Contents

BJPS is an IMS supported journal

ISSN 0103-0752

Brazilian Statistical Association

Brazilian Journal of Probability and Statistics

Volume 38 • Number 4 • December 2024

Brazilian Journal of Probability and Statistics

Volume 38 • Number 4 • December 2024

www.imstat.org/bjps

ISSN 0103-0752 (Print) ISSN 2317-6199 (Online), Volume 38, Number 4, December 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. 4, 463–478
https://doi.org/10.1214/24-BJPS613

On the restricted maximum likelihood estimation in a two-arm randomized clinical trial for binary endpoints

Marko Raseta[1,a] **and Alina Bazarova**[2,b]

[1]*Department of Molecular Genetics, Erasmus MC, Rotterdam, The Netherlands,* [a]*m.raseta@erasmusmc.nl*
[2]*Jülich Supercomputing Centre, Jülich Research Centre, Jülich, Germany,* [b]*al.bazarova@fz-juelich.de*

Abstract. Optimal design of superiority, noninferiority and equivalence two-arm clinical trials with binary endpoints remains a scientific topic of considerable theoretical and practical interest. In their seminal paper, Farrington and Manning (*Stat. Med.* **9** (1990) 1447–1454) demonstrate, upon systematic simulation, the superiority of the design based on the conditional maximum likelihood estimators (RMLE) to a number of viable alternatives. The approach based on the RMLE is the most popular option among practitioners with more than 700 citations up to date. They offer a closed-form formula for the maximum likelihood estimator, which is claimed to be a unique root to a certain cubic polynomial in a specified parameter-dependent interval. We identify a number of instances where this formula is not well-defined, and thus not universally valid. More specifically still, for infinitely many parameter combinations, the formula suggested in Farrington and Manning (*Stat. Med.* **9** (1990) 1447–1454) involves division by 0 which is an ill-defined concept. In this manuscript, we offer a full, mathematically rigorous theory of RMLE in this set-up, which has not been known to have been published previously. We thereby correct and extend the current results on the topic and provide a formula, which is indeed (universally) valid for all possible parameter combinations.

References

Blackwelder, W. C. (1982). "Proving the null hypothesis" in clinical trials. *Controlled Clinical Trials* **3**, 345–353. https://doi.org/10.1016/0197-2456(82)90024-1

Capending, M., Alberto, C., Feser, J., Mooney, J., et al (2020). Immunogenicity and safety of concurrent or sequential administration of live, attenuated SA 14-14-2 Japanise encephalitis vaccine (CD-JEV) and measles-mumps-rubella vaccine in infants 9-12 months of age in the Phillipines: A non-inferiority Phase 4 randomized clinical trial. *Vaccine: X* **6**, 100074. https://doi.org/10.1016/j.jvacx.2020.100074

Chan, I. (1998). Exact tests of equivalence and efficacy with a non-zero lower bound for comparative studies. *Statistics in Medicine* **17**, 1403–1413. https://doi.org/10.1002/(sici)1097-0258(19980630)17:12<1403::aid-sim834>3.0.co;2-y

Chan, I. (1999). Author's reply. *Statistics in Medicine* **18**, 1735–1737. https://doi.org/10.1002/(SICI)1097-0258(19990715)18:13<1735::AID-SIM148>3.0.CO;2-J

Chen, M., McNulty, A., Avery, A., Whiley, D., et al (2019). Solithromycin versus ceftriaxone plus azithromycin for the treatment of uncomplicated genital gonorrhoea (SOLITAIRE-U): A randomised phase 3 non-inferiority trial. *The Lancet: Infectious Diseases* **19**, 833–842. https://doi.org/10.1016/S1473-3099(19)30116-1

Curran, D., Matthews, S., Rowley, S., Young, J., et al (2019a). Recombinant Zoster vaccine significantly reduces the impact on quality of life caused by herpes zoster in adult autologous hematopoietic stem cell transplant recipients: A randomized placebo-controlled trial (ZOE-HSCT). *Biology of Blood and Marrow Transplantation* **25**, 2474–2481. https://doi.org/10.1016/j.bbmt.2019.07.036

Curran, D., Oostvogels, L., Heineman, T., Matthews, S., et al (2019b). Quality of life impact of an adjuvanted recombinant zoster vaccine in adults aged 50 years and older. *The Journals of Gerantology* **74**, 1231–1238. https://doi.org/10.1093/gerona/gly150

Dunnett, C. W. and Gent, M. (1977). Significance testing to establish equivalence between treatments, with special reference to data in the form of 2X2 tables. *Biometrics* **33**, 593–602. https://doi.org/10.2307/2529457

Key words and phrases. Clinical trial design, closed-form formula, restricted maximum likelihood, statistical theory, two-arm trial.

Fagerland, M. W., Lydersen, S. and Laake, P. (2015). Recommended confidence intervals for two independent binomial proportions. *Statistical Methods in Medical Research* **24**, 224–254. MR3336292 https://doi.org/10.1177/0962280211415469

Farrington, C. and Manning, G. (1990). Test statistics and sample size formulae for comparative binomial trials with null hypothesis of non-zero risk difference or non-utility relative risk. *Statistics in Medicine* **9**, 1447–1454. https://doi.org/10.1002/sim.4780091208

Haas, H., Richard, P., Eymin, C., Fiquet, A., et al (2011). Immunogenicity and safety of intramuscular versus subcutaneous administration of a combined measles, mumps, rubella, and varicella vaccine to children 12 to 18 months of age. *Human Vaccines and Immunotherapeutics* **15**, 778–785. https://doi.org/10.1080/21645515.2018.1549452

Kandasamy, R., Gerung, M., Thorson, S., Yu, L., et al (2019). Comparison of two schedules of two-dose priming with the ten-valent pneumococcal conjugate vaccine in Nepalese children: An open-label, randomised non-inferiority controlled trial. *The Lancet* **19**, 156–164. https://doi.org/10.1016/S1473-3099(18)30568-1

Kang, D., Ahu, J., Park, H., Lee, P., et al (2020). Comparison of optical tomography-guided versus intravascular ultrasound-guided percutaneus coronary intervention: Rationale and design of a randomized, controlled OCTIVUS trial. *American Heart Journal* **228**, 72–80.

Kotloff, K., Shirley, D., Creech, C., Frey, S., et al (2019). Mupirocin for staphylococcus aureus decolonization of infants in neonatal intensive care units. *Pediatrics* **143**, e20181565. https://doi.org/10.1542/peds.2018-1565

Miettinen, O. and Nurminen, M. (1985). Comparative analysis of two rates. *Statistics in Medicine* **4**, 213–216. https://doi.org/10.1002/sim.4780040211

Mokhtari, A., Moghadami, M., Seif, M. and Mirahmadizadeh, A. (2021). Association of routine hepatitis B vaccination and other effective factors with hepatitis B virus infection: 25 years since the introduction of national hepatitis B vaccination in Iran. *Iranian Journal of Medical Sciences* **46**, 93–102. https://doi.org/10.30476/ijms.2019.83112.1199

Morrison, N., Kolluri, R., Vasques, M., Madsen, M., Jones, A. and Gibson, K. (2019). Comparison of cyanoacrylate closure and radiofrequency ablation for the treatment of incompetent great saphenous veins: 36-month outcomes of the VeClose randomized controlled trial. *Phlebology: The Journal of Venous Disease* **34**, 380–390. https://doi.org/10.1177/0268355518810259

Munk, A., Skipka, G. and Stratmann, B. (2019). Testing general hypotheses under binomial sampling: The two sample case—asymptotic theory and exact procedures. *Computational Statistics and Data Analysis* **49**, 723–739. MR2141414 https://doi.org/10.1016/j.csda.2004.05.035

Nikki, A., Pluymaekers, M., Elton, A., Dudink, M., et al (2019). Early or delayed cardioversion in recent-onset atrial fibrillation. *The New England Journal of Medicine* **380**, 1499–1508. https://doi.org/10.1056/NEJMoa1900353

O'Riordan, W., Cardenas, C., Shin, E., Sirbu, A., et al (2019a). Once-daily oral omadacycline versus twice-daily oral linezolid for acute bacterial skin and skin structure infections (OASIS-2): A phase 3, double-blind, multicentre, randomised, controlled, non-inferiority trial. *The Lancet: Infectious Diseases* **19**, 1080–1090. https://doi.org/10.1016/S1473-3099(19)30275-0

O'Riordan, W., Green, S., Overcash, J., Puljiz, I., et al (2019b). Omadacycline for acute bacterial skin and skin-structure infections. *The New England Journal of Medicine* **380**, 528–538. https://doi.org/10.1056/NEJMoa1800170

Roebruck, P. and Kuhn, A. (1995). Comparison of tests and sample size formuale for proving therapeutic equivalence based on the difference of binomial probabilities. *Statistics in Medicine* **14**, 1583–1594. https://doi.org/10.1002/sim.4780141409

Roehmel, J. and Mansmann, U. (1999). Unconditional non-asymptotic one-sided tests for independent binomial proportions when the interest lies in showing non-inferiority and/or superiority. *Biometrical Journal* **41**, 149–170. https://doi.org/10.1002/(SICI)1521-4036(199905)41:2<149::AID-BIMJ149>3.0.CO;2-E

Roehmel, J. and Mansmann, U. (1999). Exact tests of equivalence and efficacy with a non-zero lower bound for comparative studies by I. S. F. Chan. *Statistics in Medicine* **18**, 1734–1735. https://doi.org/10.1002/(SICI)1097-0258(19990715)18:13<1734::AID-SIM201>3.0.CO;2-7

Santner, T. J., Pradhan, V., Senchaudhuri, P., Mehta, C. R. and Tamhane, A. (2007). Small-sample comparisons of confidence intervals for the difference of two independent binomial proportions. *Computational Statistics and Data Analysis* **51**, 5791–5799. MR2407677 https://doi.org/10.1016/j.csda.2006.10.018

Sawano, M., Katsuki, T., Kitai, T., Tamita, K., et al (2020). Beta blockers versus calcium channel blockers for provocation of vasospastic angina after drug-eluting stent implantation: A multicentre prospective randomised trial. *Coronary Artery Disease* **7**, e001406. https://doi.org/10.1136/openhrt-2020-001406

Shelby, S. (1975). *CRC Standard Mathematical Tables and Formulae*. Boca Raton: CRC Press.

Skipka, G., Munk, A. and Freitag, G. (2004). Unconditional exact tests for the difference of binomial probabilities—contrasted and compared. *Computational Statistics and Data Analysis* **47**, 757–773. MR2101550 https://doi.org/10.1016/j.csda.2003.11.013

Stets, R., Popescu, M., Gonong, J., Mitha, I., et al (2019). Omadacycline for community-acquired bacterial pneumonia. *The New England Journal of Medicine* **380**, 517–527. https://doi.org/10.1056/NEJMoa1800201

Yanagawa, T., Tango, T. and Hiejima, Y. (1994). Mantel–Haenszel-type tests for testing equivalence or more than equivalence in comparative clinical trials. *Biometrics* **50**, 859–864. https://doi.org/10.2307/2532800

Brazilian Journal of Probability and Statistics
2024, Vol. 38, No. 4, 479–502
https://doi.org/10.1214/24-BJPS614

Imputation of missing data using Gaussian linear cluster-weighted modeling

Luis Alejandro Masmela-Caita[1,a], **Thaís Paiva-Galletti**[2,b] **and Marcos Oliveira-Prates**[2,c]

[1]*Faculty of Mathematical and Natural Sciences, Universidad Distrital Francisco José de Caldas, Bogotá, Colombia,* [a]*lmasmela@udistrital.edu.co*
[2]*Department of Statistics, Federal University of Minas Gerais, Belo Horizonte—Minas Gerais, Brasil,* [b]*thaispaiva@est.ufmg.br*, [c]*marcosop@est.ufmg.br*

Abstract. The theory of missing data addresses statistical methods for handling cases where some values for the variable of interest are absent. However, much of statistical theory traditionally assumes fully observed data. One approach to manage incomplete data sets is imputation, where missing values are estimated based on certain criteria. In this work, we present a novel imputation methodology for univariate missing data patterns, leveraging fully observed auxiliary variables. The model assumes the nonobserved variable is continuous and that auxiliary variables provide valuable information to enhance imputation accuracy. Within a fully Bayesian framework, our method employs a flexible mixture of multivariate normal distributions to jointly model the response and auxiliary variables. Using the properties of Gaussian cluster-weighted modeling, we develop a predictive model to impute the missing values based on the auxiliary data. We illustrate the performance of our method through simulation studies and real-world data, demonstrating its imputation capacity across various scenarios and comparing it to existing methods in the literature.

References

Ali, M. M., Woo, J. and Nadarajah, S. (2010). Some skew symmetric inverse reflected distributions. *Brazilian Journal of Probability and Statistics* **24**, 1–23. MR2580985 https://doi.org/10.1214/08-BJPS100

Atkinson, A. C. and Riani, M. (2007). Exploratory tools for clustering multivariate data. *Computational Statistics & Data Analysis* **52**, 272–285. MR2409981 https://doi.org/10.1016/j.csda.2006.12.034

Azzalini, A. and Bowman, A. W. (1990). A look at some data on the Old Faithful Geyser. *Journal of the Royal Statistical Society Series C Applied Statistics* **39**, 357–365.

Di Zio, M., Guarnera, U. and Luzi, O. (2007). Imputation through finite Gaussian mixture models. *Computational Statistics & Data Analysis* **51**, 5305–5316. MR2370885 https://doi.org/10.1016/j.csda.2006.10.002

Enders, C. K. (2010). *Applied Missing Data Analysis*. New York: Guilford. MR2597021

Fraley, C. and Raftery, A. E. (2007). Bayesian regularization for normal mixture estimation and model-based clustering. *Journal of Classification* **24**, 155–181. MR2415725 https://doi.org/10.1007/s00357-007-0004-5

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

Gershenfeld, N. A. (1997). Nonlinear inference and Cluster-Weighted Modeling. *Annals of the New York Academy of Sciences* **808**, 18–24.

Ghahramani, Z. and Jordan, M. I. (1995). Learning from incomplete data.

Härdle, W. K., et al (1991). *Smoothing Techniques: With Implementation in S*. Berlin: Springer. MR1140190 https://doi.org/10.1007/978-1-4612-4432-5

Hennig, C. (2000). Identifiablity of models for clusterwise linear regression. *Journal of Classification* **17**, 273–296.

Hennig, C. (2003). Clusters, outliers, and regression: Fixed point clusters. *Journal of Multivariate Analysis* **86**, 183–212. MR1994728 https://doi.org/10.1016/S0047-259X(02)00020-9

Key words and phrases. Cluster-weighted modeling, Gaussian mixture models, imputation method, missing data.

Hershey, J. R. and Olsen, P. A. (2007). Approximating the Kullback Leibler divergence between Gaussian mixture models. In *2007 IEEE International Conference on Acoustics, Speech and Signal Processing-ICASSP'07 4 IV–317*. Los Alamitos: IEEE.

Hoshikawa, T. (2013). Mixture regression for observational data, with application to functional regression models. arXiv preprint. Available at arXiv:1307.0170.

Ingrassia, S., Minotti, S. C. and Vittadini, G. (2012). Local statistical modeling via a cluster-weighted approach with elliptical distributions. *Journal of Classification* **29**, 363–401. MR2972778 https://doi.org/10.1007/s00357-012-9114-3

Kim, H. J., Cox, L. H., Karr, A. F., Reiter, J. P. and Wang, Q. (2015). Simultaneous edit-imputation for continuous microdata. *Journal of the American Statistical Association* **110**, 987–999. MR3420678 https://doi.org/10.1080/01621459.2015.1040881

Kim, H. J., Reiter, J. P., Wang, Q., Cox, L. H. and Karr, A. F. (2014). Multiple imputation of missing or faulty values under linear constraints. *Journal of Business & Economic Statistics* **32**, 375–386. MR3238592 https://doi.org/10.1080/07350015.2014.885435

Kottas, A. and Fellingham, G. W. (2012). Bayesian semiparametric modeling and inference with mixtures of symmetric distributions. *Statistics and Computing* **22**, 93–106. MR2865058 https://doi.org/10.1007/s11222-010-9208-x

Kullback, S. and Leibler, R. A. (1951). On information and sufficiency. *The Annals of Mathematical Statistics* **22**, 79–86. MR0039968 https://doi.org/10.1214/aoms/1177729694

Liao, X., Li, H. and Carin, L. (2007). Quadratically gated mixture of experts for incomplete data classification. In *Proceedings of the 24th International Conference on Machine Learning*, 553–560.

Little, R. J. (1988). Missing-data adjustments in large surveys. *Journal of Business & Economic Statistics* **6**, 287–296.

Little, R. J. and Rubin, D. B. (2019). *Statistical Analysis with Missing Data 793*. New York: Wiley. MR1925014 https://doi.org/10.1002/9781119013563

Manrique-Vallier, D. and Reiter, J. P. (2017). Bayesian simultaneous edit and imputation for multivariate categorical data. *Journal of the American Statistical Association* **112**, 1708–1719. MR3750893 https://doi.org/10.1080/01621459.2016.1231612

Masmela-Caita, L. A. (2021). Imputation of Missing Data Using Gaussian Linear Cluster-Weighted Modeling. PhD thesis, Universidad Federal de Minas Gerais.

Masmela-Caita, L. A., Paiva, T. V. and Prates, M. O. (2023). Imputation of missing data using multivariate Gaussian Linear Cluster-Weighted Modeling. arXiv preprint. Available at arXiv:2308.06677.

Masmela-Caita, L. A., Paiva-Galletti, T. and Oliveira-Prates, M. (2025). Supplement to "Imputation of missing data using Gaussian linear cluster-weighted modeling." https://doi.org/10.1214/24-BJPS614SUPP

McCaw, Z. R., Aschard, H. and Julienne, H. (2022). Fitting Gaussian mixture models on incomplete data. *BMC Bioinformatics* **23**, 208. https://doi.org/10.1186/s12859-022-04740-9

Paiva, T. V. (2014). Multiple Imputation Methods for Nonignorable Nonresponse, Adaptive Survey Design, and Dissemination of Synthetic Geographies. PhD thesis, Duke University. MR3301744

Paiva, T. V. and Reiter, J. P. (2017). Stop or continue data collection: A nonignorable missing data approach for continuous variables. *Journal of Official Statistics* **33**, 579–599.

Plummer, M., Best, N., Cowles, K. and Vines, K. (2006). CODA: Convergence diagnosis and output analysis for MCMC. *R News* **6**, 7–11.

Prates, M. O., Barbosa, C. R. and Lachos, V. H. (2013). mixsmsn: Fitting finite mixture of scale mixture of skew-normal distributions. *Journal of Statistical Software* **54**, 1–20.

R Core Team (2020). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Rubin, D. B. (1987). *Multiple Imputation for Nonresponse in Surveys 81*. New York: Wiley. MR0899519 https://doi.org/10.1002/9780470316696

Schafer, J. L. (1997). *Analysis of Incomplete Multivariate Data*. Boca Raton: CRC Press. MR1692799 https://doi.org/10.1201/9781439821862

Sethuraman, J. (1994). A constructive definition of Dirichlet priors. *Statistica Sinica* **4**, 639–650. MR1309433

Sovilj, D., Eirola, E., Miche, Y., Björk, K.-M., Nian, R., Akusok, A. and Lendasse, A. (2016). Extreme learning machine for missing data using multiple imputations. *Neurocomputing* **174**, 220–231.

Tanner, M. A. and Wong, W. H. (1987). The calculation of posterior distributions by data augmentation. *Journal of the American Statistical Association* **82**, 528–540. MR0898357

Van Buuren, S. (2018). *Flexible Imputation of Missing Data*. Boca Raton: CRC Press.

Van Buuren, S. and Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software* **45**, 1–67.

Varin, C. and Vidoni, P. (2005). A note on composite likelihood inference and model selection. *Biometrika* **92**, 519–528. MR2202643 https://doi.org/10.1093/biomet/92.3.519

Wedel, M. (2002). Concomitant variables in finite mixture models. *Statistica Neerlandica* **56**, 362–375. MR2028029 https://doi.org/10.1111/1467-9574.t01-1-00072

Wei, Y., Tang, Y. and McNicholas, P. D. (2019). Mixtures of generalized hyperbolic distributions and mixtures of skew-t distributions for model-based clustering with incomplete data. *Computational Statistics & Data Analysis* **130**, 18–41. MR3860527 https://doi.org/10.1016/j.csda.2018.08.016

Brazilian Journal of Probability and Statistics
2024, Vol. 38, No. 4, 503–530
https://doi.org/10.1214/24-BJPS616

A Bayesian approach for the G-DINA model

Renato Fernandes[1,a], **Jorge Luis Bazán**[2,b] **and Mariana Cúri**[2,c]

[1]*Instituto Federal de Educação, Ciência e Tecnologia de São Paulo, Campus Araraquara, Araraquara, Brazil,* [a]*renato.fernandes@ifsp.edu.br*

[2]*Universidade de São Paulo, Instituto de Ciências Matemáticas e de Computação, São Carlos, Brazil,* [b]*jlbazan@icmc.usp.br,* [c]*mcuri@icmc.usp.br*

Abstract. The G-DINA model is a versatile cognitive diagnostic model used for individuals' classification. In this paper, we propose a Bayesian formulation of the G-DINA model and an estimation method using an MCMC Gibbs Sampling algorithm, implemented using the software JAGS. A simulation study was designed to evaluate the parameters recovery and the estimation accuracy of the Bayesian implementation of G-DINA and the results were compared with the standard frequentist approach in four scenarios. The results show that the proposed Bayesian implementation recovers all parameters and has good accuracy in the estimation, with performance similar to or better than the frequentist approach in all simulated scenarios. As an application, we propose a new methodology of classification of depression for respondents to the Beck Depression Inventory (BDI) based on an existing one, by replacing the original DINA model for the G-DINA. A comparative analysis of the application of these two methods in a data set from 1111 respondents of the BDI test was conducted. The results indicate that the methodology with the G-DINA model is more suitable for this application.

References

Aryadoust, V. (2021). A cognitive diagnostic assessment study of the listening test of the Singapore–Cambridge general certificate of education O level: Application of DINA, DINO, G-DINA, HO-DINA, and RRUM. *International Journal of Listening* **35**, 29–52.

Brooks, S. P. and Gelman, A. (1998). General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics* **7**, 434–455. MR1665662 https://doi.org/10.2307/1390675

Chen, H. and Chen, J. (2016). Exploring reading comprehension skill relationships through the G-DINA model. *Educational Psychology* **36**, 1049–1064.

Chen, J. and De la Torre, J. (2018). Introducing the general polytomous diagnosis modeling framework. *Frontiers in Psychology* **9**, 1474.

Chen, Y.-H., Senk, S. L., Thompson, D. R. and Voogt, K. (2019). Examining psychometric properties and level classification of the van Hiele geometry test using CTT and CDM frameworks. *Journal of Educational Measurement* **56**, 733–756.

Culpepper, S. A. (2015). Bayesian estimation of the DINA model with Gibbs sampling. *Journal of Educational and Behavioral Statistics* **40**, 454–476.

Culpepper, S. A. (2019). Estimating the cognitive diagnosis Q matrix with expert knowledge: Application to the fraction-subtraction dataset. *Psychometrika* **84**, 333–357. MR3947367 https://doi.org/10.1007/s11336-018-9643-8

da Silva, M. A., de Oliveira, E. S. B., von Davier, A. A. and Bazán, J. L. (2018). Estimating the DINA model parameters using the no-U-turn sampler. *Biometrical Journal* **60**, 352–368. MR3773539 https://doi.org/10.1002/bimj.201600225

de la Torre, J. (2009). DINA model and parameter estimation: A didactic. *Journal of Educational and Behavioral Statistics* **34**, 115–130.

de la Torre, J. (2011). The generalized DINA model framework. *Psychometrika* **76**, 179–199. MR2788881 https://doi.org/10.1007/s11336-011-9207-7

de la Torre, J., van der Ark, L. A. and Rossi, G. (2018). Analysis of clinical data from a cognitive diagnosis modeling framework. *Measurement and Evaluation in Counseling and Development* **51**, 281–296.

Key words and phrases. Bayesian models, BDI, cognitive diagnostic models, G-DINA, latent variable models.

Fernandes, R., Bazán, J. L. and Cúri, M. (2025). Supplement to “A Bayesian approach for the G-DINA model.” https://doi.org/10.1214/24-BJPS616SUPPA, https://doi.org/10.1214/24-BJPS616SUPPB

Fragoso, T. M. and Cúri, M. (2013). Improving psychometric assessment of the Beck depression inventory using multidimensional item response theory. *Biometrical Journal* **55**, 527–540. MR3079988 https://doi.org/10.1002/bimj.201200197

Gelman, A. and Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science* **7**, 457–472. https://doi.org/10.1214/ss/1177011136

George, A. C. and Robitzsch, A. (2015). Cognitive diagnosis models in R: A didactic. *The Quantitative Methods for Psychology* **11**, 189–205.

Henson, R. A., Templin, J. L. and Willse, J. T. (2009). Defining a family of cognitive diagnosis models using log-linear models with latent variables. *Psychometrika* **74**, 191–210. MR2507377 https://doi.org/10.1007/s11336-008-9089-5

Hong, C.-Y., Chang, Y.-W. and Tsai, R.-C. (2016). Estimation of generalized DINA model with order restrictions. *Journal of Classification* **33**, 460–484. MR3575625 https://doi.org/10.1007/s00357-016-9215-5

Hwang, J., Choi, K. M. and Hand, B. (2020). Examining domain-general use of reasoning across science and mathematics through performance on standardized assessments. *Canadian Journal of Science, Mathematics and Technology Education* **20**, 521–537.

Junker, B. W. and Sijtsma, K. (2001). Cognitive assessment models with few assumptions, and connections with nonparametric item response theory. *Applied Psychological Measurement* **25**, 258–272. MR1842982 https://doi.org/10.1177/01466210122032064

Koedsri, A. and Ngudgratoke, S. (2018). Diagnostic assessment of scientific literacy of lower secondary school students using G-DINA model. In *Re-Thinking Teacher Professional Education: Using Research Findings for Better Learning-61st World Assembly ICET 2017*, 268–273.

Lei, P.-W. and Li, H. (2016). Performance of fit indices in choosing correct cognitive diagnostic models and Q-matrices. *Applied Psychological Measurement* **40**, 405–417.

Li, H., Hunter, C. V. and Lei, P.-W. (2016). The selection of cognitive diagnostic models for a reading comprehension test. *Language Testing* **33**, 391–409.

Ma, W. and de la Torre, J. (2020). GDINA: An R package for cognitive diagnosis modeling. *Journal of Statistical Software* **93**, 1–26.

Plummer, M., et al. (2016). rjags: Bayesian graphical models using MCMC. R package version 4.

Ravand, H. (2016). Application of a cognitive diagnostic model to a high-stakes reading comprehension test. *Journal of Psychoeducational Assessment* **34**, 782–799.

Ravand, H. and Robitzsch, A. (2018). Cognitive diagnostic model of best choice: A study of reading comprehension. *Educational Psychology* **38**, 1255–1277.

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

Templin, J. L. and Henson, R. A. (2006). Measurement of psychological disorders using cognitive diagnosis models. *Psychological Methods* **11**, 287.

Vehtari, A., Gelman, A. and Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing* **27**, 1413–1432. MR3647105 https://doi.org/10.1007/s11222-016-9696-4

Vehtari, A., Gelman, A., Simpson, D., Carpenter, B. and Bürkner, P.-C. (2021). Rank-normalization, folding, and localization: An improved $\widehat{R}$ for assessing convergence of MCMC (with discussion). *Bayesian Analysis* **16**, 667–718. MR4298989 https://doi.org/10.1214/20-ba1221

Wang, W., Song, L., Chen, P., Meng, Y. and Ding, S. (2015). Attribute-level and pattern-level classification consistency and accuracy indices for cognitive diagnostic assessment. *Journal of Educational Measurement* **52**, 457–476.

Yamaguchi, K. and Okada, K. (2018). Comparison among cognitive diagnostic models for the TIMSS 2007 fourth grade mathematics assessment. *PLoS ONE* **13**, e0188691.

Zhan, P., Jiao, H., Man, K. and Wang, L. (2019). Using JAGS for Bayesian cognitive diagnosis modeling: A tutorial. *Journal of Educational and Behavioral Statistics* **44**, 473–503.

Brazilian Journal of Probability and Statistics
2024, Vol. 38, No. 4, 531–548
https://doi.org/10.1214/24-BJPS617

The Bayes estimator of a conditional density: Asymptotic behavior

Agustín G. Nogales[1,a]

[1]*Department of Mathematics, University of Extremadura, Badajoz, Spain,* [a]*nogales@unex.es*

Abstract. The problem of estimating a density, conditional or not, from both a frequentist and a Bayesian point of view, is of particular interest. These problems have been addressed by the author in previous works in a Bayesian context. Here, under mild regularity assumptions on the measurable spaces involved, the convergence to 0 of the Bayes risk (for both the L^1 and L^1-squared loss functions) of the Bayes estimator of a conditional density is proved; its strong consistency is also derived. The Bayes estimator of the conditional density is the conditional density with respect to the posterior predictive distribution. A closely related problem is that of estimating the conditional distribution itself; the asymptotic behavior of its Bayes estimator is also addressed, and a Bayesian nonparametric example is provided to show a situation where estimating the conditional distribution has an optimal solution, although the problem of estimating the conditional density is meaningless.

The results are general enough to cover discrete and continuous, parametric or nonparametric cases; and no specific supposition is made about the prior distribution.

References

Efromovich, S. (2007). Conditional density estimation in a regression setting. *The Annals of Statistics* **35**, 250–2535. MR2382656 https://doi.org/10.1214/009053607000000253

Florens, J. P., Mouchart, M. and Rolin, J. M. (1990). *Elements of Bayesian Statistics*. Basel: Marcel Dekker. MR1051656

Geisser, S. (1993). *Predictive Inference: An Introduction*. Berlin: Springer. MR1252174 https://doi.org/10.1007/978-1-4899-4467-2

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A. and Rubin, D. B. (2014). *Bayesian Data Analysis*, 3rd ed. Boca Raton: CRC Press. MR3235677

Ghosal, S. and van der Vaart, A. (2017). *Fundamentals of Nonparametric Bayesian Inference*. Cambridge UK: Cambridge University Press. MR3587782 https://doi.org/10.1017/9781139029834

Ghosh, J. K. and Ramamoorthy, R. V. (2003). *Bayesian Nonparametrics*. New York: Springer. MR1992245

Izbicki, R. and Lee, A. B. (2016). Nonparametric conditional density estimation in a high setting. *Journal of Computational and Graphical Statistics* **25**, 1297–1316. MR3572041 https://doi.org/10.1080/10618600.2015.1094393

Nogales, A. G. (2022a). On Bayesian estimation of densities and sampling distributions: The posterior predictive distribution as the Bayes estimator. *Statistica Neerlandica* **76**, 236–250. MR4423267 https://doi.org/10.1111/stan.12258

Nogales, A. G. (2022b). Optimal Bayesian estimation of a regression curve, a conditional density, and a conditional distribution. *Mathematics* **10**, 1213.

Nogales, A. G. (2022c). On consistency of the Bayes estimator of the density. *Mathematics* **10**, 636.

Nogales, A. G. (2022d). Asymptotic Behavior of the Bayes Estimator of a Regression Curve. Preprint.

Rosenblatt, M. (1969). Conditional probability density and regression estimators. In *Multivariate Analysis II* (P. R. Krishnaiah, ed.) 25–31. New York: Academic Press. MR0254987

Key words and phrases. Bayesian estimation of a conditional density, posterior predictive distribution, consistency.

Brazilian Journal of Probability and Statistics
2024, Vol. 38, No. 4, 549–561
https://doi.org/10.1214/24-BJPS618

Variance swaps pricing under the regime-switching jump-diffusion model

Anqi Zou[1,a], **Jiajie Wang**[*2,b] **and Chiye Wu**[2,c]

[1]*School of Information Science and Technology, Jinan University, Guangzhou, China,* [a]*zouanqi@stu2020.jnu.edu.cn*

[2]*Shenzhen Campus, Jinan University, Shenzhen, China,* [b]*wang_jj@sz.jnu.edu.cn,* [c]*chyewu@jnu.edu.cn*

Abstract. This paper reformulates a model for pricing continuous-sampling times variance swaps, which is the mean-reverting Gaussian model with asymmetric double exponential distribution jump. We extend it by further considering the regime-switching frame. In other words, the parameters in this model are modulated by a Markov chain. Note that the market considered is incomplete in general, which implies that more than one equivalent martingale measures exist. Utilizing differential methods such as the regime-switching Esscher transform, we identify an equivalent martingale measure called the minimal entropy martingale measure (MEMM) and derive an analytic solution for the pricing problem.

References

Andrusiv, A. and Engelbert, H.-J. (2020). On the minimal entropy martingale measure for Lévy processes. *Stochastics* **92**, 1223–1243. MR4175845 https://doi.org/10.1080/17442508.2019.1693571

Bo, L., Wang, Y. and Yang, X. (2010). Markov-modulated jump–diffusions for currency option pricing. *Insurance: Mathematics and Economics* **46**, 461–469. MR2642523 https://doi.org/10.1016/j.insmatheco.2010.01.003

Bo, W. and Ruili, S. (2013). The MEMMs for Markov-modulated GBMs. *Chinese Journal of Applied Probability and Statistics* **29**, 183–196. MR3098730

Carr, P., Geman, H., Madan, D. B. and Yor, M. (2005). Pricing options on realized variance. *Finance and Stochastics* **9**, 453–475. MR2213777 https://doi.org/10.1007/s00780-005-0155-x

Deng, G. (2007). Pricing European option in a double exponential jump-diffusion model with two market structure risks and its comparisons. *Applied Mathematics—A Journal of Chinese Universities* **22**, 127–137. MR2323844 https://doi.org/10.1007/s11766-007-0201-x

Elliott, R. J., Chan, L. and Siu, T. K. (2005). Option pricing and Esscher transform under regime switching. *Annals of Finance* **1**, 423–432.

Elliott, R. J., Siu, T. K. and Chan, L. (2007). Pricing volatility swaps under Heston's stochastic volatility model with regime switching. *Applied Mathematical Finance* **14**, 41–62. MR2321122 https://doi.org/10.1080/13504860600659222

Frittelli, M. (2000). The minimal entropy martingale measure and the valuation problem in incomplete markets. *Mathematical Finance* **10**, 39–52. MR1743972 https://doi.org/10.1111/1467-9965.00079

Gerber, H. U. and Shiu, E. S. W. (1994). Option pricing by Esscher transforms (with discussion). *Transactions—Society of Actuaries* **46**, 99–191. MR2211908 https://doi.org/10.1080/10920277.2005.10596232

Grandits, P. (1999). The p-optimal martingale measure and its asymptotic relation with the minimal-entropy martingale measure. *Bernoulli* **5**, 225–247. MR1681696 https://doi.org/10.2307/3318433

Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies* **6**, 327–343. MR3929676 https://doi.org/10.1093/rfs/6.2.327

Heston, S. L. and Nandi, S. (2000). Derivatives on volatility: some simple solutions based on observables. *Federal Reserve Bank of Atlanta WP* **2000-20**.

Howison, S., Rafailidis, A. and Rasmussen, H. (2004). On the pricing and hedging of volatility derivatives. *Applied Mathematical Finance* **11**, 317–346.

Huang, J., Zhu, W. and Ruan, X. (2014). Option pricing using the fast Fourier transform under the double exponential jump model with stochastic volatility and stochastic intensity. *Journal of Computational and Applied Mathematics* **263**, 152–159. MR3162343 https://doi.org/10.1016/j.cam.2013.12.009

Key words and phrases. Variance swaps, mean-reverting Gaussian model, regime switching Esscher transform, double exponential jump-diffusion, equivalent martingale measure.

Kou, S. G. (2002). A jump-diffusion model for option pricing. *Management Science* **48**, 1086–1101.

Liao, J., Shu, H. and Wei, C. (2017). Pricing power options with a generalized jump diffusion. *Communications in Statistics—Theory and Methods* **46**, 11026–11046. MR3717851 https://doi.org/10.1080/03610926.2016.1257138

Merton, R. C. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics* **3**, 125–144.

Protter, P. (1990). *Stochastic Integration and Differential Equations*. Berlin: Springer. MR1037262 https://doi.org/10.1007/978-3-662-02619-9

Rheinlander, T. (2005). An entropy approach to the Stein and Stein model with correlation. *Finance and Stochastics* **9**, 399–413. MR2211715 https://doi.org/10.1007/s00780-004-0149-0

Sepp, A. (2008). Pricing options on realized variance in the Heston model with jumps in returns and volatility. *Journal of Computational Finance* **11**, 33–70.

Shreve, S. E. (2004). *Stochastic Calculus for Finance II: Continuous-Time Models, Vol. 11*, 210–220. New York: Springer. MR2057928

Stübinger, J. and Endres, S. (2018). Pairs trading with a mean-reverting jump–diffusion model on high-frequency data. *Quantitative Finance* **18**, 1735–1751. MR3853698 https://doi.org/10.1080/14697688.2017.1417624

Tankov, P. (2003). *Financial Modelling with Jump Processes*. London: Chapman and Hall/CRC.

Vasicek, O. (1977). An equilibrium characterization of the term structure. *Journal of Financial Economics* **5**, 177–188.

Wu, H., Jia, Z., Yang, S. and Liu, C. (2022). Pricing variance swaps under double Heston stochastic volatility model with stochastic interest rate. *Probability in the Engineering and Informational Sciences* **36**, 564–580. MR4404154 https://doi.org/10.1017/S0269964820000662

Yang, H. (2014). Esscher transform. *Wiley StatsRef: Statistics Reference Online*.

Yang, Y., Liu, S., Wu, Y. and Wiwatanapataphee, B. (2021). Pricing of volatility derivatives in a Heston–CIR model with Markov-modulated jump diffusion. *Journal of Computational and Applied Mathematics* **393**, 113–277. MR4233389 https://doi.org/10.1016/j.cam.2020.113277

Zhang, L.-W. (2014). A closed-form pricing formula for variance swaps with mean-reverting Gaussian volatility. *ANZIAM Journal* **55**, 362–382. MR3257304 https://doi.org/10.1017/S144618111400011X

Zhu, S.-P. and Lian, G.-H. (2015). Analytically pricing volatility swaps under stochastic volatility. *Journal of Computational and Applied Mathematics* **288**, 332–340. MR3349626 https://doi.org/10.1016/j.cam.2015.04.036

Brazilian Journal of Probability and Statistics
2024, Vol. 38, No. 4, 562–575
https://doi.org/10.1214/24-BJPS619

On the directed likelihood statistic in simplex regressions

Artur J. Lemonte[1,a]

[1]*Departamento de Estatística, Universidade Federal do Rio Grande do Norte, Natal/RN, Brazil,*
[a]*arturlemonte@gmail.com*

Abstract. We address the issue of performing testing inference in small-sized samples in the class of simplex regression models, which is tailored to continuous response variables that assume values in the standard unit interval, such as rates and proportions. We consider the directed likelihood test statistic and derive a simple adjustment factor to this statistic to improve its performance in small samples. We compare the finite sample performances of the (usual and adjusted) directed likelihood test statistics using Monte Carlo simulation experiments. The numerical results favor the adjusted directed likelihood test we propose. We also provide an empirical application for illustrative purposes.

References

Atkinson, A. C. (1985). *Plots, Transformation and Regression: An Introduction to Graphical Methods of Diagnostic Regression Analysis*. New York: Oxford University Press.

Barndorff-Nielsen, O. E. (1986). Inference on full and partial parameters based on the standardized signed log likelihood ratio. *Biometrika* **73**, 307–322. MR0855891 https://doi.org/10.2307/2336207

Barndorff-Nielsen, O. E. (1991). Modified signed log likelihood ratio. *Biometrika* **78**, 557–563. MR1130923 https://doi.org/10.1093/biomet/78.3.557

Barndorff-Nielsen, O. E. and Cox, D. R. (1994). *Inference and Asymptotics*. London: Chapman and Hall. MR1317097 https://doi.org/10.1007/978-1-4899-3210-5

Barndorff-Nielsen, O. E. and Jørgensen, B. (1991). Some parametric models on the simplex. *Journal of Multivariate Analysis* **39**, 106–116. MR1128675 https://doi.org/10.1016/0047-259X(91)90008-P

Bayer, F. M. and Cribari-Neto, F. (2013). Bartlett corrections in beta regression models. *Journal of Statistical Planning and Inference* **143**, 531–547. MR2995113 https://doi.org/10.1016/j.jspi.2012.08.018

Carrasco, J. M. F. and Reid, N. (2021). Simplex regression models with measurement error. *Communications in Statistics Simulation and Computation* **50**, 3420–3435. MR4336346 https://doi.org/10.1080/03610918.2019.1626881

Cavalcante, J. P. L. (2023). The simplex regression model as a methodology of actuarial analysis. *Contemporary Journal of Economics and Management* **21**, e83379.

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 Zeileis, A. (2010). Beta regression in R. *Journal of Statistical Software* **34**, 1–24.

Dean, C. B. and Lundy, E. R. (2016). *Overdispersion*. Wiley StatsRef: Statistics Reference Online. USA: John Wiley & Sons.

Espinheira, P. L. and Silva, A. O. (2020). Residual and influence analysis to a general class of simplex regression. *Test* **29**, 523–552. MR4095039 https://doi.org/10.1007/s11749-019-00665-3

Ferrari, S. L. P. and Cribari-Neto, F. (2004). Beta regression for modelling rates and proportions. *Journal of Applied Statistics* **31**, 799–815. MR2095753 https://doi.org/10.1080/0266476042000214501

Jørgensen, B. (1997). *The Theory of Dispersion Models*. London: Chapman and Hall. MR1462891

Kutner, M. H., Nachtsheim, C. J., Neter, J. and Li, W. (2005). *Applied Linear Statistical Models*, 5th ed. Boston: McGraw-Hill Irwin.

Liu, P., Yuen, K. C., Wu, L. C., Tian, G. L. and Li, T. (2020). Zero-one-inflated simplex regression models for the analysis of continuous proportion data. *Statistics and its Interface* **13**, 193–208. MR4063156 https://doi.org/10.4310/SII.2020.v13.n2.a5

Key words and phrases. Proportions, rates, signed likelihood ratio statistic, simplex distribution, small sample asymptotics.

Loose, L. H., Bayer, F. M. and Pereira, T. L. (2017). Bootstrap Bartlett correction in inflated beta regression. *Communications in Statistics Simulation and Computation* **46**, 2865–2879. MR3640110 https://doi.org/10.1080/03610918.2015.1065326

McCullagh, P. and Nelder, J. A. (1989). *Generalized Linear Models*, 2nd ed. Chapman & Hall/CRC. MR3223057 https://doi.org/10.1007/978-1-4899-3242-6

Qiu, Z., Song, P. X. K. and Tan, M. (2008). Simplex mixed-effects models for longitudinal proportional data. *Scandinavian Journal of Statistics* **35**, 577–596. MR2468863 https://doi.org/10.1111/j.1467-9469.2008.00603.x

R Core Team (2024). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Rauber, C., Cribari-Neto, F. and Bayer, F. M. (2020). Improved testing inferences for beta regressions with parametric mean link function. *AStA Advances in Statistical Analysis* **104**, 687–717. MR4173400 https://doi.org/10.1007/s10182-020-00376-3

Santos, J. S. and Cribari-Neto, F. (2023). Improved likelihood ratio tests for varying dispersion simplex regression models. *Advanced Theory and Simulations* **6**, e2300315.

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*. New York: Oxford University Press. MR1854870

Silva, A. O., Silva, J. W. A. and Espinheira, P. L. (2022). Bootstrap-based inferential improvements to the simplex nonlinear regression model. *PLoS ONE* **17**, e0272512.

Song, P. X. K., Qiu, Z. and Tan, M. (2004). Modeling heterogeneous dispersion in marginal simplex models for continuous longitudinal proportional data. *Biometrical Journal* **46**, 540–553. MR2101142 https://doi.org/10.1002/bimj.200110052

Song, P. X. K. and Tan, M. (2000). Marginal models for longitudinal continuous proportional data. *Biometrics* **56**, 496–502.

Yi, G. Y., Wu, J. and Liu, Y. (2002). Implementation of higher-order asymptotics to S-plus. *Computational Statistics & Data Analysis* **40**, 775–800. MR1933009 https://doi.org/10.1016/S0167-9473(02)00096-8

Zhang, P., Qiu, Z. and Shi, C. (2016). simplexreg: An R package for regression analysis of proportional data using the simplex distribution. *Journal of Statistical Software* **71**, 1–21.

Zhang, W. Z. and Wei, H. J. (2008). Maximum likelihood estimation of simplex distribution nonlinear mixed models via the stochastic approximation algorithm. *The Rocky Mountain Journal of Mathematics* **38**, 1863–1875. MR2457391 https://doi.org/10.1216/RMJ-2008-38-5-1863

Zhao, Y. Y., Xu, D. K., Duan, X. D. and Dai, L. (2014). Bayesian estimation of simplex distribution nonlinear mixed models for longitudinal data. *International Journal of Applied Mathematics & Statistics* **52**, 1–10. MR3224062

Brazilian Journal of Probability and Statistics
2024, Vol. 38, No. 4, 576–604
https://doi.org/10.1214/24-BJPS620

Exponential squared loss-Lasso: A sparse noise-resilient partially linear spatial autoregressive models

Qinghui Zeng, Rui Yang, Yue Yin and Yunquan Song[a]

[1]*College of Science, China University of Petroleum, Qingdao 266580, P.R. China,* [a]*syqfly1980@163.com*

Abstract. In spatial autoregressive models, the relationship between response variable and covariable is sometimes not only linear, but also nonlinear. Therefore, a partially linear spatial autoregressive model has recently been proposed for such data with complex correlations. In this paper, we consider the estimation and variable selection problems for partially linear spatial autoregressive models. A robust variable selection method based on spline estimation and exponential squared loss is offered. Under some mild conditions, we establish the asymptotic and Oracle properties of the proposed estimator. We propose a block coordinate descent (BCD) algorithm based on the concave–convex process (CCCP) to solve the problems. Numerical studies have shown that the method is more robust and accurate when the observations are noisy or the estimated spatial weight matrix is inaccurate.

References

Beck, A. and Teboulle, M. (2009). A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM Journal on Imaging Sciences* **2**, 183–202. MR2486527 https://doi.org/10.1137/080716542

Bunea, F. and Wegkamp, M. H. (2004). Two-stage model selection procedures in partially linear regression. *Canadian Journal of Statistics* **32**. MR2064395 https://doi.org/10.2307/3315936

Cai, Z. and Xu, X. (2009). Nonparametric quantile estimations for dynamic smooth coefficient models. *Publications of the American Statistical Association* **104**, 371–383. MR2504383 https://doi.org/10.1198/jasa.2009.0102

Cheng, S. and Chen, J. (2019). Estimation of partially linear single-index spatial autoregressive model. *Statistical Papers* **62**, 495–531. MR4220889 https://doi.org/10.1007/s00362-019-01105-y

Du, J., Sun, X., Cao, R. and Zhang, Z. (2018). Statistical inference for partially linear additive spatial autoregressive models. In *Spatial Statistics*. MR3809255 https://doi.org/10.1016/j.spasta.2018.04.008

Fan, J. and Li, R. (2001). Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American Statistical Association* **96**, 1348–1360. MR1946581 https://doi.org/10.1198/016214501753382273

Forsythe, G. E., Malcolm, M. A. and Moler, C. B. (1977). Computer methods for mathematical computations. MR0458783

Harrison, D. and Rubinfeld, D. L. (1978). Hedonic housing prices and the demand for clean air. *Journal of Environmental Economics and Management* **5**, 81–102.

Koch, M. and Krisztin, T. (2011). Applications for asynchronous multi-agent teams in nonlinear applied spatial econometrics. *Journal of Internet Technology* **12**, 1007–1014.

Krisztin, T. (2018). Semi-parametric spatial autoregressive models in freight generation modeling. *Transportation Research Part E: Logistics and Transportation Review.*

Li, T. and Guo, Y. (2020). Penalized profile quasi-maximum likelihood method of partially linear spatial autoregressive model. *Journal of Statistical Computation and Simulation* **90**, 2705–2740. MR4154862 https://doi.org/10.1080/00949655.2020.1788561

Liu, X., Chen, J. and Cheng, S. (2018). A penalized quasi-maximum likelihood method for variable selection in the spatial autoregressive model. In *Spatial Statistics*. MR3809257 https://doi.org/10.1016/j.spasta.2018.05.001

Ni, X., Zhang, H. H. and Zhang, D. (2009). Automatic model selection for partially linear models. *Journal of Multivariate Analysis* **100**, 2100–2111. MR2543089 https://doi.org/10.1016/j.jmva.2009.06.009

Key words and phrases. Partially linear spatial autoregressive models, exponential squared loss, robust variable selection.

Robinson, P. M. (2006). Efficient estimation of the semiparametric spatial autoregressive model. *Journal of Econometrics* **157**, 6–17. MR2652275 https://doi.org/10.1016/j.jeconom.2009.10.031

Song, Y., Liang, X., Zhu, Y. and Lin, L. (2021). Robust variable selection with exponential squared loss for the spatial autoregressive model. *Computational Statistics & Data Analysis* **155**, 107094. MR4161786 https://doi.org/10.1016/j.csda.2020.107094

Su, L. (2012). Semiparametric gmm estimation of spatial autoregressive models. *Journal of Econometrics* **167**, 543–560. MR2892094 https://doi.org/10.1016/j.jeconom.2011.09.034

Su, L. and Jin, S. (2010). Profile quasi-maximum likelihood estimation of partially linear spatial autoregressive models. *Journal of Econometrics* **157**, 18–33. MR2652276 https://doi.org/10.1016/j.jeconom.2009.10.033

Wang, X., Jiang, Y., Huang, M. and Zhang, H. (2013). Robust variable selection with exponential squared loss. *Journal of the American Statistical Association* **108**, 632–643. MR3174647 https://doi.org/10.1080/01621459.2013.766613

Yu, P., Zhu, Z. and Zhang, Z. (2018). Robust exponential squared loss-based estimation in semi-functional linear regression models. *Computational Statistics* **34**, 503–525. MR3953670 https://doi.org/10.1007/s00180-018-0810-2

Zhang, Y. and Shen, D. (2015). Estimation of semi-parametric varying-coefficient spatial panel data models with random-effects. *Journal of Statistical Planning and Inference* **159**, 64–80. MR3299089 https://doi.org/10.1016/j.jspi.2014.11.001

Zhao, P., Gan, H., Cheng, S. and Zhou, X. (2021). Orthogonality based penalized gmm estimation for variable selection in partially linear spatial autoregressive models. *Communications in Statistics—Theory and Methods* **52**, 1676–1691. MR4549307 https://doi.org/10.1080/03610926.2021.1937652

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/