreira, C. S., Montoril, M. H. and Paula, G. A. (2022). Supplement to "Partially linear models with p-order autoregressive skew-normal errors." https://doi.org/10.1214/22-BJPS556SUPP

de Boor, C. (2001). *A Practical Guide to Spline*, revised ed. Berlin: Springer. MR1900298

Dempster, A., Laird, N. and Rubin, D. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B* **39**, 1–38. MR0501537

Dunn, P. K. and Smyth, G. K. (1996). Randomized quantile residuals. *Journal of Computational and Graphical Statistics* **5**, 236–244.

Eilers, P. H. C. and Marx, B. D. (1996). Flexible smoothing with B-splines and penalties. *Statistical Science* **11**, 89–121. MR1435485 https://doi.org/10.1214/ss/1038425655

Engle, R. F., Granger, C. and Weiss, A. (1986). Semiparametric estimates of the relation between weather and electricity sales. *Journal of the American Statistical Association* **81**, 310–320.

Ferreira, C. S., Bolfarine, H. and Lachos, V. H. (2011). Skew scale mixtures of normal distributions: Properties and estimation. *Statistical Methodology* **8**, 154–171. MR2769277 https://doi.org/10.1016/j.stamet.2010.09.001

Ferreira, C. S. and Paula, G. A. (2017). Estimation and diagnostic for skew-normal partially linear models. *Journal of Applied Statistics* **44**, 3033–3053. MR3721088 https://doi.org/10.1080/02664763.2016.1267124

Ferreira, C. S., Paula, G. A. and Lana, G. C. (2022). Estimation and diagnostic for partially linear models with first-order autoregressive skew-normal errors. *Computational Statistics* **37**, 445–468. MR4390022 https://doi.org/10.1007/s00180-021-01130-2

Key words and phrases. Autoregressive models, EM-algorithm, penalized smoothing, semiparametric models, skew-normal distribution.

Ferreira, C. S., Vilca, F. and Bolfarine, H. (2018). Diagnostics analysis for skew-normal linear regression models: Applications to a quality of life dataset. *Brazilian Journal of Probability and Statistics* **32**, 525–544. MR3812380 https://doi.org/10.1214/17-BJPS352

Ferreira, G., Castro, L. M., Lachos, V. H. and Dias, R. (2013). Bayesian modeling of autoregressive partial linear models with scale mixture of normal errors. *Journal of Applied Statistics* **40**, 1796–1816. MR3290675 https://doi.org/10.1080/02664763.2013.796349

Figueiredo, C. C., Sandoval, M. C., Bolfarine, H. and Lima, C. (2008). Skew-normal linear calibration: A Bayesian perspective. *Journal of Chemometrics* **22**, 472–480.

Genton, M. G. (2004). *Skew-Elliptical Distributions and Their Applications: A Journey Beyond Normality*. London: Chapman & Hall. MR2155324 https://doi.org/10.1201/9780203492000

Green, P. J. (1987). Penalized likelihood for general semi-parametric regression models. *International Statistical Review* **55**, 245–259. MR0963142 https://doi.org/10.2307/1403404

Green, P. J. (1990). On use of the em algorithm for penalized likelihood estimation. *Journal of the Royal Statistical Society, Series B* **52**, 443–452. MR1086796

Härdle, W., Müller, M., Sperlich, S. and Werwatz, A. (2004). *Nonparametric and Semiparametric Models*. Berlin: Springer. MR2061786 https://doi.org/10.1007/978-3-642-17146-8

Harezlak, J., Ruppert, D. and Wand, M. P. (2018). *Semiparametric Regression with R*. New York: Springer. MR3889001 https://doi.org/10.1007/978-1-4939-8853-2

Harvey, C. R., Liechty, J. C., Liechty, M. and Müller, P. (2010). Portfolio selection with higher moments. *Quantitative Finance* **10**, 469–485. MR2668886 https://doi.org/10.1080/14697681003756877

Hastie, T. and Tibshirani, R. (1990). *Generalized Additive Models*. London: Chapman and Hall. MR1082147

Johnson, N. L., Kotz, S. and Balakrishnan, N. (1994). *Continuous Univariate Distributions, Vol. 1*. New York: John Wiley. MR1326603

Johnson, R. A. and Wichern, D. W. (2007). *Applied Multivariate Statistical Analysis*, 6th ed. Upper Saddle River: Pearson Prentice Hall. MR2372475

Montgomery, D. C., Peck, E. A. and Vining, G. G. (2021). *Introduction to Linear Regression Analysis*, 6th ed. Hoboken: John Wiley & Sons. MR0646067

Opsomer, J. D. and Ruppert, D. (1999). A root-n consistent backfitting estimator for semiparametric additive modeling. *Journal of Computational and Graphical Statistics* **8**, 715–732.

R Core Team (2021). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. http://www.R-project.org/.

Relvas, C. E. M. and Paula, G. A. (2016). Partially linear models with first-order autoregressive symmetric errors. *Statistical Papers* **57**, 795–825. MR3557373 https://doi.org/10.1007/s00362-015-0680-4

Ruppert, D., Wand, M. and Carrol, R. (2003). *Semiparametric Regression*. New York: Cambridge University Press. MR1998720 https://doi.org/10.1017/CBO9780511755453

Sahu, S. K., Dey, D. K. and Branco, M. D. (2003). A new class of multivariate distributions with applications to Bayesian regression models. *Canadian Journal of Statistics* **31**, 129–150. MR2016224 https://doi.org/10.2307/3316064

Shumway, R. H. (1988). *Applied Statistical Time Series Analysis*. Englewood Cliffs: Prentice-Hall.

Wood, S. (2017). *Generalized Additive Models, an Introduction with R*, 2nd ed. Chapman and Hall. MR3726911

You, J. and Zhou, X. (2005). Bootstrap of a semiparametric partially linear model with autoregressive errors. *Statistica Sinica* **15**, 117–133. MR2125723

You, J. H. and Chen, G. (2007). Semiparametric generalized least squares estimation in partially linear regression models with correlated errors. *Journal of Statistical Planning and Inference* **137**, 117–132. MR2292845 https://doi.org/10.1016/j.jspi.2005.10.001

Zheng, S. and Li, D. (2018). Semiparametric time series regression modeling with a diverging number of parameters. *Statistica Neerlandica* **72**, 90–108. MR3790755 https://doi.org/10.1111/stan.12121

Brazilian Journal of Probability and Statistics
2022, Vol. 36, No. 4, 807–820
https://doi.org/10.1214/22-BJPS557
© Brazilian Statistical Association, 2022

Bayesian modeling for a new cure rate model based on the Nielsen distribution

Manoel Santos-Neto[1,a], **Yolanda M. Gómez**[2,b], **Diego I. Gallardo**[3,c] **and Eliardo G. Costa**[4,d]

[1]*Departamento de Estatística, Universidade Federal de Campina Grande, Campina Grande, Brazil,* [a]*manoel.ferreira@professor.ufcg.edu.br*
[2]*Departamento de Medicina, Facultad de Medicina, Universidad de Atacama, Copiapó, Chile,* [b]*yolanda.gomez@uda.cl*
[3]*Departamento de Matemática, Facultad de Ingeniería, Universidad de Atacama, Copiapó, Chile,* [c]*diego.gallardo@uda.cl*
[4]*Departamento de Estatística, Universidade Federal do Rio Grande do Norte, Natal, Brazil,* [d]*eliardo.costa@ufrn.br*

Abstract. In this paper, we proposed a new cure rate model based on the Nielsen distribution. This model has a simple form for the probability generating function, it includes as a particular case the logarithmic distribution and it is a proposal recently discussed in greater detail in the literature, so its application within the context of cure models is very attractive. The model is parameterized directly in the cure rate, facilitating the comparison among other cure rate models in the literature also parameterized in this term. The estimation is approached based on a Bayesian paradigm. A real data set is considered to illustrate the performance of our proposal.

References

Berkson, J. and Gage, R. P. (1952). Survival curve for cancer patients following treatment. *Journal of the American Statistical Association* **47**, 501–515.

Boag, J. W. (1949). Maximum likelihood estimates of the proportion of patients cured by cancer therapy. *Journal of the Royal Statistical Society, Series B* **35**, 15–53.

Cancho, V., Macera, M., Suzuki, A., Louzada, F. and Zavaleta, K. (2020). A new long-term survival model with dispersion induced by discrete frailty. *Lifetime Data Analysis* **26**, 221–244. MR4079665 https://doi.org/10.1007/s10985-019-09472-2

Cancho, V. G., de Castro, M. and Dey, D. (2013a). Long-term survival models with latent activation under a flexible family of distributions. *Brazilian Journal of Probability and Statistics* **27**, 585–600. MR3105045 https://doi.org/10.1214/12-bjps186

Cancho, V. G., de Castro, M. and Rodrigues, J. (2012). A Bayesian analysis of the Conway-Maxwell-Poisson cure rate model. *Statistical Papers* **53**, 165–176. MR2878599 https://doi.org/10.1007/s00362-010-0326-5

Cancho, V. G., Louzada, F. and Ortega, E. M. (2013b). The power series cure rate model: An application to a cutaneous melanoma data. *Communications in Statistics Simulation and Computation* **42**, 586–602. MR3020088 https://doi.org/10.1080/03610918.2011.639971

Castellares, F., Lemonte, A. J. and Santos, M. (2020). On the Nielsen distribution. *Brazilian Journal of Probability and Statistics* **1**, 90–111. MR4058972 https://doi.org/10.1214/18-BJPS414

Chen, M.-H., Ibrahim, J. G. and Sinha, D. (1999). A new Bayesian model for survival data with a surviving fraction. *Journal of the American Statistical Association* **94**, 909–919. MR1723307 https://doi.org/10.2307/2670006

Chen, M.-H., Shao, Q. M. and Ibrahim, J. G. (2000). *Monte Carlo Methods in Bayesian Computation*. New York: Springer. MR1742311 https://doi.org/10.1007/978-1-4612-1276-8

Cordeiro, G. M., Ortega, E. M. and Cunha, D. (2013). The exponentiated generalized class of distributions. *Journal of Data Science* **11**, 1–27. MR2963349

Corless, R. M., Gonnet, G. H., Hare, D. E. G., Jeffrey, D. J. and Knuth, D. E. (1996). On the Lambert *W* function. *Advances in Computational Mathematics* **5**, 329–359. MR1414285 https://doi.org/10.1007/BF02124750

Key words and phrases. Nielsen distribution, long-term survival model, maximum likelihood, reparameterization.

Fox, J. (2018) *Bayesian Estimation of Regression Models: An Appendix to Fox & Weisberg—an R Companion to Applied Regression*, 3rd ed. Last revision: 2018-10-01.

Gallardo, D. and Bourguignon, M. (2021). A note on the nielsen distribution. *Brazilian Journal of Probability and Statistics*, 788–790. MR4350960 https://doi.org/10.1214/21-bjps507

Gallardo, D. I., Castro, M. D. and Gómez, H. (2021). An alternative promotion time cure model with overdispersed number of competing causes: An application to melanoma data. *Mathematics* **15**, 1815.

Gallardo, D. I., Gómez, H. and Bolfarine, H. (2017). A new cure rate model based on the Yule-Simon distribution with application to a melanoma data set. *Journal of Applied Statistics* **44**, 1153–1164. MR3638437 https://doi.org/10.1080/02664763.2016.1194385

Gallardo, D. I., Gómez, Y. and Castro, M. D. (2018). A flexible cure rate model based on the polylogarithm distribution. *Journal of Statistical Computation and Simulation* **88**, 2137–2149. MR3804193 https://doi.org/10.1080/00949655.2018.1451850

Gallardo, D. I., Gómez, Y. G. H. and Castro, M. D. (2020). On the use of the modified power series family of distributions in a cure rate model context. *Statistical Methods in Medical Research* **29**, 1831–1845. MR4114816 https://doi.org/10.1177/0962280219876962

Gelfand, A. E. (1996). Model determination using sampling based methods. In *Markov Chain Monte Carlo in Practice* (W. Gilks, S. Richardson and D. Spiegelhalter, eds.) 145–161. Boca Raton: Chapman & Hall. MR1397969

Gelfand, A. E., Dey, D. K. and Chang, H. (1992). Model determination using predictive distributions with implementation via sampling-based methods (with discussion). In *Bayesian Statistics 4* (J. M. Bernardo, J. O. Berger, A. P. Dawid and A. F. M. Smith, eds.) 147–169. Oxford: Oxford University Press. MR1380275

Gronau, Q. F., Singmann, H. and Wagenmakers, E.-J. (2020). bridgesampling: An R package for estimating normalizing constants. *Journal of Statistical Software* **92**, 1–29.

Leao, J., Bourguignon, M., Gallardo, D., Rocha, R. and Tomazella, V. (2020). A new cure rate model with flexible competing causes with applications to melanoma and transplantation data. *Statistics in Medicine* **24**, 3272–3284. MR4157839 https://doi.org/10.1002/sim.8664

Lehmann, E. (1953). The power of rank tests. *The Annals of Mathematical Statistics* **24**, 23–43. MR0054208 https://doi.org/10.1214/aoms/1177729080

McElreath, R. (2020). *Statistical Rethinking: A Bayesian Course with Examples in R and Stan. Texts in Statistical Science*, 2nd ed. Boca Raton: CRC Press.

R Core Team (2021). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Rodrigues, J., Cancho, V. G., de Castro, M. and Louzada-Neto, F. (2009b). On the unification of the long-term survival models. *Statistics & Probability Letters* **79**, 753–759. MR2662300 https://doi.org/10.1016/j.spl.2008.10.029

Rodrigues, J., de Castro, M., Cancho, V. G. and Balakrishnan, N. (2009a). COM-Poisson cure rate survival model and an application to a cutaneous melanoma data. *Journal of Statistical Planning and Inference* **139**, 3605–3611. MR2549108 https://doi.org/10.1016/j.jspi.2009.04.014

Scheike, T. H. and Martinussen, T. (2006). *Dynamic Regression Models for Survival Data*. NY: Springer. MR2214443

Scheike, T. H. and Zhang, M.-J. (2011). Analyzing competing risk data using the R timereg package. *Journal of Statistical Software* **38**, 1–15.

Sinha, D., Chen, M.-H. and Ghosh, S. (1999). Bayesian analysis and model selection for interval-censored survival data. *Biometrics* **55**, 585–590. MR1705161 https://doi.org/10.1111/j.0006-341X.1999.00585.x

Stan Development Team (2021a). RStan: the R interface to Stan. R package version 2.26.4.

Stan Development Team (2021b). Stan modeling language users guide and reference manual. Version 2.28.

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

Williams, J. and Lagakos, S. (1977). Models for censored survival analysis: Constant-sum and variable-sum models. *Biometrika* **64**, 215–224. MR0655435 https://doi.org/10.1093/biomet/64.2.215

Yakovlev, A. Y. and Tsodikov, A. D. (1996). *Stochastic Models of Tumor Latency and Their Biostatistical Applications*. Singapore: World Scientific.

AIMS AND SCOPE

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

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

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

GENERAL INFORMATION

Journal Homepage: http://www.imstat.org/journals-and-publications/brazilian-journal-of-probability-and-statistics/

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

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

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

Correspondence. Mail concerning membership, subscriptions, nonreceipt claims, copyright permissions, advertising or back issues should be sent to the IMS Dues and Subscription Office, 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:
http://www.imstat.org/individual-membership/

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

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

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