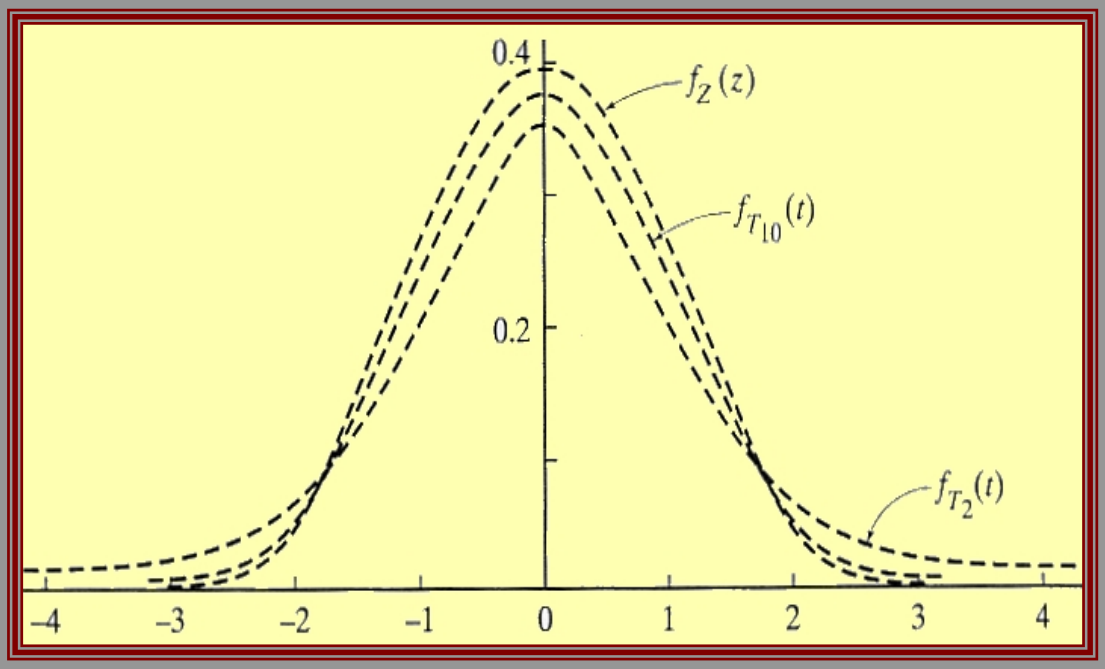
ISSN 1726-3328

# JPSS

**A comprehensive journal of probability and statistics**
**for theorists, methodologists, practitioners, teachers, and others**




# JOURNAL OF PROBABILITY AND STATISTICAL SCIENCE

**Volume 13 Number 1** **February 2015**

**ISSN 1726-3328**

# JPSS

## Journal of Probability and Statistical Science

A Comprehensive Journal of Probability and Statistics

for Theorists, Methodologists, Practitioners, Teachers, and Others

**Volume 13 Number 1**

**February 2015**

# *JPSS* Editorial Board

# *JPSS* Editorial Board

(continued)

機統學誌 JPSS ISSN 1726-3328

Journal of Probability and Statistical Science

Published by: Susan Rivers' Cultural Institute, Hsinchu, Taiwan, ROC

(Cosponsored by: The International Chinese Association of Quantitative Management, Taiwan, ROC)

**Aims and Scope** The ***Journal of Probability and Statistical Science*** (***JPSS***, ISSN 1726-3328) is a modified version of the ***Journal of Propagations in Probability and Statistics*** (***JPPS***, ISSN 1607-7083). ***JPSS***, like its predecessor ***JPPS***, is a multipurpose and comprehensive journal of probability and statistics that publishes papers of interest to a broad audience of theorists, methodologists, practitioners, teachers, and any other users of probability and/or statistics. The scope of ***JPSS*** is intended to be quite broad, including all the major areas of probability and statistics. Research papers involving probability and/or statistics, either theoretical or applied in nature, will be seriously considered for publication. Additionally, papers involving innovative techniques or methods in teaching probability and/or statistics will also be considered. Papers submitted for publication consideration will be peer reviewed. Initially, we will publish semiannually, one issue each in February and August. Hopefully, as time accrues, we will be able to publish quarterly. It is the goal of ***JPSS*** to publish a wide range of works involving probability and/or statistics (theoretical and/or applied in nature) and provide widespread availability of such to a broad audience of people interested in probability and/or statistics.

## Submission and Review Policies

1. Three hard copies of the manuscript written in English should be mailed to the **Editor-in-Chief, an Editor, a Coordinating Editor,** or to the **Managing Editor** at the address provided in the ***JPSS*** Editorial Board. Alternatively, submission of manuscript by email attachment is also acceptable.
2. A manuscript is accepted only with the understanding that the text has not appeared in publication, and that it is not being simultaneously reviewed by any other journal.
3. Submitted manuscripts are refereed by a double-blind process, meaning that the reviewers will not know the names of the authors and vice versa.
4. If an article is accepted for publication, the author(s) will be required to provide an electronic copy of the paper, in **Micro-soft Word** or **PCTEX** format, on a floppy disk or through an email attachment. The authors will also be requested to transfer their copyright on certain conditions to the publisher.

**Publisher** Harold C. H. Chen, Head, Susan Rivers' Cultural Institute. Address: 26, Lane 2, Chien Mei Road, Hsinchu, Taiwan, ROC. Phone: (03)5716594, Fax: (03)5712524.

魏蘇珊文教事業機構發行

總公司: 中華民國臺灣新竹市建美路 2 巷 26 號。

JPSS

Journal of Probability and Statistical Science

**Volume 13 Number 1 February 2015**

## Table of Contents

*Journal of Probability and Statistical Science*
**13**(1), 1-23, Feb. 2015

# A Representation of Hermitian Quadratic Forms in Singular Normal Vectors and Related Distributional Results

Serge B. Provost and A. Akbar Mohsenipour
*The University of Western Ontario*

**ABSTRACT** No completely general representation of noncentral indefinite Hermitian quadratic forms in possibly singular complex normal vectors is currently available in the statistical literature. In this paper, such quadratic forms are represented as the difference of two real positive definite quadratic forms plus an independently distributed linear combination of normal random variables. An accurate moment-based approximation to their density function is also proposed. Several particular cases of interest are also considered, including those involving real quadratic forms or expressions in singular normal vectors for which a similar decomposition applies. Four numerical examples illustrate the distributional results.

***Keywords*** Hermitian quadratic forms; Cumulant generating function; Singular normal vectors; Real quadratic expressions; Moments; Density approximation; Gamma distribution.

## 1. Introduction

Hermitian quadratic forms in several complex normal variables frequently arise in binary hypothesis testing problems, especially in the performance analysis of systems whose inputs are affected by random noise such as radars, sonars, communications receivers and signal acquisition devices. This is explained for instance in [12, 5, 13]. As pointed out by [3], the decision variables in many systems can also be characterized by means of Hermitian quadratic forms in complex Gaussian vectors. Moreover, as explained in Section 2.16 of [26], several statistics used for testing hypotheses on the parameters of complex random vectors

Received July 2012, revised May 2014, in final form August 2014.
Authors of this article are affiliated with the Department of Statistical & Actuarial Sciences at the University of Western Ontario, London, Ontario, Canada N6A 5B7; email address: provost@stats.uwo.ca and amohsen@alumni.uwo.ca.

 ISSN 1726-3328

*Journal of Probability and Statistical Science*
**13**(1), 25-34, Feb. 2015

# Empirical Likelihood Confidence Regions in QVF Measurement Error Models

Naima Shifa — *DePauw University*
Hanfeng Chen — *Bowling Green State University*
Yuehua Wu — *York University*

**ABSTRACT** In applications in many areas, the data sets are contaminated or corrupted by the mismeasured covariate variables resulting in the so-called measurement errors. In this article, we propose an empirical likelihood method for constructing confidence regions in measurement error models. The quasi variance functions are used to establish functional constraints for the empirical likelihood method. The proposed empirical likelihood confidence regions in measurement error models as a nonparametric approach has clear advantages over the parametric model approaches. The distribution of the response error as well as the distribution of the error prone covariates does not need to be specified. Simulation studies in a simple linear model and a logit regression model are conducted. The results show that the proposed method performs well.

***Keywords*** Berkson's error model; Empirical likelihood method; Linear regression model; Logit regression; Measure error models; Quasi variance function.

## 1. Introduction

Problems of measurement error spread through all research areas from agriculture to astronomy. A common source of measurement error problem comes from incorrect response to a survey question, incorrect coding of a correct response and the use of a correctly measured variable as a proxy for another theoretically valid but unobserved variable. The statistical consideration of models containing measurement errors began as early as 1877 (see [8]). There is a large literature in all branches of science concerned with inference in linear and nonlinear models with measurement errors. [6] proposed a simulation based estimation

Received July 2013, revised August 2014, in final form October 2014.
Naima Shifa is an Assistant Professor in the Mathematics Department at DePauw University, Greencastle, IN 46135-0037, USA; email: naimashifa@depauw.edu. Hanfeng Chen is a Professor in the Department of Mathematics and Statistics at Bowling Green State University, Bowling Green, OH 43551, USA; email: hchen@bgsu.edu. Yuehua Wu is affiliated to the Department of Mathematics and Statistics at York University, Toronto, Ontario M3J 1P3, Canada; email: wuyh@mathstat.yorku.ca.
The last two authors' research is partially supported by the Natural Sciences and Engineering Research Council of Canada.

 ISSN 1726-3328

*Journal of Probability and Statistical Science*
**13**(1), 35-52, Feb. 2015

# Bivariate Generalized Exponential Distribution Derived from Farlie-Gumbel-Morgenstern and Clayton Copulas Using Censored Data and Covariates

Jorge Alberto Achcar
*Universidade Estadual Paulista and Universidade de São Paulo*

Fernando Antônio Moala  Leandro Fernandes Coladello
*Universidade Estadual Paulista*

**ABSTRACT** In this paper, we introduce a Bayesian analysis for a bivariate generalized exponential distribution derived from the Farlie-Gumbel-Morgensten in presence of censored data and covariates. The generalized exponential distribution could be a good alternative to analyze lifetime data in comparison to usual existing parametric lifetime distributions as Weibull or Gamma distributions. A Bayesian approach is carried out assuming different prior distributions for the parameters and MCMC (Markov Chain Monte Carlo) methods. In this paper we also derive a bivariate Generalized Exponential distribution from the Clayton copula and a comparison between the both distributions is carried out. A simulation study is illustrated to test the performance of the proposed models and one example with real lifetime data set is introduced to illustrate the proposed study.

***Keywords*** Bivariate generalized exponential distribution; Copula function; Bayesian analysis; Censored data; Covariates.

## 1. Introduction

In medical, engineering or other lifetime data applications, we could have more than one lifetime associated to each unit. A special situation is the presence of two lifetimes $T_1$ and $T_2$

Received September 2013, revised September 2014, in final form October 2014.
Jorge Alberto Achcar is affiliated to the Departamento de Matemática, Estatística e Computação, UNESP, at Universidade Estadual Paulista, Presidente Prudente, SP, Brazil and the Departamento de Medicina Social, FMRP, Universidade de São Paulo, Avenida Bandeirantes, 3900, CEP: 14048-900, Ribeirão Preto, SP, Brasil; email: achcar@fmrp.usp.br. Fernando Antônio Moala and Leandro Fernandes Coladello are affiliated to the Departamento de Matemática, Estatística e Computação, UNESP, at Universidade Estadual Paulista, Presidente Prudente, SP, Brazil.

 ISSN 1726-3328

*Journal of Probability and Statistical Science*
**13**(1), 53-68, Feb. 2015

# Statistical Estimation of Multiple Poisson Rate

Dadao Feng, Chuoxin Ma and Maozai Tian
*Renmin University of China  Lanzhou University of Finance and Economics*

**ABSTRACT** In this paper, we investigate various confidence intervals for the multiple Poisson rate (MPR). Eight confidence intervals (namely, logarithm Wald, conditional Wald, binomial Wald, uniform Beyesian, exact, score and SA-based confidence intervals) are derived. For comparison, we use Monte Carlo method to calculate the coverage probabilities and the average interval widths of the eight confidence intervals when the real parameter value ranges from 0.1% to 1%. Our simulation results suggest that the SA-based confidence interval generally performs better. Particularly, when the disease incidence is around 0.1‰, the width of the SA-based confidence interval is 35% significantly shorter than others. Additionally, real data analysis is applied to the 1978-1982 New York State Leukemia dataset.

***Keywords*** Multiple Poisson rate; Saddlepoint approximation; Monte Carlo simulation; Confidence interval.

## 1. Introduction

The multiple Poisson rate (MPR) is commonly used in epidemiological and medical applications as an indicator of rare disease incidence. It reflects the baseline risk of contracting certain disease among different groups [1]. For instance, in the research of the 1978-1982 New York State Leukemia dataset [1], one may wish to estimate the incidence of Leukemia and investigate the correlation between the disease rate and certain aspects of the population (such as income, population density, etc.). Let $X_i \sim \text{Poisson}(\tau_i)$ be the number of people having the risk to have Leukemia during the period between 1978 and 1982 in the *i*th district in New York State, $Y_i \sim \text{Poisson}(\beta\tau_i)$ be the number of people contracting the disease, $i = 1, \ldots, n$. $X_i$ and $Y_i$ are independent. $\tau_i$, which is called the additional factor, measures the population density of the *i*th district and $\beta$, which is called the overall effect,

Received September 2013, revised August 2014, in final form September 2014.
Dadao Feng, Chuoxin Ma are affiliated to the School of Statistics at Lanzhou University of Finance and Economics, Lanzhou, 730101, Gansu, China. Maozai Tian (corresponding author) is affiliated to the Center for Applied Statistics in the School of Statistics at Renmin University of China, Beijing, 100872, China; e-mail: mztian@ruc.edu.cn.

 ISSN 1726-3328

*Journal of Probability and Statistical Science*
**13**(1), 69-83, Feb. 2015

# The Most Conservative Dirichlet Prior Distribution for Trinomial Sampling

Paul Chiou
*Lamar University*

Weiwen Miao
*Haverford College*

**ABSTRACT** The incorporation of prior information about a parameter into a statistical procedure is an essential feature of Bayesian statistics. However, the manner in which this is done is often arbitrary. In order to reduce such arbitrariness, Dyer and Chiou [7] have introduced a methodology based on information-theoretic concepts to incorporate prior information in binomial sampling. Due to the mathematical tractability, Chiou and Miao [5] have proposed the Rényi information measure of order one-half instead of the Kullback-Leibler information measure to choose a beta prior distribution for binomial sampling. In this article, we will study the choice of a Dirichlet prior distribution (Lindley [15]; Good [9]) for trinomial sampling based on the Rényi information in three situations: (1) no prior estimate of $\theta_i$, where $\theta_i$, $i$ = 1; 2; 3, is the probability that the outcome belongs to the $i$th category, is available, (2) a prior point estimate of $\theta_i$ for some category is available, and (3) a prior point estimate of $\theta_i$ and $\theta_j$ for two categories are available.

***Keywords*** Conservative prior distribution; Information theory; Ranking of prior distributions; Rényi information measure.

## 1. Introduction

Suppose that sample data $\tilde{x}$ are obtained by observing the random vector $\tilde{X}$ whose sampling distribution is assumed to belong to some specified family of probability density functions (pdfs) $F = \{f_{\tilde{\theta}}(\tilde{x});\ \tilde{x} \in \chi,\ \tilde{\theta} \in \Omega\}$, where $\chi$ is the sample space and $\Omega$ is the parameter space. The parameter $\tilde{\theta}$, which indexes the members of $F$ is unknown. Under the Bayesian approach, $\tilde{\theta}$ is assumed to be a value taken on by the random vector $\tilde{\Theta}$ whose prior pdf is $g(\tilde{\theta})$. The essence of $g(\tilde{\theta})$ is to incorporate into the statistical procedure all

Received October 2014, revised January 2015, in final form January 2015.
Paul Chiou is a Professor in the Department of Mathematics at Lamar University, Beaumont, TX 77710, USA; email: paul.chiou@lamar.edu. Weiwen Miao is an Associate Professor in the Department of Mathematics and Statistics at Haverford College, Haverford, PA 19041, USA; email: wmiao@haverford.edu.

 ISSN 1726-3328

*Journal of Probability and Statistical Science*
**13**(1), 85-102, Feb. 2015

# Inference for Lomax Distribution Based on Trimmed Samples: A Bayesian Approach

Tabassum Naz Sindhu
Muhammad Aslam
*Quaid-i-Azam University*

Navid Feroze
*Allama Iqbal Open University*

**ABSTRACT** The Lomax distribution has an important position in the field of lifetime testing. This distribution is useful for modeling and analyzing the lifetime data in medical, biological and engineering sciences. Trimmed samples are widely utilized in several areas of statistical practice, especially when some sample values at either or both extremes might have been adulterated. In this article, the problem of estimating the parameter of Lomax distribution based on trimmed samples under informative and non-informative priors has been addressed based on a Bayesian framework. Elicitation of hyper-parameter through prior predictive approach has also been discussed. Posterior predictive distributions along with posterior predictive intervals and credible intervals have also been derived under different priors. A comparison among different estimators has been made using the Monte Carlo simulation. A real life data example has also been discussed.

***Keywords*** Inverse transformation method; Doubly censored samples; Loss functions; Posterior predictive distributions; Credible intervals; Predictive intervals.

## 1. Introduction

The Lomax distribution was originally proposed as a second kind of the Pareto distribution by Lomax [1]. It is considered to be a good model in biomedical problems. The Lomax distribution has an important position in the field of lifetime testing. This distribution is useful for modeling and analyzing the lifetime data in medical, biological and engineering sciences: see Habibullah and Ahsanullah [2]. It has been widely used for stochastic modeling of decreasing failure rate life components. For its applications as lifetime distribution and extensions, reader may refer to Marshall and Olkin [3]. The distribution function of Lomax distribution is given by

Received March 2013, revised November 2014, in final form December 2014.
Tabassum Naz Sindhu and Muhammad Aslam are affiliated to the Department of Statistics at Quaid-i-Azam University 45320, Islamabad 44000, Pakistan. Navid Feroze (corresponding author) is affiliated to the Department of Mathematics and Statistics at Allama Iqbal Open University, Islamabad 44000, Pakistan; email: navidferoz@hotmail.com.

 ISSN 1726-3328

*Journal of Probability and Statistical Science*
**13**(1), 103-120, Feb. 2015

# Risk Functions of Rayleigh Model under Progressive Type II Censored Samples: Empirical Bayes Approach

D. R. Barot
*Ahmedabad University*

M. N. Patel
*Gujarat University*

**ABSTRACT** This paper deals with the comparison of empirical Bayes estimates of reliability performances on the basis of risk efficiencies when the data are progressively Type II censored from Rayleigh distribution. The estimates relative to squared error and general entropy loss functions are considered, and their risk functions are obtained under both the mentioned loss function. The proposed comparison method is illustrated by a real data set. The performance of the estimates is examined and compared through an extensive Monte Carlo simulation study. The findings of the study indicate that the empirical Bayes estimates relative to general entropy loss function are preferable than those relative to squared error loss function.

***Keywords*** Empirical Bayes estimates; Rayleigh model; Progressively Type II censored samples; Squared error loss function; General entropy loss function; Risk efficiency criterion.

## 1. Introduction and Prediction Criterion

Empirical Bayes (EB) method of estimation is quite efficient and very popular in theory and practice of statistics. In this method, an experimenter usually does not specify an unknown prior hyper-parameter, but often tempted to use some estimate of a prior hyper-parameter. This approach was first formulated by Robbins [19], and has been described extensively by many authors, including Maritz and Lewin [17], Casella [10], Carlin and Louis [9], Lehmann and Casella [16] among others. Ali Mousa [1] obtained the EB estimates for the Burr type XII model based on Type II censored data. Asgharzadeh and Valiohi [3] studied the EB approach for estimation and prediction problems in the case of proportional hazards family under progressive Type II censoring.

□Received January 2014, revised November 2014, in final form January 2015.
□D. R. Barot is affiliated to the Department of Statistics in the H. L. Institute of Commerce at Ahmedabad University, Ahmedabad 380009, India; email: hl.dineshbarot1@gmail.com. M. N. Patel is affiliated to the Department of Statistics in the School of Sciences at Gujarat University, Ahmedabad 380009, India; email: mnpatel.stat@gmail.com.

 ISSN 1726-3328

*Journal of Probability and Statistical Science*
**13**(1), 121-134, Feb. 2015

# Structured versus Unstructured Covariance Patterns in Modelling Longitudinal Data

K. Radhakrishnan Nair
*Nehru Arts and Science College*

M. Manoharan
*University of Calicut*

**ABSTRACT** Modelling of covariance has become increasingly important in longitudinal data analysis. Misspecification of covariance structure may lead to a large loss of efficiency of estimators of mean parameter. We compare the performances of various covariance structures, in the context of modelling longitudinal data, by comparing their maximum likelihoods, AIC, AICC and BIC, for varied structures of correlation among the responses at different measurement occasions. Various nested covariance models are compared using the maximum likelihoods and non-nested models are compared using AIC, AICC and BIC criteria. Models are statistically validated using two sets of data - real as well as simulated ones. Using Bayesian approach, we show that, often models with lesser parameters may perform better than the unstructured model and that the larger number of covariance parameters involved in it does not always make the unstructured model superior to others. Also as the correlation decays, the performance of unstructured model becomes poorer and poorer.

***Keywords*** Longitudinal data; Covariance structure; Nested models; REML; AIC; AICC; BIC.

## 1. Introduction and Motivation

The last forty years have seen remarkable advances in methods for analyzing longitudinal and clustered data. The problems in which we are interested, when we consider longitudinal studies are similar to those we face under the cross-sectional counterparts and may be classified as analysis of variance (ANOVA) or more generally as regression problems. The basic difference between the two approaches is that in cross sectional studies we deal with independent observations, while in the longitudinal cases it is necessary to consider a possible statistical dependence among them.

Received September 2012, revised June 2014, in final form August 2014.
K. Radhakrishnan Nair is affiliated with the Department of Statistics at Nehru Arts and Science College, Kanhangad, Kerala 671 314, India; email: rknasc@gmail.com. M. Manoharan is affiliated with the Department of Statistics at University of Calicut, Kerala 673 635, India; email: mano30@rediffmail.com.

 ISSN 1726-3328

*Journal of Probability and Statistical Science*
**13**(1), 135-152, Feb. 2015

# On Generating a New Family of Distributions Using the Logit Function

Raid Al-Aqtash — *Marshall University*
Felix Famoye, Carl Lee — *Central Michigan University*

**ABSTRACT** In this article, a method for generating new family of distributions using the logit function is investigated. The properties of symmetry, quantile function, and Shannon entropy of the family are presented. Some new members in the family of distributions are provided. A member of the new family, the Gumbel-Weibull distribution, is defined and studied. Some properties of the Gumbel-Weibull distribution, including shapes, hazard function, moments, skewness, and Shannon entropy are given. Two real data sets are used to illustrate the application of Gumbel-Weibull distribution and the results are compared to other commonly used distributions.

***Keywords*** Beta-generated; Generalized class; Moments; Quantiles; Shannon entropy.

## 1. Introduction

A class of distributions, namely the beta-generated distributions, was developed and studied by Eugene *et al*. [16]. These authors used the property that the beta random variable lies between 0 and 1 to define the class. The cumulative distribution function (CDF) of the beta-generated class is given by

$$G(x)=\int_0^{F(x)}\frac{1}{B(\alpha,\beta)}t^{\alpha-1}(1-t)^{\beta-1}dt,\ \alpha,\beta>0, \tag{1.1}$$

where $F(x)$ is the CDF of any continuous random variable $X$. Many research articles on beta-generated distributions have been published since the 2002 paper by Eugene *et al*. [16]. Examples include beta-Pareto (Akinsete *et al*. [1]), beta-Gumbel (Nadarajah and Kotz [25]), beta-gamma (Kong *et al*. [20]), and beta-Cauchy (Alshawarbeh *et al*. [3]). For detailed review of this family and other methods for generating distributions, one may refer to Lee *et al*. [22].

Received July 2013, revised April 2014, in final form September 2014.
Raid Al-Aqtash is affiliated to the Department of Mathematics at Marshall University, Huntington, West Virginia 25755, USA; email: alaqtash@marshall.edu. Felix Famoye and Carl Lee are affiliated to the Department of Mathematics at Central Michigan University, Mount Pleasant, Michigan 48859, USA; emails: felix.famoye@cmich.edu and lee1c@cmich.edu.

 ISSN 1726-3328