Contents

Y. LAZAR and B. ALMUTAIRI
Dirac distributions related to sums of independent nonidentically uniform random variables ................................................................. 435

T. GOYAL, S. K. MAURYA and S. NADARAJAH
Geometric generated family of distributions: A review ................................. 442

D. A. SPADE
A Monte Carlo integration approach to estimating drift and minorization coefficients for Metropolis–Hastings samplers ..................................... 466

B. VINESHKUMAR and N. UNNIKRISHNAN NAIR
Inferring association from reliability functions: An approach based on copulas .............................................................. 484

M. LÖWE and S. TERVEER
A Central Limit Theorem for incomplete U-statistics over triangular arrays ... 499

D. SENGUPTA, S. ROY and T. BANERJEE
Testing of Poisson mean with under-reported counts .................................. 523

G. JONGBLOED, F. VAN DER MEULEN and L. PANG
Nonparametric Bayesian estimation of a concave distribution function with mixed interval censored data ............................................. 544

J. HOKAMA, P. MORETTIN, H. BOLFARINE and M. GALEA
Inference in a linear functional relationship with replications ............... 569

D. TRAN
Behavior of the Fréchet mean and Central Limit Theorems on spheres .... 590

E. SÖNMEZ
Graph distances of continuum long-range percolation .......................... 609

J. B. KADANE
Partitioning some multivariate distributions ........................................... 625

M. GALEA, F. VILCA and C. BORELLI ZELLER
Hypotheses tests on the skewness parameter in a multivariate generalized hyperbolic distribution ........................................ 630
Dirac distributions related to sums of independent nonidentically uniform random variables

Youssef Lazar$^{1,2}$ and Bander Almutairi$^3$

$^1$University of East Anglia, Norwich, UK
$^2$Imam Mohammad Ibn Saud Islamic University (IMSIU), Riyadh, KSA, ylazar77@gmail.com
$^3$King Saud University, KSU, Riyadh, KSA, baalmutairi@ksu.edu.sa

Abstract. The aim of this note is to give an elegant proof of a result due to E. G. Olds which concerns the density distribution of the sum of independent uniform random variables non-identically distributed. The proof uses both analytical and combinatorial properties of Dirac distributions and their convolutions. The method is new and can apply to other situations.

References


Hall, P. (1927). The distribution of means for samples of size $N$ drawn from a population in which the variate takes values between 0 and 1, all such values being equally probable. *Biometrika* 19, 240–245.


Roach, S. A. (1963). The frequency distribution of the sample mean where each member of the sample is drawn from a different rectangular distribution. *Biometrika* 50, 508–513. MR0163381 https://doi.org/10.1093/biomet/50.3-4.508


Key words and phrases. Sums of uniform distributions, convolutions, Dirac distributions.
Geometric generated family of distributions: A review

Teena Goyal¹, Sandeep Kumar Maurya² and Saralees Nadarajah³

¹Department of Mathematics and Statistics, Banasthali Vidyapith, Rajasthan, India, teenagoyalnewai@gmail.com
²Department of Statistics, Central University of South Bihar, Bihar, India, sandeepmaurya.maurya48@gmail.com
³School of Mathematics, University of Manchester, Manchester, UK, Saralees.Nadarajah@manchester.ac.uk

Abstract. The present article represents a review of the geometric generated family of distributions. Based on this family of distribution, several distributions are proposed. The family can be proposed by using the compounding concept of zero truncated geometric distribution with any other model or family of distributions. Here, we provide a complete survey on this family of distributions and also listed the contributory related research work, their sub-models, hazard rates, and utilized real datasets. We also address 10 power series distributions, 60 distributions based on the geometric family of distribution. These numbers show the importance of the geometric family of distribution.

References


Key words and phrases. Lifetime models, geometric distribution, power series distribution, compounding technique, real data application.

A Monte Carlo integration approach to estimating drift and minorization coefficients for Metropolis–Hastings samplers

David A. Spade

Department of Mathematical Sciences, University of Wisconsin–Milwaukee, 3200 N. Cramer Street, EMS E403, Milwaukee, Wisconsin 53211, USA, spade@uwm.edu

Abstract. Bayesian statistical methodology has become highly popular in a myriad of applications over the past several decades. In Bayesian statistics, it is often required to draw samples from intractable probability distributions. Markov chain Monte Carlo (MCMC) algorithms are common methods of obtaining samples from these distributions. When an MCMC algorithm is used, it is important to be able to obtain an answer to the question of how many iterations the chain must run before it is “close enough” to its target distribution to allow approximate sampling from this distribution. Several methods of approaching this question exist in the literature. Some rely on the output of the chain, and some are based on Markov chain theory. These techniques suffer from major practical limitations. This work provides a computational method of bounding the mixing time of a Metropolis–Hastings algorithm. This approach extends the work of Spade (Statistics and Computing 26 (2016) 761–781) and Spade (Markov Processes and Related Fields 26 (2020) 487–516) to general versions of the Metropolis–Hastings algorithm, while examining the convergence behavior of such samplers under symmetric and asymmetric proposal densities.

References


Key words and phrases. Markov chain Monte Carlo, mixing time, geometric ergodicity, minorization.
Inferring association from reliability functions: 
An approach based on copulas

B. Vineshkumar\(^1\) and N. Unnikrishnan Nair\(^2\)

\(^1\)Department of Statistics, Government Arts College, Thiruvananthapuram-695014, India, Vinesh910@gmail.com

\(^2\)Department of Statistics, Cochin University of Science and Technology, Cochin-682022, India, unnikrishnannair4@gmail.com

Abstract. Nair, Sankaran and John (Metron 76 (2018) 133–153) have defined and studied the properties of reliability functions in terms of copulas. In the present paper, we investigate the utility of such functions in inferring the time-dependent association of bivariate distributions. We consider the Clayton measure of association for the study. A general expression for this measure in terms of the generator of Archimedean copulas is given, and a method of finding nature of association using the generators is provided. We derive the relationship of the association measure with the ageing property of the distribution, associated with the generator. We analyze how the hazard rate of survival copulas can be utilized in studying the association between two random variables. Applications of the results in real life situations are discussed.

References


Key words and phrases. Archimedean copula, Clayton measure, hazard rate of survival copulas.


A Central Limit Theorem for incomplete U-statistics over triangular arrays

Matthias Löwe* and Sara Terveer†

Fachbereich Mathematik und Informatik, Universität Münster, Einsteinstraße 62, 48149 Münster, Germany,

*maloewe@uni-muenster.de; †sara.terveer@uni-muenster.de

Abstract. We analyze the fluctuations of incomplete U-statistics over a triangular array of independent random variables. We give criteria for a Central Limit Theorem (CLT, for short) to hold in the sense that we prove that an appropriately scaled and centered version of the U-statistic converges to a normal random variable. Our method of proof relies on a martingale CLT. An application, a CLT for the hitting time for random walks on random graphs, will be presented in Löwe and Terveer (2020).

References


Key words and phrases. Central Limit Theorem, U-statistics, incomplete U-statistics, triangular arrays.
Testing of Poisson mean with under-reported counts

Debjit Sengupta¹,*, Surupa Roy¹ and Tathagata Banerjee²

¹Department of Statistics, St. Xavier’s College, 30, Mother Teresa Sarani, Kolkata-700016, India, *debjits10@gmail.com
²Indian Institute of Management, Vastrapur, Ahmedabad 380015, India

Abstract. For modelling unbounded count data, Poisson distribution is a natural choice. However, count data arising in various fields of scientific research are often under-reported. In such situations, inference carried out on the basis of Poisson model will result in biased parameter estimates and suboptimal tests. A modified Poisson model is developed to accommodate the possible undercount. For model-identifiability a double sampling scheme of data collection has been adopted. The focus of this paper is to develop asymptotically optimal tests for the Poisson mean in presence of undercount. Simulation study is conducted to compare the performance of the tests with respect to level and power and also to investigate the impact of ignoring undercount on each of the tests. The findings are validated using real life data.

References


Key words and phrases. Count data, under-reporting, validation data, maximum likelihood, optimal tests.


Nonparametric Bayesian estimation of a concave distribution function with mixed interval censored data

Geurt Jongbloed*, Frank van der Meulen† and Lixue Pang‡

Delft Institute of Applied Mathematics, Delft University of Technology, Van Mourik Broekmanweg 6, 2628 XE Delft, The Netherlands, *g.jongbloed@tudelft.nl; †f.h.vandermeulen@tudelft.nl; ‡lixuepang3@gmail.com

Abstract. Assume we observe a finite number of inspection times together with information on whether a specific event has occurred before each of these times. Suppose replicated measurements are available on multiple event times. The set of inspection times, including the number of inspections, may be different for each event. This is known as mixed case interval censored data. We consider Bayesian estimation of the distribution function of the event time while assuming it is concave. We provide sufficient conditions on the prior such that the resulting procedure is consistent from the Bayesian point of view. We also provide computational methods for drawing from the posterior and illustrate the performance of the Bayesian method in both a simulation study and two real datasets.

References


*Key words and phrases.* Bayesian nonparametrics, Dirichlet process, Markov Chain Monte Carlo, posterior consistency, shape constrained inference.


Inference in a linear functional relationship with replications

Julio Hokama\textsuperscript{1}, Pedro Morettin\textsuperscript{2,*}, Heleno Bolfarine\textsuperscript{2,†} and Manuel Galea\textsuperscript{3}

\textsuperscript{1}Universidade Federal Rural do Rio de Janeiro, Rio de Janeiro, Brazil, jhokama41@gmail.com
\textsuperscript{2}Universidade de São Paulo, São Paulo, Brazil, pamorettin@gmail.com; \textsuperscript{†}hbolfar08@gmail.com
\textsuperscript{3}Pontificia Universidad Católica de Chile, Santiago, Chile, mgalea@mat.uc.cl

Abstract. In this paper, we consider a model for data analysis with measurement errors. The main objective of this work is to develop statistical inference tools, such as parameter estimation and hypothesis tests in a linear functional relationship with replicated observations. For this purpose, we use the maximum likelihood method in the presence of incidental parameters, and the unbiased estimating equations approach. Both approaches lead to explicit expressions for the asymptotic covariance matrices of the estimators of the model parameters. A simulation study is performed to assess the empirical behavior of estimators and of a Wald statistic. The methodology is illustrated with a real data set.

References


\textbf{Key words and phrases.} Linear functional relationship, replicated observations, maximum likelihood estimation, unbiased estimating equations, asymptotic normality, Wald statistic.
Behavior of the Fréchet mean and Central Limit Theorems on spheres

Do Tran

Institute for Mathematical Stochastics, University of Göttingen, Goldschmidtstrasse 7, 37077 Göttingen, Germany, do.tranvan@uni-goettingen.de

Abstract. We compute higher derivatives of the Fréchet function on spheres with an absolutely continuous and rotationally symmetric probability distribution. Consequences include (i) a practical condition to test if the mode of the symmetric distribution is a local Fréchet mean; (ii) a central limit theorem on spheres with practical assumptions and an explicit limiting distribution; and (iii) an answer to the question of whether the smeary effect can occur on spheres with absolutely continuous and rotationally symmetric distributions: with the method presented here, it can in dimension at least 4.

References


Key words and phrases. Spherical statistic, Central Limit Theorem, Fréchet mean, smeary.
Graph distances of continuum long-range percolation

Ercan Sönmez

Department of Statistics, University of Klagenfurt, Universitätsstraße 65–67, 9020 Klagenfurt, Austria,
ercan.soenmez@aau.at

Abstract. We consider a version of continuum long-range percolation on finite boxes of \(\mathbb{R}^d\) in which the vertex set is given by the points of a Poisson point process and each pair of two vertices at distance \(r\) is connected with probability proportional to \(r^{-s}\) for a certain constant \(s\). We explore the graph-theoretical distance in this model. The aim of this paper is to show that this random graph model undergoes phase transitions at values \(s = d\) and \(s = 2d\) in analogy to classical long-range percolation on \(\mathbb{Z}^d\), by using techniques which are based on an analysis of the underlying Poisson point process.

References


Institute of Mathematical Statistics. MR3791470


Newman, C. M. and Schulman, L. S. (1986). One dimensional \(1/|j − i|^s\) percolation models: The existence of a transition for \(s \leq 2\). *Communications in Mathematical Physics* **104**, 547–571. MR0841669


Key words and phrases. Diameter, graph distance, random graphs, long-range percolation, random connection model, Poisson process.

Partitioning some multivariate distributions

Joseph B. Kadane

Department of Statistics & Data Science, Carnegie Mellon University, Baker Hall 232, Pittsburgh, Pennsylvania 15213, USA, kadane@stat.cmu.edu

Abstract. This is a study of the behavior under partition of the sample space of three multivariate distributions: multinomial, multinomial-Dirichlet, and Dirichlet. A general theorem is given, of which all three are special cases.

References


Key words and phrases. Multinomial distribution, Dirichlet distribution, Dirichlet-multinomial distribution, partition.
Hypotheses tests on the skewness parameter in a multivariate generalized hyperbolic distribution

Manuel Galea\textsuperscript{1}, Filidor Vilca\textsuperscript{2} and Camila Borelli Zeller\textsuperscript{3}

\textsuperscript{1}Pontificia Universidad Católica de Chile, Santiago, Chile, mgalea@mat.uc.cl
\textsuperscript{2}Universidade Estadual de Campinas, São Paulo, Brazil, fily@ime.unicamp.br
\textsuperscript{3}Universidade Federal de Juiz de Fora, Minas Gerais, Brazil, camila.zeller@ufjf.edu.br

Abstract. The class of generalized hyperbolic (GH) distributions is generated by a mean-variance mixture of a multivariate Gaussian with a generalized inverse Gaussian (GIG) distribution. This rich family of GH distributions includes some well-known heavy-tailed and symmetric multivariate distributions, including the Normal Inverse Gaussian and some members of the family of scale-mixture of skew-normal distributions. The class of GH distributions has received considerable attention in finance and signal processing applications. In this paper, we propose the likelihood ratio (LR) test to test hypotheses about the skewness parameter of a GH distribution. Due to the complexity of the likelihood function, the EM algorithm is used to find the maximum likelihood estimates both in the complete model and the reduced model. For comparative purposes and due to its simplicity, we also consider the Gradient (G) test. A simulation study shows that the LR and G tests are usually able to achieve the desired significance levels and the testing power increases as the asymmetry increases. The methodology developed in the paper is applied to two real datasets.

References


\textit{Key words and phrases}. Likelihood ratio test, gradient test, EM algorithm, generalized hyperbolic distribution, normal inverse Gaussian distribution.


AIMS AND SCOPE

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

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

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

GENERAL INFORMATION


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.


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 bjets@mattsonpublishing.com.

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

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


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