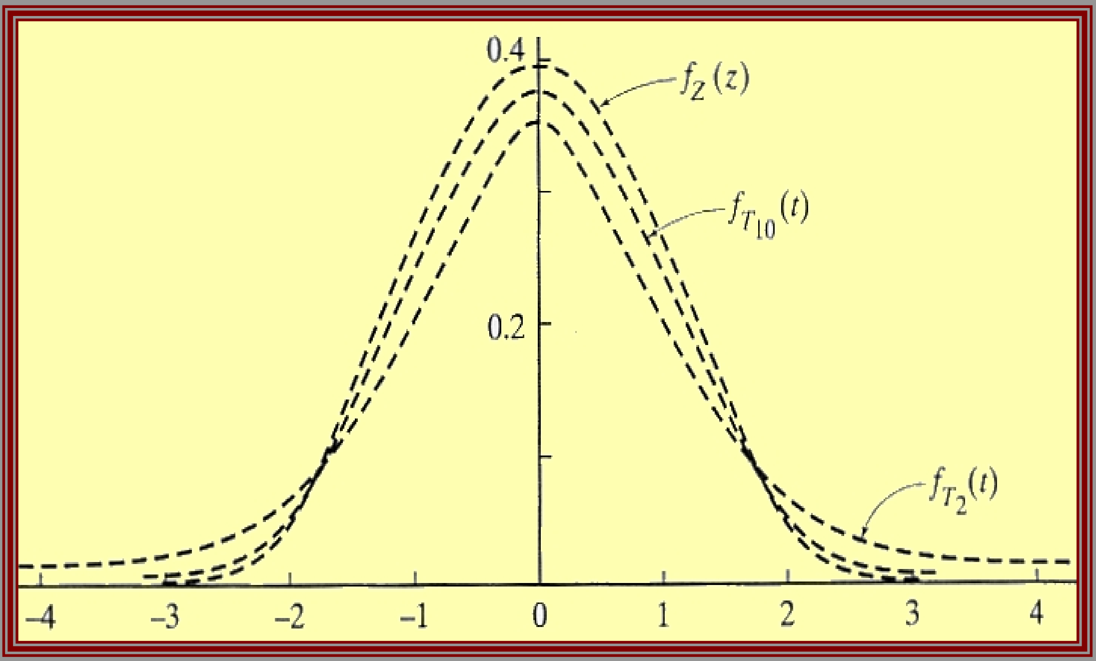
ISSN 1726-3328

JPSS

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



JOURNAL OF PROBABILITY AND STATISTICAL SCIENCE

Volume 6 Number 2 August 2008

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 6 Number 2

August 2008

JPSS Editorial Board

JPSS Editorial Board

(continued)

JPSS

Journal of Probability and Statistical Science

Volume 6 Number 2 August 2008

<u>Table of Contents</u>

Journal of Probability and Statistical Science
6(2), 125-134, Aug. 2008

On the Weak Law of Large Numbers for Double Arrays of Banach Space Valued Random Elements

Nguyen Van Quang
University of Vinh

Nguyen Van Huan
University of Dong Thap

ABSTRACT The aim of this paper is to extend the "classical degenerate convergence criterion" and the Feller weak law of large numbers to adapted double arrays of p-uniformly smooth Banach space valued random elements. Our result is more general and stronger than some well-known ones.

Keywords *P*-uniformly smooth Banach space; Doubly adapted array of random elements; Weak law of large numbers; Convergence in probability; Martingale difference; Sum of i.i.d. random elements.

1. Introduction

The WLLN has been extended to array of random variables or random elements (for random variables, see Hong and Lee [7], Hong and Oh [8], Sung [13], and Sung *et al*. [14, 15]; for random elements, see Adler *et al*. [1, 2], Ahmed *et al*. [3], Hong *et al*. [9, 11], and Sung *et al*. [14, 15]). The aim of this paper is to extend the "classical degenerate convergence criterion" and the Feller weak law of large numbers to adapted double arrays of p-uniformly smooth Banach space valued random elements. Our result is more general and stronger than some well-known ones.

Let us begin with some notations and definitions. Throughout this paper, the symbol C will denote a generic constatnt ($0 < C < \infty$) which is not necessarily the same one in each appearance. A real separable Banach space $\mathbf{E}$ is said to be p-uniformly smooth ($1 \le p \le 2$) if

---

□ Received December 2007, revised March 2008, in final form April 2008.
□ Nguyen Van Quang is affiliated to the Department of Mathematics at University of Vinh, 182 Le Duan, Vinh, Nghean, Vietnam; email: nvquang@hotmail.com. Nguyen Van Huan is affiliated to the Department of Mathematics at University of Dong Thap, 783 Pham Huu Lau, Cao Lanh, Dong Thap, Vietnam; email: vanhuandhdt@yahoo.com.
□ AMS 2000 Subject Classifications: 60F05, 60G50, 60G42.

 ISSN 1726-3328

Journal of Probability and Statistical Science
6(2), 135-154, Aug. 2008

Some Shrunken Testimators for the Variance of a Normal Distribution in Double Stage Samples under an Asymmetric Loss Function

Gyan Prakash D. C. Singh B. N. Pandey
Harish Chandra P. G. College *Banaras Hindu University*

ABSTRACT The present paper investigates the properties of some shrunken testimators for the variance of a normal distribution in double stage samples, when initial estimate of variance σ^2 is available in the form of point estimate σ_0^2 and their risks are compared with an improved estimator under the asymmetric loss function.

Keywords Normal distribution; Shrunken testimator; Shrunken factor; LINEX loss function and level of significance.

1. Introduction

Let $x_{1i}\ (i=1,2,\cdots,n_1)$ and $x_{2j}\ (j=1,2,\cdots,n_2)$ be the two independent random samples of size n_1 and n_2 respectively, drawn from a normal distribution with mean μ and variance σ^2. The unbiased estimator for σ^2 based on sample *i* of size n_i, *i* = 1, 2, is

$$s_i^2 = \frac{1}{v_i}\sum_{j=1}^{n_i}(x_{ij}-\bar{x}_i)^2$$

with $MSE(s_i^2) = 2\sigma^4 / v_i$, where $v_i = n_i - 1$. The pooled unbiased estimator for σ^2 based on two samples of size n_1 and n_2 is

$$S^2 = \frac{v_1 s_1^2 + v_2 s_2^2}{v_1 + v_2}$$

with $MSE(S^2) = 2\sigma^4 /(v_1 + v_2)$.

□ Received December 2006, revised November 2007, in final form February 2008.
□ Gyan Prakash (corresponding author; email: ggyanji@yahoo.com) and D. C. Singh are affiliated to the Department of Statistics at Harish Chandra P. G. College, Varanasi, India. B. N. Pandey is affiliated to the Department of Statistics at Banaras Hindu University, Varanasi, India.

 ISSN 1726-3328

Journal of Probability and Statistical Science
6(2), 155-168, Aug. 2008

A Note on the Maximum Likelihood Box-Cox Transformation Parameter

Mezbahur Rahman and Larry M. Pearson
Minnesota State University

ABSTRACT The Box-Cox transformation is a well-known family of power transformations that brings a set of data into agreement with the normality assumption of the residuals and hence the response variable of a postulated model in regression analysis. This paper points out a potential innocent mistake in maximizing the likelihood in obtaining the transformation parameter using the Newton-Raphson method and gives an accurate algorithm while using the Newton-Raphson root finding method. In addition, a workable bound of the transformation parameter is suggested which can be used in the grid search method to obtain a more accurate estimate.

Keywords Moments for the ordered standard normal variates; Normality tests; Shapiro-Wilk *W* statistic.

1. Introduction

In regression analysis, often the key assumption regarding normality of the error variable and hence the response variable are violated. The commonly used remedy is the Box-Cox family of power transformations (see [2]). The process is to select a parameter in the Box-Cox transformation which maximizes the normal likelihood using the data at hand and then apply regression analysis on the transformed response variable. There is no role of the estimates of the location and the scale parameters which were derived in the process of estimating the power transformation parameter in regression analysis. In practice, the regression model parameters are usually estimated separately after the necessary Box-Cox power transformation parameter is selected.

□ Received December 2007, revised April 2008, in final form May 2008.
□ Authors of this paper are affiliated with the Department of Mathematics and Statistics at Minnesota State University, Mankato, MN 56001, USA; emails: mezbahur.rahman@mnsu.edu and larry.pearson@mnsu.edu.

 ISSN 1726-3328

Journal of Probability and Statistical Science
6(2), 169-185, Aug. 2008

Simplified Estimating Functions for Discretely Sampled IG-OU Processes

Shibin Zhang
Shanghai Maritime University

Xinsheng Zhang
Fudan University

ABSTRACT The set of IG-OU processes is an important subclass of non-Gaussian processes of Ornstein-Uhlenbeck type, which are recommended to model the volatility of financial assets. After proving self-decomposability of the transition distribution of IG-OU processes, conditional moments are obtained. For the discretely sampled IG-OU process, we propose to use an estimating function, which is based on both a simple estimating function and a martingale estimating function, to estimate the parameter. The simple estimating function is constructed by the score function of the marginal distribution of the process, and the martingale estimating function is constructed by using properties of conditional moments of the process. The consistency and asymptotic normality of the estimator are also discussed. In addition, the theoretical study is validated by simulations.

Keywords Asymptotic normality; Consistency; Estimating function; Inverse Gaussian; Process of Ornstein-Uhlenbeck type; Self-decomposability.

1. Introduction

Non-Gaussian processes of Ornstein-Uhlenbeck type form an important subclass of Markov processes with jumps. Certain of these processes are recently introduced into the financial literature by Barndorff-Nielsen and Shephard [2, 3]. In this paper, we consider the stochastic volatility process suggested in [2], which is given by the solution to the following differential equation:

$$\mathrm{d}X(t) = -\lambda X(t)\mathrm{d}t + \mathrm{d}Z(\lambda t), \qquad (1.1)$$

□ Received November 2006, revised January 2008, in final form June 2008.
□ Shibin Zhang is an Assistant Professor in the Department of Mathematics at Shanghai Maritime University, Shanghai 200135, P. R. China; email: sbzhang@dbc.shmtu.edu.cn. Xinsheng Zhang is a Professor in the Department of Statistics at Fudan University, Shanghai 200433, P. R. China; email: xszhang@fudan.edu.cn.
□ AMS 2000 Subject Classifications: 62M05, 60G17.

 ISSN 1726-3328

Journal of Probability and Statistical Science
6(2), 187-200, Aug. 2008

On a Matrix-Variate Generalized Type-1 Dirichlet Model

Seemon Thomas Alex Thannippara A. M. Mathai
St. Thomas College *McGill University*

ABSTRACT A generalized matrix-variate Dirichlet integral over a random matrix and the associated probability density are examined and a large number of properties of the density are studied. The extension considered is different from the extensions of Dirichlet proba- bility model considered by other authors. The model proposed is shown to be mathematically and statistically interesting with potential for applications in many areas.

Keywords Random matrix; Generalized Dirichlet model; Real matrix-variate distributions; Jacobians of matrix transformations.

1. Introduction

The standard real scalar type-1 Dirichlet density with the parameters $(\alpha_1,\cdots,\alpha_k;\alpha_{k+1})$ is given by

$$f(x_1,\cdots,x_k)=c_k x_1^{\alpha_1-1}\cdots x_k^{\alpha_k-1}(1-x_1-\cdots-x_k)^{\alpha_{k+1}-1}, \quad (1)$$

for $x_i\geq 0,\ i=1,\cdots,k,\ x_1+\cdots+x_k\leq 1,\ \Re(\alpha_j)>0,\ j=1,\cdots,k+1$ and $f_1(x_1,\cdots,x_k)=0$ elsewhere, where $\Re(\cdot)$ denotes the real part of $(\cdot)$. In statistical problems, usually the parameters are real. Then the conditions reduce to $\alpha_j>0,\ j=1,\cdots,k+1$. Since the results are abailable for the parameters in the complex domain, we will keep the conditions $\Re(\alpha_j)>0,\ j=1,\cdots,k+1$. The normalizing constant c_k can be evaluated by integrating out variables one at a time and it is such that

$$c_k^{-1}=\frac{\Gamma(\alpha_1)\cdots\Gamma(\alpha_{k+1})}{\Gamma(\alpha_1+\cdots+\alpha_{k+1})}.$$

□ Received July 2007, revised March 2008, in final form June 2008.
□ Seemon Thomas (email: seemon@stcp.ac.in) and Alex Thannippara are affiliated with the Department of Statistics at St. Thomas College, Palai, Mahatma Gandhi University Arunapuram P. O., Kerala, India. A. M. Mathai is affiliated with McGill University, Montreal, Canada and Center for Mathematical Sciences, Trivandrum and Pala Campuses, India.

 ISSN 1726-3328

Journal of Probability and Statistical Science
6(2), 201-214, Aug. 2008

Autoregressive Processes of Order *p* with Values on the Positive Half-Line

Nadjib Bouzar
University of Indianapolis

ABSTRACT We introduce a class of $\mathbf{R}_+$-valued autoregressive processes of order *p* (AR(*p*)) by using the generalized multiplication $\odot_C$ of van Harn and Steutel [9]. Stationary AR(*p*) processes with Mittag-Leffler-type marginal distributions are developed. It is shown in particular that these processes are uniquely defined if one assumes a specialized form of Self-decomposability on the marginal distribution. An $\mathbf{R}_+$-valued stationary AR(*p*) process with a compound-exponential marginal is presented. Related integer-valued AR(*p*) processes are also discussed.

Keywords Autoregressive; Stationarity; Semigroup; Self-decomposability; Mittag-Leffler distribution.

1. Introduction

Lawrance and Lewis [11] introduced the autoregressive processes of order *p* (AR(*p*)):

$$X_n = \begin{cases} \eta_1 X_{n-1} + \varepsilon_n & \text{with probability } c_1, \\ \eta_2 X_{n-2} + \varepsilon_n & \text{with probability } c_2, \\ \quad\cdots & \qquad\cdots \\ \eta_p X_{n-p} + \varepsilon_n & \text{with probability } c_p, \end{cases} \tag{1}$$

where $(\varepsilon_n, n \in \mathbf{Z})$ is a sequence if iid $\mathbf{R}_+$-valued rv's, $0 < \eta_i < 1$, $0 \le c_i \le 1$, i = 1, 2, …, p and $c_1 + \cdots + c_p = 1$. The authors obtained the distribution of the innovation sequence $(\varepsilon_n, n \in \mathbf{Z})$ for the stationary AR(2) process with an exponential marginal as a solution to a functional equation. Using the

Received July 2007, revised December 2007, in final form January 2008.
Nadjib Bouzar is affiliated with the Department of Mathematics and Computer Science at the University of Indianapolis, Indianapolis, IN 46227,USA; email: nbouzar@uindy.edu.
Research partially supported by a research grant from the College of Arts and Sciences, University of Indianapolis.

 ISSN 1726-3328

Journal of Probability and Statistical Science
6(2), 215-226, Aug. 2008

Simulation of the Log-Likelihood Random Variables with Applications

Amir Naderi
Isfahan University of Technology

ABSTRACT Likelihood functions guide the analysis in almost all areas of statistical applications. The distribution of the log-likelihood random variable is developed for the location, the scale and the location-scale normal models and algorithms for generating random numbers from these distributions are provided. Simulation results, using the R language, are presented and a new method of likelihood based confidence region construction is developed.

Keywords Likelihood function; Log-likelihood function; Sample distributions; Random number generators; The R language; The duality between tests and confidence regions.

1. Introduction

With model $f_\theta(x)$, the observed likelihood function from x_0 is $cf_\theta(x_0)$ where c is an arbitrary positive constant; this can be formalized in terms of an equivalence class $L_\theta(x_0) = \{cf_\theta(x_0), c \in R^+\}$ of functions on the parameter space, say Ω. One way of avoiding the nuisance constant c is to choose a representative from the equivalence class. For example, if there is a θ_0 in Ω for which $f_{\theta_0}(x) > 0$ for all sample points, then we can use the explicit representative $\ell_\theta(x_0) = f_\theta(x_0) / f_{\theta_0}(x_0)$ in place of the equivalence class, see Fraser [3]. Barndorff-Nielsen *et al*. [1], in a similar fashion, use the Radon-Nikodym derivative $q_\theta(x_0)$ with respect to a finite dominating measure as an appropriate expression of the likelihood function. McDunnough and Naderi [5] have shown that $L_\theta(x_0)$ and $q_\theta(x_0)$ (and in a similar fashion $\ell_\theta(x_0)$) induce the same partition of the sample space as that used by Lehmann and Scheffé [4] in their construction of the minimal sufficient statistics. If X is a random vector that

---

□ Received March 2007, revised January 2008, in final form February 2008.
□ Amir Naderi is an Associate Professor in the Department of Mathematical Sciences at Isfahan University of Technology, Isfahan 84156, Iran; email: amir@cc.iut.ac.ir.

 ISSN 1726-3328

Journal of Probability and Statistical Science
6(2), 227-242, Aug. 2008

A Discrete Linear Trend-Change Model for Unequally Spaced Data

Tze-San Lee; *Western Illinois University*
Arjun K. Gupta; *Bowling Green State University*
Jin Xu; *East China Normal University*

ABSTRACT Here a change-point model is presented for detecting changes in the linear trend of unequally spaced time-dependent data. Unobservable errors in the data are modeled by the Ornstein-Uhlenbeck process. No assumptions are made about the unknown parameters in the model. All unknown parameters in the model are estimated by employing the method of maximum likelihood estimation. Testing the significance of the unknown change-point is accomplished by using the likelihood ratio test. Two real world data sets, the U.S. graduate enrollments and the growth of body mass index for the U.S. preschool children, are analyzed which illustrates the usefulness of the proposed model.

Keywords Body mass index; Change-point; Graduate enrollment; Likelihood ratio test; Linear trend; Ornstein-Uhlenbeck process.

1. Introduction

Time-trend analysis is a very important tool for studying economic time series or investigating epidemiologic issues. For example, a need for trend detection and forecasting is desirable by medical doctors working in the renal intensive therapy unit where relatively sudden and unexpected deaths have occurred during the night when there is less medical and nursing cover (Hill-Endresen [14]). Moreover, if data are collected over an extended period, in either cross sectional or prospective studies, time trends in laboratory measurements can have

---

□ Received December 2007, revised May 2008, in final form June 2008.

□ Tze-San Lee was a Professor in the Department of Mathematics at Western Illinois University, Macomb, IL 61455, USA; Present address: CDC/NCEH, Mail Stop F-58, Chamblee, GA 30341, USA; e-mail: tjl3@cdc.gov. Arjun K. Gupta is a Distinguished Professor in the Department of Mathematics and Statistics at Bowling Green State University, Bowling Green, OH 43403, USA; e-mail: gupta@bgsu.edu. Jin Xu is an Associate Professor in the Department of Statistics and Actuarial Science at East China Normal University, Shanghai 200241, China; e-mail: jxu@stat.ecnu.edu.cn.

 ISSN 1726-3328

Journal of Probability and Statistical Science
6(2), 243-257, Aug. 2008

Recurrence Relations for Moment Generating Functions of Generalized Order Statistics from Doubly Truncated Continuous Distributions

Abd EL-Baset A. Ahmad
University of Assiut

Tahani A. Abu-Shal
Girls College of Science

ABSTRACT In this paper, a recurrence relation for joint moment generating functions of generalized order statistics, based on a random sample drawn from a class of doubly truncated continuous distributions, is obtained. Joint moments of generalized order statistics (ordinary order statistics and k-records) have been obtained. Recurrence relations for single and product moment generating functions and joint moments of generalized order statistics have been derived. Results for joint moment generating functions and joint moments of generalized order statistics from a number of doubly truncated distributions such as three parameters inverted Weibull (exponential, Rayleigh as special cases), logistic, extreme-value, inverted Pareto, among others, can be derived as special cases of our results. Specializations to right, left and non-truncated classes of continuous distributions have been made.

Keywords Generalized order statistics; Order statistics; K-records; Moment generating function; Moments; Truncated distributions.

1. Introduction

Kamps [15] introduced the concept of *generalized order statistics* (*gos's*). Ordinary order statistics (*oos's*), k-records (ordinary record values (*orv's*) when k = 1), sequential order statistics, ordering via truncated distributions and censoring schemes can be discussed as special cases of the *gos's*. Kamps's book [15] gave several applications such as a variety disciplines, recurrence

---

□ Received March 2007, revised January 2008, in final form February 2008.

□ Abd EL-Baset A. Ahmad is affiliated to the Department of Mathematics and Statistics at Taif University, Saudi Arabia; permanent address: Department of Mathematics, University of Assiut, Egypy; email: abahmad2002@yahoo.com. Tahani A. Abu-Shal is affiliated to the Mathematics Department at Girls College of Science, Jeddah, Saudi Arabia.

 ISSN 1726-3328

Journal of Probability and Statistical Science
6(2), 259-269, Aug. 2008

An Improved Software Failure Model Based on Non-homogeneous Poisson Processes

Syed A. Hossain
Rider University

ABSTRACT Yamada *et al.* [15] suggested a stochastic model for a software failure phenomenon based on non-homogeneous Poisson process (NHPP). The model suffers form the problem of improper probability density function for the time to failure. The main concern of this paper is to address this issue. A modification of it is proposed which rids the model of the problem of improper density function. Parameters of the model are estimated by maximum likelihood (ML) estimation method. Necessary and sufficient conditions for the existence of solution of maximum likelihood equations are derived. The results of the model are applied to real-time software failure data and compared with Goel-Okumoto (G-O), Modified G-O and Yamada *et al.* models. Three real life examples are considered. In all these examples the modified Yamada *et al.* model performed better than the Yamada *et al* model.

Keywords Software reliability, NHPP, Maximum likelihood estimation, Improper probability density function.

1. Introduction

Yamada, Ohaba and Osaki [15] offered a NHPP model in which the observed growth curve of the time to failures is S-shaped. This model is equivalent to other exponential class of NHPP models like Jelinski and Moranda [6] (J-M), Musa [10], Schneidewind [14], G-O [4], Hossain and Dahiya [5], and Dahiya-Hossain [3]. One of the basic NHPP models is the G-O model in which the failure detection rate is a constant. In the Yamada *et al.* model it is an increasing function of time. Ohba [13] model describes a software failure process with a mutual dependence of detected failures. In this process, as more and more failures are detected, more

□ Received December 2007, revised June 2008, in final form July 2008.
□ Syed A. Hossain is an Associate Professor in the Management Science Department at Rider University, Lawrenceville, NJ 08648, USA; e-mail: shossain@rider.edu.

 ISSN 1726-3328

Journal of Probability and Statistical Science
6(2), 271-280, Aug. 2008

On Size-Biased Generalized Negative Binomial Distribution

Anwar Hassan
University of Kashmir

Khurshid Ahmad Mir
Amar Singh College

ABSTRACT In this paper, a size-biased generalized negative binomial distribution (SBGNBD) is defined and studied. A recurrence relationship for the moments of SBGNBD is established. The Bayes estimator for a parametric function of one parameter when other two parameters are known of size-biased generalized negative binomial distribution is derived. The prior information on one parameter may be given by a beta distribution. It has been illustrated that the parameters in the prior distribution can be assigned by a computer using Monte Carlo and R-software.

Keywords Generalized negative binomial distribution; Size-biased generalized negative binomial distribution; Zero-truncated generalized negative binomial distribution; Size-biased negative binomial distribution; Goodness of fit; Bayes estimation.

1. Introduction

Jain and Consul [21] defined generalized negative binomial distribution (GNBD). It was subsequently obtained by Consul and Shenton ([7], [8]) as a particular family of Lagrangian distribution. The parameter space of the distribution was further modified by Consul and Gupta [6] when they studied some of its interesting properties. The probability function of the GNBD is given by

$$P_1(X=x)=\frac{m}{m+\beta x}\binom{m+\beta x}{x}\alpha^x(1-\alpha)^{m+\beta x-x};\ x=0,1,2,\cdots \tag{1.1}$$

where $0<\alpha<1$, $m>0$ and $|\alpha\beta|<1$.

□ Received June 2006, revised September 2007, in final form April 2008.

□ Anwar Hassan is a Reader in the Department of Statistics at the University of Kashmir, Srinagar (J&K), India; e-mail: anwar.hassan2007@rediffmail.com. Khurshid Ahmad Mir is a Senior Lecturer in the Department of Statistics at Amar Singh College, Srinagar (J&K) India; e-mail: khrshdmir@yahoo.com.

 ISSN 1726-3328