## Contents

**Special issue in Honor of Heleno Bolfarine**

Preface to the Special Issue ................................................................. 247

**J. M. F. CARRASCO, S. L. P. FERRARI and R. B. ARELLANO–VALLE**
Multiplicative errors-in-variables beta regression .................................. 249

**E. GÓMEZ–DÉNIZ, O. VENEGAS and H. W. GÓMEZ**
Beyond the lognormal distribution with properties and applications .......... 263

**L. BENITES, C. B. ZELLER, H. BOLFARINE and V. H. LACHOS**
Regression modeling of censored data based on compound scale mixtures of normal distributions ......................................................... 282

**M. ZEVALLOS**
On the asymptotic distribution of sample autocovariance differences of long-memory processes ......................................................... 313

**S. PAREEK, K. DAS and S. MUKHOPADHYAY**
Likelihood-based missing data analysis in crossover trials ....................... 329

**X. ZHANG and H. SHU**
Maximum likelihood estimation for the reflected stochastic linear system with a large signal ................................................................. 351

**A. DE LA CRUZ HUAYANAY, J. L. BAZÁN and C. A. RIBEIRO DINIZ**
Longitudinal binary response models using alternative links for medical data ................................................................. 365

**G. MARTÍNEZ-FLÓREZ, M. PACHECO and A. J. LEMONTE**
Influence diagnostics for the power-normal Tobit model ......................... 393

**J. GAMAIN, D. A. C. MOLLINEDO and C. A. TUDOR**
High-dimensional regime for Wishart matrices based on the increments of the solution to the stochastic heat equation ............................. 412

**B. N. B. DE LIMA, V. SIDORAVICIUS and M. E. VARES**
Dependent percolation on $\mathbb{Z}^2$ ....................................................... 431

**S. I. LÓPEZ and L. P. R. PIMENTEL**
On the two-point function of the one-dimensional KPZ equation ............... 455
Preface to the Special Issue
Multiplicative errors-in-variables beta regression

Jalmar M. F. Carrasco¹,a, Silvia L. P. Ferrari²,b and Reinaldo B. Arellano–Valle³,c

¹Department of Statistics, Federal University of Bahia, Brazil, ²carrasco.jalmar@ufba.br
²Department of Statistics, University of São Paulo, Brazil, ³silviaferrari@usp.br

Abstract. This paper deals with beta regression models with a covariate that is not directly observed; instead, it is replaced by a surrogate covariate that underpredicts its actual value. We propose a multiplicative errors-in-variables model tailored for this situation and develop calibration regression and pseudo-likelihood-based inference for the unknown parameters. The impact of ignoring the measurement error and the performance of the inference methods are evaluated through simulations and a real data illustration.

References


Key words and phrases. Beta regression, maximum pseudo-likelihood, multiplicative measurement error, product of independent beta random variables, regression calibration.


Beyond the lognormal distribution with properties and applications

Emilio Gómez–Déniz\textsuperscript{1,a}, Osvaldo Venegas\textsuperscript{2,b} and Héctor W. Gómez\textsuperscript{3,c}

\textsuperscript{1}Department of Quantitative Methods in Economics and TIDES Institute, University of Las Palmas de Gran Canaria, 35017 Las Palmas de Gran Canaria, Spain, \textsuperscript{a}emilio.gomez-deniz@ulpgc.es

\textsuperscript{2}Departamento de Ciencias Matemáticas y Físicas, Facultad de Ingeniería, Universidad Católica de Temuco, Temuco, Chile, \textsuperscript{b}ovenegas@uct.cl

\textsuperscript{3}Departamento de Matemáticas, Facultad de Ciencias Básicas, Universidad de Antofagasta, Antofagasta, Chile, \textsuperscript{c}hector.gomez@uantof.cl

Abstract. In this paper, a new family of continuous random variables with positive support is introduced. Its density function has the capacity to incorporate features of unimodality and bimodality. Special attention is paid to the lognormal distribution which is included as a particular case. Its density function is given in closed-form, allowing probabilities, moments and other related measures such as skewness and kurtosis coefficients to be computed easily. In addition, a stochastic representation of the family that enables us to generate random variates of this model is also presented. Some properties related with the right tail and actuarial aspects of the distribution are also shown. This new family of distributions is numerically illustrated with data taken from the Medical Expenditure Panel Survey (MEPS), conducted by the US Agency of Health Research and Quality and with a well-known data set which has been studied widely in the actuarial literature.

References


Key words and phrases. Bimodal distribution, folded normal distribution, hyperbolic function, lognormal distribution.


Regression modeling of censored data based on compound scale mixtures of normal distributions

Luis Benites\textsuperscript{1,a}, Camila B. Zeller\textsuperscript{2,b}, Heleno Bolfarine\textsuperscript{3,c} and Víctor H. Lachos\textsuperscript{4,d}

\textsuperscript{1}\textit{Departamento de Ciencias, Pontificia Universidad Católica del Perú, Lima, Perú, a}lbenitess@pucp.edu.pe
\textsuperscript{2}\textit{Departamento de Estatística, Universidade Federal de Juiz de Fora, Brazil, b}camila.zeller@uffj.edu.br
\textsuperscript{3}\textit{Departamento de Estatística, Universidade de São Paulo, Butanta, São Paulo, Brazil, c}hbolfar@ime.usp.br
\textsuperscript{4}\textit{Department of Statistics, University of Connecticut, Storrs, Connecticut, U.S.A., d}hlachos@uconn.edu

\textbf{Abstract.} In the framework of censored regression models, the distribution of the error term can depart significantly from normality, for instance, due to the presence of multimodality, skewness and/or atypical observations. In this paper we propose a novel censored linear regression model where the random errors follow a finite mixture of scale mixtures of normal (SMN) distribution. The SMN is an attractive class of symmetrical heavy-tailed densities that includes the normal, Student-t, slash and the contaminated normal distribution as special cases. This approach allows us to model data with great flexibility, accommodating simultaneously multimodality, heavy tails and skewness depending on the structure of the mixture components. We develop an analytically tractable and efficient EM-type algorithm for iteratively computing the maximum likelihood estimates of the parameters, with standard errors and prediction of the censored values as a by-products. The proposed algorithm has closed-form expressions at the E-step, that rely on formulas for the mean and variance of the truncated SMN distributions. The efficacy of the method is verified through the analysis of simulated and real datasets. The methodology addressed in this paper is implemented in the R package \texttt{CensMixReg}.

\textbf{References}


\textit{Key words and phrases.} Censored regression model, EM-type algorithms, finite mixture models, heavy-tails distributions, limit of detection, Tobit model.


On the asymptotic distribution of sample autocovariance differences of long-memory processes

Mauricio Zevallos

Department of Statistics, University of Campinas, Campinas, SP, Brazil, amadeus@unicamp.br

Abstract. This paper presents a procedure to calculate, in terms of analytic functions, the asymptotic covariance matrix of sample autocovariance differences of stationary autoregressive fractionally integrated moving average process with Gaussian and non-Gaussian errors. Furthermore, an application of minimum distance estimation of Gaussian autoregressive fractionally integrated moving average models is presented.

References


Key words and phrases. ARFIMA, central limit theorem, minimum distance estimation.
Likelihood-based missing data analysis in crossover trials

Savita Pareek\textsuperscript{a}, Kalyan Das\textsuperscript{b} and Siuli Mukhopadhyay\textsuperscript{c}

Department of Mathematics, Indian Institute of Technology Bombay, Mumbai 400 076, India, \textsuperscript{a}savita@math.iitb.ac.in, \textsuperscript{b}kalyan@math.iitb.ac.in, \textsuperscript{c}siuli@math.iitb.ac.in

Abstract. A multivariate mixed-effects model seems to be the most appropriate for gene expression data collected in a crossover trial. It is, however, difficult to obtain reliable results using standard statistical inference when some responses are missing. Particularly for crossover studies, missingness is a serious concern as the trial requires a small number of participants. A Monte Carlo EM (MCEM)-based technique was adopted to deal with this situation. In addition to estimation, MCEM likelihood ratio tests are developed to test fixed effects in crossover models with missing data. Intensive simulation studies were conducted prior to analyzing gene expression data.

References


Key words and phrases. Crossover trials, Monte Carlo EM algorithm, MCEM likelihood ratio tests.


Maximum likelihood estimation for the reflected stochastic linear system with a large signal

Xuekang Zhang\textsuperscript{1,a} and Huisheng Shu\textsuperscript{2,b}

\textsuperscript{1}School of Mathematics-Physics and Finance, and Key Laboratory of Advanced Perception and Intelligent Control of High-end Equipment, Ministry of Education, Anhui Polytechnic University, Wuhu, 241000, China, \textsuperscript{a}xkzhang@ahpu.edu.cn

\textsuperscript{2}College of Science, Donghua University, Shanghai, 201620, China, \textsuperscript{b}hsshu@dhu.edu.cn

Abstract. This paper deals with maximum likelihood estimation for the drift of the reflected stochastic linear system with a large signal. The law of iterated logarithm, consistency, and the asymptotic distributions of the maximum likelihood estimators in both the stationary and the non-stationary cases are studied based on the continuous observation.

References


Key words and phrases. Maximum likelihood estimation, reflected stochastic linear system, large signal, law of iterated logarithm, consistency, asymptotic distributions.


Longitudinal binary response models using alternative links for medical data

Alex de la Cruz Huayanay\textsuperscript{1,a}, Jorge L. Bazán\textsuperscript{2,b} and Carlos A. Ribeiro Diniz\textsuperscript{3,c}

\textsuperscript{1}Inter-Institutional Graduation Program in Statistics, USP/UFSCar, São Carlos, Brazil, \textsuperscript{a}aldehu@usp.br
\textsuperscript{2}Department of Applied Mathematics and Statistics, University of São Paulo, São Carlos, Brazil, \textsuperscript{b}jlbazan@icmc.usp.br
\textsuperscript{3}Department of Statistics, Federal University of São Carlos, São Carlos, Brazil, \textsuperscript{c}dcad@ufscar.br

Abstract. Motivated for a medical data about schizophrenia symptoms where an imbalanced binary response is observed, we introduce a broad class of link functions, called power and reverse power, as an alternative to analyse longitudinal binary data, particularly when it is imbalanced as is common in medical data. Bayesian estimation using an MCMC procedure through the No-U-Turn Sampler algorithm is proposed. Posterior predictive checks, Bayesian randomized quantile residuals, and a Bayesian influence measures are considered for model diagnostics. Different models are compared using selection model criteria. A simulation study is developed to analyse the prior sensitivity of the variance of the random effect and to assess the performance of the proposed model in the presence of imbalanced data. Finally, an application of the methodology studied in a set of medical data on the presence of schizophrenia symptom “thought disorder” is considered. In this data set, the presence of symptoms is much less than the absence, thus we show, in practice, the usefulness of using alternative link functions in imbalanced data.

References


Key words and phrases. Asymmetric link, Bayesian diagnostic, binary response, imbalanced data, mixed-effects model, longitudinal data, health data.


Influence diagnostics for the power-normal Tobit model

Guillermo Martínez-Flórez\textsuperscript{1,a}, Mario Pacheco\textsuperscript{2,b} and Artur J. Lemonte\textsuperscript{3,c}

\textsuperscript{1}Departamento de Matemáticas y Estadística, Universidad de Córdoba, Montería, Colombia, \textsuperscript{a}gmartinezflorez11@gmail.com
\textsuperscript{2}Facultad de Estadística, Universidad Santo Tomás, Bogotá, Colombia, \textsuperscript{b}mariopachecolopez@gmail.com
\textsuperscript{3}Departamento de Estatística, Universidade Federal do Rio Grande do Norte, Natal/RN, Brazil, \textsuperscript{c}arturlemonte@gmail.com

Abstract. Diagnostic analysis tools are studied for the censored power-normal Tobit regression model. We follow the Cook's (J. R. Stat. Soc., Ser. B, Stat. Methodol. \textbf{48} (1986) 133–169) approach, and several perturbation schemes are considered to detect influential observations. In particular, closed-form expressions of the normal curvatures for studying local influence are obtained under some perturbation schemes. The approach pursued also considers separate analysis for regression and scale-asymmetry parameters. Further, we define residuals to identify departures from the model assumptions, as well as to assess the overall goodness-of-fit of the censored power-normal Tobit regression model. The diagnostic measures developed are applied in a real data set for illustrative purposes.

References


Key words and phrases. Influence diagnostic, deviance residual, local influence, power-normal model.


High-dimensional regime for Wishart matrices based on the increments of the solution to the stochastic heat equation

Julie Gamain\textsuperscript{1,a}, David A. C. Mollinedo\textsuperscript{2,c} and Ciprian A. Tudor\textsuperscript{1,3,b}

\textsuperscript{1}CNRS, Université de Lille, Laboratoire Paul Painlevé UMR 8524, F-59655 Villeneuve d’Ascq, France, \textsuperscript{a}julie.gamain@univ-lille.fr, \textsuperscript{b}ciprian.tudor@univ-lille.fr

\textsuperscript{2}Universidade Tecnológica Federal do Paraná, Brazil

\textsuperscript{3}Simion Stoilow Institute of Mathematics of the Romanian Academy, Bucharest, Romania, \textsuperscript{c}davida@utfpr.edu.br

Abstract. We consider a $n \times d$ random matrix $X_{n,d}$ whose entries are the spatial increments of the solution to the stochastic heat equation with space-time white noise. We analyze the limit behavior of the associated Wishart matrix, by showing that it converges almost surely to a diagonal matrix (with equal diagonal terms) and the renormalized Wishart matrix satisfies a central limit theorem. Our techniques are based on the analysis on Wiener chaos, Malliavin calculus and Stein’s method.

References


Key words and phrases. Wishart matrix, stochastic heat equation, Wiener chaos, multiple stochastic integrals, Malliavin calculus, high-dimensional regime.


Dependent percolation on $\mathbb{Z}^2$

Bernardo N. B. de Lima$^{1,a}$, Vladas Sidoravicius$^2$ and Maria Eulália Vares$^{3,b}$

$^1$Departamento de Matemática, Universidade Federal de Minas Gerais, Av. Antônio Carlos 6627, CEP 30123-970, Belo Horizonte, MG, Brazil, $^a$bnblima@mat.ufmg.br

$^2$Shanghai New York University, Pudong New District, Shanghai, China

$^3$Instituto de Matemática, Universidade Federal do Rio de Janeiro, Av. Athos da Silveira Ramos 149, CEP 21941-909, Rio de Janeiro, RJ, Brazil, $^b$eulalia@im.ufrj.br

Abstract. We consider a dependent percolation model on the square lattice $\mathbb{Z}^2$. The range of dependence is infinite in vertical and horizontal directions. In this context, we prove the existence of a phase transition. The proof exploits a multi-scale renormalization argument that is defined once the environment configuration is suitably good and, which, together with the main estimate for the induction step, comes from Kesten, Sidoravicius and Vares (Electronic Journal of Probability 27 (2022) 1–49). This paper is inspired by de Lima (Ph.D.Thesis, Informes de Matemática. IMPA, Série C-26/2004) where the simpler case of a deterministic environment was considered. It has various applications, including an alternative proof for the phase transition on the two dimensional random stretched lattice proved by Hoffman (Comm. Math. Phys. 254 (2005) 1–22).

References


Key words and phrases. Dependent percolation, multiscale renormalization, random environment.


On the two-point function of the one-dimensional KPZ equation

Sergio I. López¹,ᵃ and Leandro P. R. Pimentel²,ᵇ

¹Facultad de Ciencias, Universidad Nacional Autónoma de México, Mexico City, Mexico, ᵃsilo@ciencias.unam.mx
²Institute of Mathematics, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil, ᵇleandro@im.ufrj.br

Abstract. In this short communication, we show that basic tools from Malliavin calculus can be applied to derive the two-point function of the slope of the one-dimensional KPZ equation, starting from a two-sided Brownian motion with an arbitrary diffusion parameter, in terms of the polymer end-point annealed distribution associated to the stochastic heat equation. We also prove that this distribution is given in terms of the derivative of the variance of the solution of the KPZ equation.

References


Key words and phrases. KPZ equation, two-point function, Malliavin calculus.


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


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/.


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

Correspondence. Mail concerning membership, subscriptions, nonreceipt claims, copyright permissions, advertising or back issues should be sent to the IMS Dues and Subscription Office, PO Box 729, Middletown, Maryland 21769, USA. Mail concerning submissions or editorial content should be sent to the Editor at nancy@ime.unicamp.br. Mail concerning the production of this journal should be sent to: Geri Mattson at bjps@mattsonpublishing.com.

Individual and Organizational Memberships: https://www.imstat.org/individual-membership/


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