Contents

چهل و پنجمین

کنفرانس ریاضی ایران

۴-۷ شهریور ۱۳۹۳

آمار، احتمال و فرآیندهای تصادفی

Estimation for the Burr Type III Distribution Based on Record Values

M.Abdi *Academic member of School of Mathematics and Computing, Higher Education complex of Bam, me.abdi@bam.ac.ir.

A. Asgharzadeh, Academic member of Department of Statistics, University of Mazasandaran, a.asgharzadeh@umz.ac.ir

Abstract: In this paper, we study the estimation problem for the Burr type III distribution based on record values. The maximum likelihood method is used to derive the point estimators of the parameters. An exact confidence interval and a joint confidence region are constructed for the parameters of Burr type III distribution based on record data. A numerical example is presented to illustrate the proposed methods.
Keywords: Confidence Interval, Joint Confidence Region, Maximum Likelihood Estimation, Record Values.

1 INTRODUCTION

Let $X_1, X_2, ...$ be a sequence of independent and identically distributed (iid) continuous random variables with cdf $F(x)$ and pdf $f(x)$. An observation X_j is called an upper (lower) record value of this sequence if its value exceeds (is lower than) that of all previous observations. Generally, let us define $T_1 = 1,\ U_1 = X_1$, and for $n \geq 2$

$$T_n = \min\{j > T_{n-1} : X_j > X_{T_{n-1}}\}, \quad U_n = X_{T_n}. \tag{1}$$

Then the sequence $\{U_n\}(\{T_n\})$ is known as upper record statistics (upper record times). Similarly, the lower record times S_n and the lower record values L_n are defined as follows: $S_1 = 1,\ L_1 = X_1$, and for $n \geq 2,\ S_n = \min\{j > S_{n-1} : X_j < X_{S_{n-1}}\},\ L_n = X_{S_n}$.

The statistical study of record statistics started with Chandler (1952) and has now spread in different directions. Record data arise in a wide variety of practical situations and there are several situations pertaining to meteorology, hydrology, sporting and athletic events wherein only record values may be recorded. For more details and applications in the record values, see Arnold et al. (1998).

In this paper, we restrict attention to the Burr type III distribution. The probability density function (pdf) and cumulative distribution function (cdf) of the two-parameter Burr type III distribution are given, respectively, by

$$F(x;\theta,c) = (1 + x^{-c})^{-\theta}, \quad x, \theta, c > 0, \tag{2}$$

and

$$f(x;\theta,c) = \theta c x^{-(c+1)}(1 + x^{-c})^{-(\theta+1)} \quad x, \theta, c > 0. \tag{3}$$

Interval estimation of the parameters of the Burr type III distribution has not yet been studied based on record values. The main purpose of this paper is to construct an exact confidence interval c

*Corresponding Author

and an exact joint confidence region for (c, θ) based on lower record values.

2 Point Estimation

Suppose we observe m lower record values $X_{L(1)} = x_1, X_{L(2)} = x_2, \cdots, X_{L(m)} = x_m$ from the Burr type III distribution with pdf (3). The Log-likelihood function is given by

$$\begin{aligned} \ln L(c, \theta) &= m \ln c\, \theta - \theta \ln(1 + x_m^{-c}) \\ &- (c+1) \sum_{i=1}^{m} \ln x_i - \sum_{i=1}^{m} \ln(1 + x_i^{-c}) \end{aligned} \tag{4}$$

Taking the derivative with respect to c and θ and equating to zero, we obtain the likelihood equations for c and θ as

$$\frac{dL(c,\theta)}{d\theta} = \frac{m}{\theta} - \ln(1 + x_m^{-c}) = 0, \tag{5}$$

$$\begin{aligned} \frac{dL(c,\theta)}{dc} &= \frac{m}{c} + \theta \frac{x_m^{-c} \ln x_m}{(1 + x_m^{-c})} - \sum_{i=1}^{m} \ln x_i \\ &+ \sum_{i=1}^{m} \frac{x_i^{-c} \ln x_i}{(1 + x_i^{-c})} = 0, \end{aligned} \tag{6}$$

From (5), we obtain the maximum likelihood estimate (MLE) of θ as a function of c, say $\widehat{\theta}(c)$, as

$$\widehat{\theta}(c) = \frac{m}{\ln(1 + x_m^{-c})} \tag{7}$$

Substituting $\widehat{\theta}(c)$ in (6) we have

$$\frac{m}{c} + \frac{m\, x_m^{-c} \ln x_m}{(1 + x_m^{-c}) \ln(1 + x_m^{-c})} - \sum_{i=1}^{m} \ln x_i + \sum_{i=1}^{m} \frac{x_i^{-c} \ln x_i}{(1 + x_i^{-c})} = 0 \tag{8}$$

Since (8) can not be solved analytically with respect to c, some numerical methods such as Newton-Raphson or fix-point methods should be used to find the MLE of c ($\widehat{c}_{MLE}$). It can be shown that the solution of (8) can be obtained as a fixed point solution of the following equation

$$h(c) = c, \tag{9}$$

where, $h(c)$ is given by

$$\begin{aligned} h(c) &= m \left[\sum_{i=1}^{m} \ln x_i - \frac{m\, x_m^{-c} \ln x_m}{(1 + x_m^{-c}) \ln(1 + x_m^{-c})} \right. \\ &- \left. \sum_{i=1}^{m} \frac{x_i^{-c} \ln x_i}{(1 + x_i^{-c})} \right]^{-1}. \end{aligned} \tag{10}$$

We apply iterative procedure to find the solution of (9). Once $\hat{c}_{MLE}$ is obtained, the MLE of θ, say $\hat{\theta}_{MLE}$, can be obtained from (8) as $\hat{\theta}_{MLE} = \hat{\theta}(\hat{c}_{MLE})$.

3 Confidence Interval and Joint Confidence Region

Let$X_{L(1)} > X_{L(2)} > \cdots > X_{L(m)}$ be the first m observed lower record values from the Burr type III distribution. In this section, a $100(1-\alpha)\%$ confidence interval for parameter c and a $100(1-\alpha)\%$ joint confidence region for (c, θ) are constructed based on the observed lower records. For notation simplicity, we will write X_i for $X_{L(i)}$. Let $Y_i = -\ln F[X_i] = \theta \ln[1 + X_i^{-c}]$, $i = 1, ..., m$. It can be shown that that $Y_1 < Y_2 < \cdots < Y_m$ are the first m upper record values from a standard exponential distribution. Moreover, the spacings $Z_1 = Y_1, Z_2 = Y_2 - Y_1, \cdots, Z_m = Y_m - Y_{m-1}$ are iid random variables from a standard exponential distribution (see Arnold et al. (1998)). Hence

$$\begin{aligned} V &= 2Z_1 = 2\, Y_1, \\ U &= 2 \sum_{i=2}^{m} Z_i = 2\, (Y_m - Y_1), \end{aligned}$$

respectively have a χ^2 distribution with 2 degrees of freedom and a χ^2 distribution with $2m-2$ degrees of freedom. We can also find that U and V are independent random variables. Let

$$\begin{aligned} P_1 = \frac{U/2(m-1)}{V/2} &= \frac{U}{(m-1)V} \\ &= \frac{1}{m-1} \left(\frac{Y_m - Y_1}{Y_1} \right) \end{aligned} \tag{11}$$

and

$$P_2 = U + V = 2Y_m. \tag{12}$$

It is easy to show that P_1 has an F distribution with $2m-2$ and 2 degrees of freedom and P_2 has a chi-square distribution with $2m$ degrees of freedom. Furthermore, P_1 and P_2 are independent, see Johnson et al. (1994, P. 350).

To derive the exact confidence interval for c and the exact joint confidence region for (c,θ), we need the following two lemmas.

Lemma 3.1. *For any $0 < x_m < x_1 < \infty$, the function*

$$g(c) = \frac{\ln(1+x_m^{-c})}{\ln(1+x_1^{-c})},$$

is a strictly increasing function of c for any $c > 0$.

Lemma 3.2. *Suppose that $0 < x_m < x_1 < \infty$. Then for any $m > 1$,*
(1) The function $P_1(c) = \frac{1}{m-1}\left[\frac{\ln(1+x_m^{-c})}{\ln(1+x_1^{-c})} - 1\right]$ is strictly increasing in $c > 0$.
(2) For $x_1 \geq 1$, and any $t > 0$, the equation $P_1(c) = t$ has a unique solution for some $c > 0$.
(3) For $x_1 < 1$ and any $0 < t < \frac{1}{m-1}\left[\frac{\ln(x_m)}{\ln(x_1)} - 1\right]$, the equation $P_1(c) = t$ has a unique solution for some $c > 0$.

Let $F_\alpha(\upsilon_1, \upsilon_2)$ be the percentile of F distribution with right-tail probability α and υ_1 and υ_2 degrees of freedom and $\chi^2_\alpha(\upsilon)$ denote the percentile of χ^2 distribution with right-tail probability α and υ degrees of freedom.

Next theorems gives an exact confidence interval for parameter c and an exact joint confidence region for the parameters c and θ.

Theorem 3.3. *Suppose that $\underline{X} = (X_1, X_2, \ldots, X_m)$ be the first m observed lower record values from the Burr type III distribution in (2). Then, for any $0 < \alpha < 1$,*

$$\left(\varphi[\underline{X}, F_{1-\frac{\alpha}{2}}(2m-2,2)], \varphi[\underline{X}, F_{\frac{\alpha}{2}}(2m-2,2)]\right)$$

is a $100(1-\alpha)\%$ confidence interval for c, where $\varphi(\underline{X}, t)$ is the solution of c for the equation

$$\frac{1}{m-1}\left[\frac{\ln(1+x_m^{-c})}{\ln(1+x_1^{-c})} - 1\right] = t$$

Theorem 3.4. *Suppose that $\underline{X} = (X_1, X_2, \ldots, X_m)$ be the first m observed lower record values from the Burr type III distribution in (2). Then, the following inequalities determine $100(1-\alpha)\%$ joint confidence region for (c,θ):*

$$\varphi[\underline{X}, F_{\frac{1+\sqrt{1-\alpha}}{2}}(2m-2,2)] < c < \varphi[\underline{X}, F_{\frac{1-\sqrt{1-\alpha}}{2}}(2m-2,2)]$$

$$\frac{\chi^2_{\frac{1+\sqrt{1-\alpha}}{2}}(2m)}{2\ln(1+x_m^{-c})} < \theta < \frac{\chi^2_{\frac{1-\sqrt{1-\alpha}}{2}}(2m)}{2\ln(1+x_m^{-c})},$$

where $0 < \alpha < 1$, and $\varphi(\underline{X}, t)$ is the solution of c for the equation

$$P_1(c) = \frac{1}{m-1}\left[\frac{\ln(1+x_m^{-c})}{\ln(1+x_1^{-c})} - 1\right] = t$$

4 Numerical Example

In this example we consider one real life data set to illustrate the proposed methods of estimation. The data represent 24 observations on the period between successive earthquakes in the last century in North Anatolia fault zone. These data are analyzed by Kus (2007).

```
1163, 501, 2039, 4863, 3258, 616, 217, 143,

323, 398, 9, 182, 159, 67, 633, 2117, 756,

896, 461, 3709, 409, 8592, 1821, 979
```

Here, we checked the validity of the Burr type III model based on the parameters $\hat{c} = 0.59785, \hat{\theta} = 30.08471$ (MLEs of parameters), using the Kolmogorov-Smirnov (K-S) test. It is observed that the K-S distance is $K-S = 0.145263$ with a corresponding $p-value = 0.6569509$. So, the Burr type III model provides a good fit to the above data. For the above data, we observe the following five lower record values:

```
1163, 501,  217, 143, 9.
```

Using the formula described in Section 2, we obtain the MLEs of the parameters c and θ to be $\hat{c}_{MLE} = 0.36887$ and $\hat{\theta}_{MLE} = 13.5921$, respectively. By Theorem 3.3 and using the S-PLUS package, the 95% confidence interval for c is $(0.12896, 1.05174)$ with confidence length 0.92278. By Theorem 3.4 and using the S-PLUS package for solving non-linear equation, the 95% joint confidence region for c and θ is determined by the following inequalities:

$$0.10591 < c < 1.18956,$$
$$\frac{0.12634}{2\ln(1+x_m^{-c})} < \theta < \frac{78.11416}{2\ln(1+x_m^{-c})}.$$

with area 60.93374. Figure 1, shows the 95% joint confidence region for (c, θ).

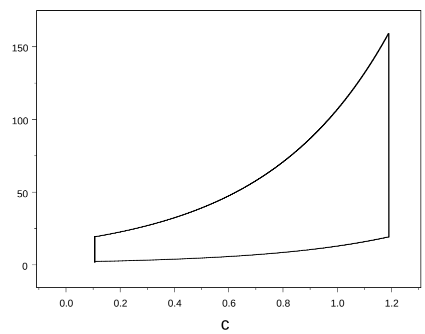

Figure 1: Joint confidence region for c and θ

References

[1] B. C. Arnold, Balakrishnan, N. and Nagaraja, H. N. Records, John Wiley & Sons, New York, 1998.

[2] K.N. Chandler, "The distribution and frequency of record values," Journal of Royal Statist, Soc., pp. 220-228, 1952.

[3] N.L. Johnson, Kotz, S., and Balakrishnan, N., Continuous Univariate Distribution, John Wiley and Sons, Inc., New York, 1994.

[4] C. Kus, "A new lifetime distribution," Comput. Statist. Data Anal., (9), pp. 44974509, 2007.

The 45th Annual Iranian Mathematics Conference
August 26-29 2014
چهل و پنجمین
کنفرانس ریاضی ایران
۴ - ۷ شهریور ۱۳۹۳

A New Generalization of the Lindley Distribution

E. Mahmoudi*
Academic member of School of Mathematics, Yazd University
emahmoudi@yazd.ac.ir
S. Abolhosseini
M. Sc. student of School of Mathematics, Yazd University
s-abolhoseini@yahoo.com

Abstract: In this paper we introduce the new generalization of the Lindely distribution and obtain several properties of the new distribution such as its probability density function, its reliablity and failure rate functions and moments. EM-algorithm is presented in this paper.
Keywords: Bonferroni and Lorenze curves, EM-algorithm, Power series, Survival function.

1 INTRODUCTION

Recently, attempts have been made to define new families of probability distributions that extend well-known families of distributions and at the same time provide great flexibility in modeling data in practice. The exponential-geometric (EG), exponential-Poisson (EP), exponential-logarithmic (EL), exponential-power series (EPS), Weibull-geometric (WG) and Weibull-power series (WPS) distributions were introduced and studied by Adamidis and Loukas [1], Kus [8], Tahmasbi and Rezaei [16], Chahkandi and Ganjali [5], Barreto-Souza et al. [3] and Morais and Barreto-Souza et al. [15], respectively.

Barreto-Souza and Cribari-Neto [2] and Louzada et al. [9] introduced the exponentiated exponential-Poisson (EEP) and the complementary exponential-geometric (CEG) distributions where the EEP is the generalization of the EP distribution and the CEG is complementary to the EG model proposed by Adamidis and Loukas [1]. Recently, Cancho et al. [4] introduced the two-parameter Poisson-exponential (PE) lifetime distribution with increasing failure rate. Mahmoudi and Jafari [10] introduced the generalized exponential-power series (GEPS) distribution by compounding the generalized exponential (GE) distribution with the power series distribution. Also exponentiated Weibull-logarithmic (EWL), exponentiated Weibull-geometric (EWG) and exponentiated Weibull-power series (EWP) distributions has been introduced and analyzed by Mahmoudi and Sepahdar [11] and Mahmoudi and Shiran [12, 13].

In this paper, we propose a new two-parameters distribution, referred to as the Lindley binomial (LB) distribution, which contains as special sub-models the Lindley.

The paper is organized as follows. In Section 2, we define the LB distribution. The density, survival and hazard rate functions of the new distribution is obtained in this section. We derive moments of the LB distribution in Section 3. Section 4 is devoted to the Bonferroni and Lorenze curves of the LB distribution. Estimation of the parameters by EM-algorithm and inference for large sample are presented in section 5.

*Corresponding Author

2 LINDLEY Binomial DISTRIBUTION

Suppose that the random variable X has the Lindley distribution where its cdf and pdf are given by

$$F_X(x) = 1 - (1 + \frac{\gamma x}{\gamma+1})e^{-\gamma x}, \quad x > 0, \tag{1}$$

$$f_X(x) = \frac{\gamma^2}{\gamma+1}(1+x)e^{-\gamma x}, \quad x > 0. \tag{2}$$

Given N, let $X_1, \cdots, X_N$ be independent and identify distributed random variables from Lindley distribution. Let N is distributed according to binomial distribution with pdf

$$P(N=n) = \frac{\binom{m}{n}\theta^n}{(\theta+1)^m - 1}, \; n = 1, 2, \cdots, \; 0 < \theta, \; n < m.$$

Let $Y = \min(X_1, \cdots, X_N)$, then the cdf of $Y|N = n$ is given by

$$F_{Y|N=n}(y) = 1 - ((1 + \frac{\gamma y}{\gamma+1})e^{-\gamma y})^n,$$

The Lindley binomial distribution, denote by LB(θ, γ), is defined by the marginal cdf of Y, i.e.

$$F(y) = \frac{(\theta+1)^m - (\theta(1+\frac{\gamma y}{\gamma+1})e^{-\gamma y} + 1)^m}{(\theta+1)^m - 1}. \tag{3}$$

The pdf of LB distribution is given by

$$f(y) = \frac{m\theta\frac{\gamma^2}{\gamma+1}e^{-\gamma y}(1+y)(\theta(1+\frac{\gamma y}{\gamma+1})e^{-\gamma y}+1)^{m-1}}{(\theta+1)^m - 1}, \tag{4}$$

where $\theta > 0, \gamma > 0$.
The survival and hazard rate functions of LB distribution are given, respectively, by

$$S(y) = \frac{(\theta(1+\frac{\gamma y}{\gamma+1})e^{-\gamma y}+1)^m - 1}{(\theta+1)^m - 1}, \tag{5}$$

and

$$h(y) = \frac{m\theta e^{-\gamma y}\frac{\gamma^2}{\gamma+1}(1+y)(\theta(1+\frac{\gamma y}{\gamma+1})e^{-\gamma y}+1)^{m-1}}{(\theta(1+\frac{\gamma y}{\gamma+1})e^{-\gamma y}+1)^m - 1}. \tag{6}$$

Proposition 2.1. *The limiting distribution of LB (θ, γ) where $\theta \to 0^+$ is*

$$\lim_{\theta \to 0^+} F(y) = 1 - (1 + \frac{\gamma x}{\gamma+1})e^{-\gamma x},$$

which is the cdf of Lindley distribution.

3 MOMENTS OF LB DISTRIBUTION

Some of the most important features and characteristics of a distribution can be studied through its moments such as tending, dispersion, skewness and kurtosis.We obtain the moment generating function of the LB distribution. Suppose that $Y \sim LB(\theta, \gamma)$ and $X_{(1)} = \min(X_1, \cdots, X_n)$, where $X_i \sim L(\gamma)$ for $i = 1, 2, \cdots, n$, then

$$\begin{aligned} M_X(t) &= \textstyle\sum_{n=1}^{\infty} P(N=n) M_{X_{(1)}}(t) \\ &= \textstyle\sum_{n=1}^{\infty} P(N=n) \sum_{i=0}^{n-1} \binom{n-1}{i} (\frac{\gamma}{\gamma+1})^{n-i} n\gamma \\ &\times \left[\frac{\Gamma(n-i)}{(n\gamma-t)^{n-i}} + \frac{\Gamma(n-i+1)}{(n\gamma-t)^{n-i+1}}\right] \\ &= \frac{\theta^n\binom{m}{n}}{(\theta+1)^m-1} \textstyle\sum_{n=1}^{\infty}\sum_{i=0}^{n-1}\binom{n-1}{i}\frac{\gamma^{n-i+1}}{(\gamma+1)^{n-i}} \\ &\times \left[\frac{\Gamma(n-i)}{(n\gamma-t)^{n-i}} + \frac{\Gamma(n-i+1)}{(n\gamma-t)^{n-i+1}}\right]. \end{aligned} \tag{7}$$

One can use $M_X(t)$ to obtain the kth moment about zero of the LB distribution. We have

$$\begin{aligned} E(Y^k) &= \textstyle\sum_{n=1}^{\infty} P(N=n) E(X_{(1)}^k) \\ &= \textstyle\sum_{n=1}^{\infty}\sum_{i=0}^{n-1}\binom{n-1}{i}\frac{\binom{m}{n}n\theta^n\gamma^{i+2}}{(\gamma+1)^{i+1}((\theta+1)^m-1)} \\ &\times \left[\frac{\Gamma(k+i+2)}{(n\gamma)^{k+i+2}} + \frac{\Gamma(k+i+1)}{(n\gamma)^{k+i+1}}\right]. \end{aligned} \tag{8}$$

The mean and variance of the LB distribution are given, respectively, by

$$\begin{aligned} E(Y) &= \textstyle\sum_{n=1}^{\infty}\sum_{i=0}^{n-1}\binom{n-1}{i} n \frac{\binom{m}{n}\theta^n\gamma^{i+2}}{(\gamma+1)^{i+1}((\theta+1)^m-1)} \\ &\times \left[\frac{\Gamma(i+3)}{(n\gamma)^{i+3}} + \frac{\Gamma(i+2)}{(n\gamma)^{i+2}}\right], \end{aligned} \tag{9}$$

and

$$\begin{aligned} Var(Y) &= \textstyle\sum_{n=1}^{\infty}\sum_{i=0}^{n-1}\binom{n-1}{i} n \binom{m}{n} \frac{\theta^n\gamma^{i+2}}{(\gamma+1)^{i+1}((\theta+1)^m-1)} \\ &\times \left[\frac{\Gamma(i+4)}{(n\gamma)^{i+4}} + \frac{\Gamma(i+3)}{(n\gamma)^{i+3}}\right] - E^2(Y), \end{aligned} \tag{10}$$

where $E(Y)$ is given in Eq. (9).

4 Bonferroni and Lorenz curves

Study of income inequality has gained a lot of importance over the last many years. Lorenz curve and the associated Gini index are undoubtedly the most popular indices of income inequality. However, there are certain measures which despite possessing interesting characteristics have not been used often for measuring inequality. Bonferroni curve and scaled total time on test transform are two such measures, which have the advantage of being represented graphically in the unit square and can also be related to the Lorenz curve and Gini ratio (Giorgi, [6]). These two measures have some applications in reliability and life testing as well (Giorgi and Crescenzi, [7]). The Bonferroni and Lorenz curves and Gini index have many applications not only in economics to study income and poverty, but also in other fields like reliability, medicine and insurance.

For a random variable X with cdf $F(.)$, the Bonferroni curve is given by

$$B_F[F(x)] = \frac{1}{\mu F(x)} \int_0^x u f(u) du.$$

From the relationship between the Bonferroni curve and the mean residual lifetime, the Bonferroni curve of the LB distribution is given by

$$\begin{aligned} B_F[F(x)] &= \tfrac{1}{\mu F(x)} \tfrac{m\theta(\gamma)^2}{((\theta+1)^m-1)(\gamma+1)} \sum_{i=0}^{m-1} \sum_{j=0}^{i} \\ &\times \binom{m-1}{i} \binom{i}{j} (\theta)^i (\tfrac{\gamma}{\gamma+1})^j \\ &\times \Big[\tfrac{1}{(\gamma(i+1))^{j+2}} \Gamma_{\gamma(i+1)x}(j+2) \\ &+ (\tfrac{1}{\gamma(i+1)})^{j+3} \Gamma_{\gamma(i+1)x}(j+2) \Big], \end{aligned}$$

where μ is the mean of the LB distribution.
The Lorenz curve of the LB distribution can be obtained via the expression $L_F[F(x)] = B_F[F(x)]F(x)$.
The scaled total time on test transform of a distribution function F is defined by

$$S_F[F(t)] = \frac{1}{\mu} \int_0^t S(u) du.$$

If $F(t)$ denotes the cdf of the LB distribution then

$$\begin{aligned} S_F[F(t)] &= \tfrac{1}{\mu((1+\theta)^m-1)} \Big[\sum_{i=0}^{m} \sum_{j=0}^{i} \binom{m}{i} \binom{i}{j} (\theta)^i \\ &\times (\tfrac{\gamma}{\gamma+1})^j (\tfrac{1}{\gamma i})^{j+1} \Gamma_{(j+1)} + t \Big]. \end{aligned}$$

The cumulative total time can be obtained by using formula $C_F = \int_0^1 S_F[F(t)] f(t) dt$ and the Gini index can be derived from the relationship $G = 1 - C_F$.

5 EM-algorithm

The MLEs of the parameters θ and γ must be derived numerically. Newton-Raphson algorithm is one of the standard methods to determine the MLEs of the parameters. To employ the algorithm, second derivatives of the log-likelihood are required for all iterations. The EM-algorithm is a very powerful tool in handling the incomplete data problem (McLachlan and Krishnan, [14]). It is an iterative method by repeatedly replacing the missing data with estimated values and updating the parameters. It is especially useful if the complete data set is easy to analyze.

Let the complete-data be $Y_1, \cdots, Y_n$ with observed values $y_1, \cdots, y_n$ and the hypothetical random variable $Z_1, \cdots, Z_n$. The joint probability density function is such that the marginal density of $Y_1, \cdots, Y_n$ is the likelihood of interest. Then, we define a hypothetical complete-data distribution for each (Y_i, Z_i) $i = 1, \cdots, n$, with a joint probability density function in the form

$$\begin{aligned} g(y, z; \Theta) &= f(y|z) f(z) = z \tfrac{\gamma^2}{\gamma+1} (1+y) e^{-z\gamma y} \\ &\times (1 + \tfrac{\gamma y}{\gamma+1})^{z-1} \tfrac{\binom{m}{z} \theta^z}{(\theta+1)^m - 1} \end{aligned}$$

where $\Theta = (\theta, \gamma)$, $y > 0$ and $z \in \mathbb{N}$.
Under the formulation, the E-step of an EM cycle requires the expectation of $(Z|Y; \Theta^{(r)})$ where $\Theta^{(r)} = (\theta^{(r)}, \gamma^{(r)})$ is the current estimate of Θ (in the rth iteration).
The pdf of Z given Y, say $g(z|y)$ is given by

$$g(z|y) = \frac{z e^{-\gamma(z-1)y} (1 + \frac{\gamma y}{\gamma+1})^{z-1} \binom{m}{z} \theta^{z-1}}{m(\theta(1 + \frac{\gamma y}{\gamma+1}) e^{-\gamma y} + 1)^{m-1}},$$

with the expectation

$$E[Z|Y = y] = 1 + \frac{\theta e^{-\gamma y}(1+\frac{\gamma y}{\gamma+1})(m-1)}{\theta e^{-\gamma y}(1+\frac{\gamma y}{\gamma+1})+1}.$$

The EM cycle is completed with the M-step by using the maximum likelihood estimation over Θ, with the missing Z's replaced by their conditional expectations given above.
The log-likelihood for the complete-data is

$$\begin{aligned} & l_n^*(y_1, \cdots, y_n; z_1, \cdots, z; \Theta) \\ & \propto n \log(\tfrac{\gamma^2}{\gamma+1}) - \sum_{i=1}^n \gamma z_i y_i \\ & + \sum_{i=1}^n (z_i - 1) \log(1 + \tfrac{\gamma y_i}{\gamma+1}) \\ & - n \log((\theta+1)^m - 1) + \sum_{i=1}^n z_i \log(\theta). \end{aligned}$$

The components of the score function

$$U_n^*(\Theta) = (\frac{\partial l_n^*}{\partial \theta}, \frac{\partial l_n^*}{\partial \gamma})^T$$

are given by

$$\begin{aligned} \frac{\partial l_n^*}{\partial \theta} &= \frac{\sum_{i=1}^n z_i}{\theta} - \frac{nm(\theta+1)^{m-1}}{(\theta+1)^{m-1}-1}, \\ \frac{\partial l_n^*}{\partial \gamma} &= \frac{n(\gamma+2)}{\gamma(\gamma+1)} - \sum_{i=1}^n z_i y_i \\ &\quad + \sum_{i=1}^n (z_i - 1)\frac{y_i}{(\gamma(1+y_i)+1)(\gamma+1)}. \end{aligned}$$

From a nonlinear system of equations $U_n^*(\Theta) = \mathbf{0}$, we obtain the iterative procedure of the EM-algorithm as

$$\hat{\theta}^{(t+1)} = \frac{((\hat{\theta}^{(t+1)}+1)^m - 1)\sum_n^{i=1} \hat{z_i}^t}{nm(\hat{\theta}^{(t+1)}+1)^{m-1}},$$

where $\hat{\theta}^{(t+1)}$ and $\hat{\gamma}^{(t+1)}$ are found numerically. Hence, for $i = 1, \cdots, n$, we have that

$$z_i^{(t)} = 1 + \frac{(\hat{\theta}^{(t)} e^{-\hat{\gamma}^t y}(1 + \frac{\hat{\gamma}^{(t)} y}{\hat{\gamma}^{(t)}+1}))(m-1)}{(\hat{\theta}^{(t)} e^{-\hat{\gamma}^t y}(1 + \frac{\hat{\gamma}^{(t)} y}{\hat{\gamma}^{(t)}+1}) + 1)}.$$

References

[1] K. Adamidis, and S. Loukas, "A lifetime distribution with decreasing failure rate", Statistics and Probability Letters, vol. 39, pp. 35-42, 1998.

[2] W. Barreto-Souza, and F. Cribari-Neto, "A generalization of the exponential-Poisson distribution", Statistics and Probability Letters, vol. 79, pp. 2493-2500, 2009.

[3] W. Barreto-Souza, A. L. Morais, and G. M. Cordeiro, "The Weibull-geometric distribution", Journal of Statistical Computation and Simulation, vol. 81, pp. 645-657, 2011.

[4] V. G. Cancho, F. Louzada-Neto, and G. D. C. Barriga, "The Poisson-exponential lifetime distribution", Computational Statistics and Data Analysis, vol. 55, pp. 677-686, 2011.

[5] M. Chahkandi, and M. Ganjali, "On some lifetime distributions with decreasing failure rate", Computational Statistics and Data Analysis, vol. 53, pp. 4433-4440, 2009.

[6] G.M. Giorgi, Concentration index, Bonferroni. Encyclopedia of Statistical Sciences, vol. 2, Wiley, New York, 1998, pp. 141-146.

[7] G.M. Giorgi, M. Crescenzi, A look at the Bonferroni inequality measure in a reliability framework. Statistica LXL 4

[8] C. Kus, "A new lifetime distribution", Computational Statistics and Data Analysis, vol. 51, pp. 4497-4509, 2007.

[9] F. Louzada, M. Roman, and V. G. Cancho, "The complementary exponential geometric distribution: model, properties, and comparison with its counter part", Computational Statistics and Data Analysis, vol. 55, pp. 2516-2524, 2011.

[10] E. Mahmoudi, and A. A. Jafari, "Generalized exponential-power series distributions", Computational Statistics and Data Analysis, vol. 56, pp. 4047-4066, 2012.

[11] E. Mahmoudi and A. Sepahdar, "Exponentiated Weibull-logarithmic distribution", submitted for publication.

[12] E. Mahmoudi, and M. Shiran, "Exponentiated Weibull-geometric distribution and its applications", submitted for publication, http://arxiv.org/abs/1206.4008v1.

[13] E. Mahmoudi, and M. Shiran, "Exponentiated Weibull-power series distribution", submitted for publication.

[14] G.J. McLachlan and T. Krishnan, The EM Algorithm and Extension,Wiley, NewYork, 1997.

[15] A. L. Morais, and W. Barreto-Souza, "A compound class of Weibull and power series distributions", Computational Statistics and Data Analysis, vol. 55, pp. 1410-1425, 2011.

[16] R. Tahmasbi and S. Rezaei, "A two-parameter lifetime distribution with decreasing failure rate", Computational Statistics and Data Analysis, vol. 52, pp. 3889-3901, 2008.

The 45th Annual Iranian Mathematics Conference
August 26-29 2014

Shannon Entropy in Order Statistics and Their Concomitants

M. Madadi*, Academic member of School of Mathematics,

Shahid Bahonar University, madadi@mail.uk.ac.ir

M. Naghavy, M.Sc, maryamnaghavy@gmail.com

Abstract: In this paper, we first derive two results on the Shannon entropy contained in the order statistics and their concomitants of a sequence of iid continuous random variables. We then compute this entropy for the general form of the Farlie-Gumbel-Morgenstern distribution.
Keywords: Concomitants of Order statistics; Farlie-Gumbel-Morgenstern Distribution; Shannon Entropy.

---

1 INTRODUCTION

Let $\{(X_i, Y_i) : i = 1, 2, ...\}$ be a sequence of bivariate random variables from a continuous distribution. If we arrange the X-values in ascending order, the corresponding Y-values are called the concomitants of the relevant order statistics. Concomitants of order statistics arise in several applications. In selection procedures, items or subjects may be chosen on the basis of their X characteristic, and an associated characteristic Y that is hard to measure or can be observed only later may be of interest. For example, X may be the score of a candidate on a screening test, and Y is the measure of his/her final performance. The first study of uncertainty (information) measure was undertaken by Nyquist (1924, 1924) and Hartley (1928), although they were introduced by Clausius in the year 1850 in the context of classical thermodynamics. Later Shannon (1946) studied the properties of information sources and the communication channels used to transmit their output, and defined an entropy known as Shannon entropy. For an absolutely continuous random variable X having pdf $f_X(x)$, the Shannon entropy is defined as

$$H(X) = -\int_{-\infty}^{+\infty} f_X(x) \ln f_X(x) dx$$

The Shannon entropy of a random variable X is a mathematical measure of information which measures the average reduction of uncertainty of X. Because of its descriptive character, analytical expressions for univariate distributions have been obtained, among others, by Lazo and Rathie (1978) and Cover and Thomas (1991). For multivariate distributions, formulas for the Shannon entropy have appeared in papers by Ahmed and Gokhale (1989) and Darbellay and Vajda (2000). Shannon entropy has been used to study the different emergent behaviors exhibited by the system in a cellu-

*Corresponding Author

The 45th Annual Iranian Mathematics Conference
August 26-29 2014
چهل و پنجمین کنفرانس ریاضی ایران
۴ - ۷ شهریور ۱۳۹۳

lar automata model. Mainly the laser spiking and the laser constant operation [8]. Recently several authors have investigated the Shannon entropy in record values (Madadi and Tata 2011). The concept of entropy can be successfully used to quantify the amount of information regarding the parent distribution that one may obtain by observing an additional record value. The organization of this article is as follows: In Section 2 we obtain general formulas for the Shannon entropy for order statistics and their concomitants and, in particular, for FGM distributions.

2 Entropy of The Order Statistics and Their concomitants

In this section we introduce order statistics and concomitants of order statistics and then compute relevant Shannon entropy.

2.1 Entropy of Order Statistics

Definition 2.1.1 Let $X_{1:n},...,X_{n:n}$ denote the order statistics of a random sample $X_1,...,X_n$ from a distribution with cdf $F_X(x)$ and pdf $f_X(x)$. Then the pdf of $X_{j:n}$; $1 \le j \le n$ is

$$\begin{aligned} f_{X_{j:n}}(x) &= \frac{n!}{(j-1)!(n-j)!}\Big(F_X(x)\Big)^{j-1} \\ &\quad \times\Big(1-F_X(x)\Big)^{n-j} f_X(x) \end{aligned}$$

and the joint pdf of $X_{i:n}$ and $X_{j:n}$; $1 \le i < j \le n$ is

$$\begin{aligned} f_{X_{i:n},X_{j:n}}(x,y) &= \frac{n!}{(i-1)!(j-i-1)!(n-j)!} \\ &\quad \times\Big(F_X(x)\Big)^{i-1} \\ &\quad \times\Big(1-F_X(y)\Big)^{n-j} \\ &\quad \times\Big(F_X(y)-F_X(x)\Big)^{j-i-1} \\ &\quad \times f_X(x)\ f_X(y). \end{aligned}$$

Lemma 2.1.1 The entropies of the j-th order statistic and joint order statistics are respectively given by

$$\begin{aligned} H(X_{j:n}) &= -\ln c_j + c_j(j-1)I(j:n) \\ &\quad + c_j(n-j)I(n-j+1:n) \\ &\quad + c_j\varphi_X(j:n) \\ H(X_{i:n},X_{j:n}) &= -\ln c_{i,j} + c_i(i-1)I(i:n) \\ &\quad + c_j(n-j)I(n-j+1:n) \\ &\quad + (j-i-1)\,[c_jI(j:n) \\ &\quad + (j-i)\binom{j-1}{i-1} \\ &\quad \times I(j-i:j-1)] \\ &\quad + c_i\varphi_X(i:n) + c_j\varphi_X(j:n), \end{aligned}$$

where

$$\begin{aligned} c_j &= \frac{n!}{(j-1)!(n-j)!} \\ c_{i,j} &= \frac{n!}{(i-1)!(j-i-1)!(n-j)!} \end{aligned}$$

and

$$\begin{aligned} \varphi_X(j:n) &= \int_{-\infty}^{+\infty}\Big(-\ln f_X(x)\Big)\Big(F_X(x)\Big)^{j-1} \\ &\quad \times\Big(1-F_X(x)\Big)^{n-j} f_X(x)dx \\ I(j:n) &= \sum_{m=0}^{n-j}\binom{n-j}{m}\frac{(-1)^m}{(m+j)^2}. \end{aligned}$$

2.2 Entropy of Order Statistics and its Concomitants

Definition 2.2.1 Let $X_1,...,X_n$ be a random sample from a continuous population with cdf $F_X(x)$ and pdf $f_X(x)$ and suppose that $(X_1,Y_1),(X_2,Y_2),...,(X_n,Y_n)$ is a random sample from a joint distribution with cdf $F_{X,Y}(x,y)$ and pdf $f_{X,Y}(x,y)$. We denote the the r-th order statistic corresponding to the $X-$values by $X_{r:n}$ and the corresponding concomitants by $Y_{[r:n]}$. Then the joint distribution of $(X_{r:n},Y_{[r:n]})$ is given by

$$\begin{aligned} f_{X_{r:n},Y_{[r:n]}}(x,y) &= \frac{n!}{(r-1)!(n-r)!}\Big(F_X(x)\Big)^{r-1} \\ &\quad \times\Big(1-F_X(x)\Big)^{n-r} f_{X,Y}(x,y)\ ; \end{aligned}$$

where $r = 1, 2, ..., n$.
Also the joint pdf of the collection of $C(X,Y) = \{(X_{r_1:n}, Y_{[r_1:n]}), (X_{r_2:n}, Y_{[r_2:n]}), ..., (X_{r_k:n}, Y_{[r_k:n]})\}$ is

$$\begin{aligned} f(x_1,..,x_k;y_1,..,y_k) &= \frac{n!}{(r_1-1)!(n-r_k)!} \\ &\times\Big(F_X(x_1)\Big)^{r_1-1} \\ &\times\Big(1-F_X(x_k)\Big)^{n-r_k} \\ &\times\prod_{i=2}^{k}\frac{\Big(F_X(x_i)-F_X(x_{i-1})\Big)^{r_i-r_{i-1}-1}}{(\,r_i-r_{i-1}-1)!} \\ &\times\prod_{i=1}^{k} f_{X,Y}(x_i,y_i) \end{aligned}$$

Theorem 2.2.1 The entropy of $(X_{r:n}, Y_{[r:n]})$ is given by

$$\begin{aligned} H(X_{r:n},Y_{[r:n]}) &= -\ln c_r + c_r(r-1)I(r:n) \\ &\quad + c_r(n-r)I(n-r+1:n) \\ &\quad + \Psi(r:n) \end{aligned}$$

where

$$\begin{aligned} \Psi(r:n) &= E\Big(-\ln f_{X,Y}(X_{r:n},Y_{[r:n]})\Big) \\ &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} c_r\Big(-\ln f_{X,Y}(x,y)\Big) \\ &\quad \times\Big(F_X(x)\Big)^{r-1}\Big(1-F_X(x)\Big)^{n-r} \\ &\quad \times f_{X,Y}(x,y)dxdy \end{aligned}$$

Theorem 2.2.2 The entropy of $C(X,Y)$; $r_1 < r_2 < ... < r_k$ is

$$\begin{aligned} H\Big(C(X,Y)\Big) &= \sum_{i=1}^{k} H(X_{r_i:n},Y_{[r_i:n]}) + \ln c \\ &\quad - \sum_{i=2}^{k} c_{r_i}(r_i-1)I(r_i:n) \\ &\quad - \sum_{i=1}^{k-1} c_{r_i}(n-r_i) \\ &\quad \times I(n-r_i+1:n) \\ &\quad + \sum_{i=2}^{k}(r_i-r_{i-1}-1) \\ &\quad \times [c_{r_i}I\,(r_i:n)+(r_i-r_{i-1}) \\ &\quad \times\binom{r_i-1}{r_{i-1}-1} \\ &\quad \times I\,(r_i-r_{i-1}:r_i-1)] \\ &\quad + \sum_{i=2}^{k}\ln(r_i-r_{i-1}-1)! \\ &\quad + \sum_{i=1}^{k}\ln c_{r_i} \end{aligned}$$

We note that $H\Big(C(X,Y)\Big) - \sum_{i=1}^{k} H(X_{r_i:n},Y_{[r_i:n]})$ does not depend on the distribution.

2.3 Farlie-Gumbel-Morgenstern Family

In this section we compute the Shannon entropy of concomitants of order statistics for the Farlie-Gumbel-Morgenstern (FGM) family of distributions [13]. The general form of the joint pdf of these distributions

$$\begin{aligned} f_{X,Y}(x,y) &= f_X(x)f_Y(y)\Big[1+\alpha\Big(1-2F_X(x)\Big) \\ &\quad \times\Big(1-2F_Y(y)\Big)\Big], \end{aligned}$$

where $-1 < \alpha < 1$ and $f_X(x)$, $f_Y(y)$, $F_X(x)$ and $F_Y(y)$ are respectively marginal the pdfs and cdfs of X and Y.

Theorem 2.3.1 The Shannon entropy contained in the concomitants of the FGM family is

$$
\begin{aligned}
(i)\ H(X_{r:n}, Y_{[r:n]}) &= -\ln c_r + c_r(r-1)I(r:n) \\
&+c_r(n-r)I(n-r+1:n) \\
&+c_r\varphi_X(r:n) + \varphi_Y(1:1) \\
&\times\Big(1+\alpha-\frac{2\alpha r}{n+1}\Big) \\
&+\Big(\frac{4\alpha r}{n+1}-2\alpha\Big)\varphi_Y(2:2) \\
&+c_r G_\alpha(r:n)+\frac{1}{2}
\end{aligned}
$$

$$
\begin{aligned}
(ii)\ H(Y_{[r:n]}) &= \varphi_Y(1:1) \\
&\times\Big(1+\alpha\Big(\frac{n-2r+1}{n+1}\Big)\Big) \\
&-2\alpha\varphi_Y(2:2)\Big(\frac{n-2r+1}{n+1}\Big) \\
&+\frac{\Big((n+1)+\alpha(n-2r+1)\Big)^2}{4(n-2r+1)(n+1)} \\
&\times\Big(-\ln\Big(1+\alpha\Big(\frac{n-2r+1}{n+1}\Big)\Big)\Big) \\
&+\frac{\Big((n+1)-\alpha(n-2r+1)\Big)^2}{4(n-2r+1)(n+1)} \\
&\times\Big(-\ln\Big(1-\alpha\Big(\frac{n-2r+1}{n+1}\Big)\Big)\Big) \\
&+\frac{1}{2},
\end{aligned}
$$

$$
\begin{aligned}
G_\alpha(r:n) &= \int_0^1 \frac{u^{r-1}(1-u)^{n-r}}{4\alpha(1-2u)} \\
&\times\Big(1-\alpha(1-2u)\Big)^2 \\
&\times\ln\Big(1-\alpha(1-2u)\Big)du \\
&-\int_0^1 \frac{u^{r-1}(1-u)^{n-r}}{4\alpha(1-2u)} \\
&\times\Big(1+\alpha(1-2u)\Big)^2 \\
&\times\ln\Big(1+\alpha(1-2u)\Big)du \qquad (1)
\end{aligned}
$$

We now mention some simple properties of $G_\alpha(r:n)$ definition in (1):

(1) $G_\alpha(r:n) = G_\alpha(n-r+1:n)$.

(2) $G_\alpha(r:n) = G_{-\alpha}(r:n)$.

(3) $G_\alpha(r:n)$ does not depend on the distribution $F(.)$.

References

[1] Ahmed, N. A. and Gokhale, D. V. , Entropy Expressions and Their Estimators for Multivariate Distributions, , IEEE Transactions on Information Theory, 35, 688-692, 1989.

[2] Cover, T. M. and Thomas, J. A. , Elements of Information Theory, Wiley, New York, 1991.

[3] Darbellay, G. A. and Vajda, I. , Entropy Expressions for Multivariate Continuous Distributions, IEEE Transactions on Information Theory 46, 709-712, 2000.

[4] Guisado J. L., Jim enez-Morales F., Guerra J. M. , Application of Shannon's Entropy to Classify Emergent Behaviors in a Simulation of Laser Dynamics Behaviors in a Simulation of Laser Dynamics, ICCMSE '03 Proceedings of The International Conference on Computational Methods in Sciences and Engineeringtable of Contents, 213- 216, 2003.

[5] Hartley, R.T.V. , Transmission of Information, Bell System Technical Journal J, 7, 535-563, 1928.

[6] Lazo, A.C.G. and Rathie, P.N. , On the Entropy of Continuous Distribution, IEEE Transactions on Information Theory, 24, 120-122, 1978.

[7] Johnson, N.L. and Kotz, S. , On Some Generalized Farlie-Gumbel-Morgenstern Distributions, Communication Statistics Theory Methods, 4, 415-427, 1972.

[8] Madadi, M. and Tata, M. , Shannon Information in Record Data, Metrika, 74, 11-31, 2011.

[9] Nyquist, H., Certain Affecting Telegraph Speed, Bell System Technical Journal, 3, 324-346, 1924.

[10] Shannon, C. E., A Mathematical Theory of Communication, Bell System Technical Journal, 27, 379-432, 1946.

The 45th Annual Iranian Mathematics Conference
August 26-29 2014
چهل و پنجمین
کنفرانس ریاضی ایران
۴ - ۷ شهریور ۱۳۹۳

Oscillation Space Embedding and the Estimation of Box Dimension: A Correction on Previous Attempts

Ali R. Taheriyoun*, Academic member of Department of Statistics,

Shahid Beheshti University, a_taheriyoun@sbu.ac.ir

Abstract: Based on an embedding on a Besov space, many fractal dimension estimations have been proposed using wavelet transformation. Many of them were diminished by the counter-example on the embedding but corrected by defining a bounded variational seminormed space. Thus, we correct the fractal function estimation approaches which have employed the wrong embedding. The results are given for the problem of estimation of multi-dimensional functions instead of simple univariate case. We also obtain the consistency of the proposed estimator in the probability space induced by random functions. The results are examined through a simulation study on the index-β family of incremental stationary Gaussian fields. Moreover, under the laboratory situations, the procedure of equilibration of heap in a liquid is studied in view of this estimator.

Keywords: box dimension; Hölder continuity; index-β Gaussian field; spatial adaptation; wavelet.

1 INTRODUCTION

Put a surface in an XYZ Cartesian system and consider XY plane as time index while Z represents surface height in each index. We therefore analyze the surface as a noisy path of a random field. This idea would be extended similarly to higher dimensional real value data, which mainly appear in specially astronomical problems. Regardless of physics theories on the noise creation due to uncertainty, noise is an inseparable part of observations even in precise tools. We then often confront with noisy time plots which may cause misunderstanding in exploring the pattern of the surface. Let A be a non-empty bounded subset in $\mathbb{R}^{N+1}$, $N \in \mathbb{N}$, and $\mathcal{C}_\delta(A)$ be the smallest number of sets of diameter lower than δ which cover A. Then the box dimension of A is $\dim_{\mathrm{B}} A = \lim_{\delta \to 0} \frac{\log \mathcal{C}_\delta(A)}{-\log \delta}$. The $\liminf$ and $\limsup$ are computed when the limit does not exist. The results are then called lower and upper box dimensions and denoted by $\overline{\dim}_{\mathrm{B}}$ and $\underline{\dim}_{\mathrm{B}}$, respectively. It follows by the definition that $\dim_{\mathrm{B}} A \leq N+1$ and there is nothing to guarantee that the dimension remains an integer number.

A variety of approaches has been proposed for computing the box dimension of signals or surfaces. The statistical literatures containing theoretical approaches to accomplish good estimators are restricted and we mention [11] and [9] as essentially opening solutions for the statistical box dimension estimation. Statistical results such as asymptotic variance and bias of capacity based box dimension were motivated by [5].

Linking the box dimension to wavelet coefficients

*Corresponding Author

is the basis of our estimator. This linkage was performed by [3] for real value functions on a bounded interval $I \subset \mathbb{R}$ The result is generalized here for real value functions on $\mathbf{I} \subset \mathbb{R}^N$ and then a consistent estimator grows up by applying an inferential solution to the ideal spatial adaptation problem. This procedure was introduced by [12] for simple one-dimensional stochastic processes which is corrected and extended here.

2 Preliminaries

Let $X(\mathbf{t})$, $\mathbf{t} \in \mathbb{R}^N$, be an N-dimensional random field and let $x(\mathbf{t})$ be a realization of $X(\mathbf{t})$. Practically, the regular observations are collected through a bounded subset of $\mathbb{R}^N$, hence without loss of generality, assume that $\mathbf{t} \in \mathbf{I_0} = [0,1]^N$. At resolution $\mathbf{j} = (j_1, \ldots, j_{N+1}) \in \mathbb{N}^{N+1}$ and translation $\mathbf{k} = (k_1, \ldots, k_N) \in \mathbb{N}^N$, let $Q_{\mathbf{j},\mathbf{k}}$ be a dyadic sub-cube of $\mathbf{I_0}$ i.e.

$$Q_{\mathbf{j},\mathbf{k}} = \Big\{ \mathbf{t} \in \mathbf{I_0} \Big| \frac{k_i}{M^{j_i}} \le t_i \le \frac{k_i+1}{M^{j_i}}; i = 1, \ldots, N; \\ k_i = 0, \ldots, M^{j_i} - 1 \Big\},$$

where M is a large enough integer number. Since the resolution will be employed in covering the *graph* of random field, we let $\mathbf{j} \in \mathbb{N}^{N+1}$. We also denote the oscillation of x over the set Q by

$$\begin{aligned} \operatorname{osc}(x;Q) &= \sup_Q \{x(\mathbf{t}) - x(\mathbf{s}) | \mathbf{t}, \mathbf{s} \in Q\} \\ &= \sup_Q x - \inf_Q x. \end{aligned}$$

Let $j_. = \sum_{i=1}^N j_i$ and assume that, $\mathbf{I_0}$ is gridded by $N^{j.}$ numbers of the disjoint sub-cubes with the same volumes. Thus, the total oscillation of x is defined by $\operatorname{Osc}(x;\mathbf{j}) = \sum_{\|Q\|=M^{-j.}} \operatorname{osc}(x;Q)$ where the summation is over all dyadic sub-cubes Q with volume equal to $M^{-j.}$. Define $G_x(A) := \{(\mathbf{t}, x(\mathbf{t})) | \mathbf{t} \in A\}$ which is the graph of x on A. According to [7] we have

$$\mathcal{C}_{M^{-j.}}\big(G_x(\mathbf{I_0})\big) \sim M^{j.} + M^{j_{N+1}} \operatorname{Osc}(x;\mathbf{j}), \tag{1}$$

where $A \sim B$ if $A = O(B)$ and $B = O(A)$. The remaining is the relationship between oscillation and wavelet which traces back to [3] where the oscillation of x is related to the Besov space of particular index. The results were confronted by a counter-example by [7] and have been corrected by [6]. Consider the representation

$$x = \sum_{Q \in \mathcal{Q}} \langle x, \varphi_Q \rangle \psi_Q, \tag{2}$$

for x where $\varphi_{Q_{\mathbf{j},\mathbf{k}}}(\cdot) = 2^{j./2}\varphi(\mathbf{2^j} \cdot - \mathbf{k})$ and $\mathbf{2^j} = (2^{j_1}, \ldots, 2^{j_N})$. The summation represents the wavelet decomposition of x, and φ is called mother wavelet or wavelet, briefly. The sequence $w_Q = \langle x, \varphi_Q \rangle$ is also called wavelet coefficient. The oscillation is related to the wavelet by

$$\begin{aligned} &\liminf_{j\to\infty} \frac{\log \operatorname{Osc}(x,j)}{\log 2^{-j}} \\ &= \limsup_{j\to\infty} \frac{\log \sum_{\mathbf{k}} \sup_{Q_{j',\mathbf{k}'}} |w_{Q_{j',\mathbf{k}'}}|}{j \log 2}. \end{aligned} \tag{3}$$

Employing (1) and (3) or directly from [6], for any continuous real value sample path x on $[0,1]^N$ we have with probability one, $\overline{\dim}_{\mathrm{B}} G_x([0,1]^N) =$

$$\max \Big\{ N, 1 + \limsup_{j\to\infty} \frac{\log \sum_{\mathbf{k}} \sup_{Q_{j',\mathbf{k}'}} |w_{Q_{j',\mathbf{k}'}}|}{j \log 2} \Big\}. \tag{4}$$

3 Fractal dimension estimation

3.1 Consistent estimation

Our sample includes the values of x only on the nodes of a lattice $\mathbb{L}$ and there is no observation within sub-cubes. Let $\mathbf{n} = (n_1, \ldots, n_N)$ be the gridding level which means that the interval $[0,1]$ on the ith axis of the Cartesian system divided into $n_i = 2^{J_i+1}$ equidistant parts where $J_i > 1$ is an integer number. For convenience, let $\mathbf{l}/\mathbf{n}$ denotes the node $(l_1/n_1, \ldots, l_N/n_N)$ for $0 \le l_i \le n_i$, $i = 1, \ldots, N$. The foregoing sampling assumptions confirms that the data $y(\mathbf{l}/\mathbf{n})$ is observed from the nonparametric regression model

$$y(\frac{\mathbf{l}}{\mathbf{n}}) = x(\frac{\mathbf{l}}{\mathbf{n}}) + \varepsilon(\frac{\mathbf{l}}{\mathbf{n}}), \tag{5}$$

for $\mathbf{l}/\mathbf{n} \in \mathbb{L}$ where $\varepsilon(\mathbf{l}/\mathbf{n})$ are independently distributed as $N(0,\sigma^2)$ and x is the real path of X which we would like to find wavelet coefficients for. We need to compute the empirical wavelet coefficients of Y, named w'_Q. Threshold w'_Q by a soft or hard thresholding rules according to the functions $\eta_S(w') = \mathrm{sgn}(w')(|w'| - \lambda_{n^*})$ or $\eta_S(w') = |w'|1(|w'| > \lambda_{n^*})$, respectively. Let us now reconstruct the function by the generated coefficients. This function estimation was introduced by [4] and was called Wavelet shrinkage. With reference to [4] and due to some restrictions for dyadic sub-cubes we have

$$w'_Q = w_Q + u_Q, \tag{6}$$

where u is the empirical wavelet coefficient of ε. For the N-dimensional sequence $\mathbf{j}_n = (j_{1n}, \ldots, j_{Nn})$ in which the ith element is $O(\sqrt{\log n})$ we have the following theorem.

Theorem 3.1. *Under the model (5) if y is a continuous noisy sample path of X on $[0,1]^N$, then*

$$T(y;\mathbf{j}_n) = 1 + \frac{\log_2\left(\sum_{\mathbf{k}} \sup_{Q_{\mathbf{j}'_n,\mathbf{k}'}} |\eta_S(w'_{Q_{\mathbf{j}'_n,\mathbf{k}'}})|\right)}{j_{.n}} \xrightarrow{p} \overline{\dim}_B G_X([0,1]^N). \tag{7}$$

Employing (7), one may estimate the upper box dimension of the graph of a noisy random field, consistently.

Concerning the index-β family: We show that the wavelet coefficients for $Q_{\mathbf{j},\mathbf{k}}$ is rewritten as

$$|w_{Q_{\mathbf{j},\mathbf{k}}}| \le c2^{-j./2}2^{-\beta j_{(1)}} \int \|\mathbf{v}\|^\beta \, |\varphi(\mathbf{v})| \, d\mathbf{v} + c2^{-j./2}2^{-\beta j_{(1)}} \|\mathbf{k}\|^\beta \int |\varphi(\mathbf{v})| \, d\mathbf{v},$$

in this family. The order of noise term in [12] is $n^{-1/2}$ while it is shown that this order is constant here. Thereupon, our estimator is more flexible in constructing precise decisions by using high reolutions. Note that the method of computing the empirical wavelet coefficients is different from the one used by [12]. On the other hand, in low resolutions the variation of w' is controlled by w. It is worth mentioning that in low resolutions the new problem is letting $\mathbf{j}$ to infinity which may cause failure in preparing a large enough $\mathbf{j}$ to hold the convergence true. Thus, a noticeable point is the behavior of convergence. Surprisingly, the order of noise, is not affected by dimension growth of the random field. More precisely, the noise term affects on estimation only via σ^2.

4 Differences with the denoising problem

One may denoise the surface before performing any estimation procedure to ensure that the estimated parameter of roughness measurement is specialized to the surface and is not affected by noise. An efficient method of denoising is based on the wavelet approximation. Projecting the surface onto spaces generated by mother and father wavelets, this method decomposes the surface into two approximation and detail parts, respectively. Since the surface is observed on discrete nodes of the lattice, the projection is strongly affected by the width of the windows which used for estimating the approximation coefficients with respect to mother wavelet. Just like our method, the problem of resolution is arisen here again. In low resolutions, we may loose some parts of the surface and in high resolutions the approximated surface is not well-denoised. According to the wavelet shrinkage technique, denoising is succeeded optimally. After denoising, we will confront the difficulties of using intrinsic estimators. Therefore, if denoising is based on wavelet methods, it can be seen an equivalence between this approach and the one introduced in this paper.

In some problems, one may need to control the roughness instead of measurement. An upper (or lower) bound for the box dimension is required in such cases. Considering the bounded variation property and (1) together is useful to this end. Let

V^α be the class of all functions $x : [0,1]^N \to \mathbb{R}$ satisfying

$$\sup_{j.\in\mathbb{N}} \frac{\mathrm{Osc}(x;\mathbf{j})}{M^{(j.+j_{N+1})(1-\alpha)}} < \infty, \tag{8}$$

and hence the space V^α is non-decreasing in α. Assume that x is continuous on $[0,1]^N$. Using (1), for any $0 < \alpha < 1$

$$\overline{\dim}_{\mathrm{B}} G_x([0,1]^N) = N + 1 - \alpha,$$

if and only if $x \in \bigcap_{\beta<\alpha} V^\beta \setminus \bigcup_{\beta>\alpha} V^\beta$. For the sake of conciseness, the proof is referred to [8] where care to be taken with deducing the last sentence of the proof of sufficiency. This equation can make an upper bound for the box dimension.

References

[1] R. J. Adler. (1981). The Geometry of Random Fields, New York: Wiley.

[2] G. Chan and A. T. A. Wood. (2004). Estimation of fractal dimension for a class of non-Gaussian stationary processes and fields, Ann. Statist. 32, pp 1222-1260.

[3] A. Deliu and B. Jawerth. (1992). Geometrical dimension versus smoothness, Const. Approx. 8, 211-222.

[4] D. L. Donoho, and I. M. Johnstone. (1994). Ideal spatial adaptation by wavelet 344 shrinkage, Biometrika, 81, pp 425-455.

[5] P. Hall, and A. T. A. Wood. (1993). On the performance of box-counting estimators of fractal dimension, Biometrika, 80, pp 246-252.

[6] S. Jaffard. (1998). Oscillation spaces: Properties and applications to fractal and multifractal functions, J. Math. Phys., 39, pp 4129-4141.

[7] A. Kamont, and B. Wolnik. (1999). Wavelet expansions and fractal dimensions. Constr. Approx., 15, pp 97-108.

[8] P. R. Massopust. (1994). Fractal Functions, Fractal Surfaces, and Wavelets. New York.:Academic Press

[9] Y. Ogata, and K. Katsura. (1991). Maximum likelihood estimates of the fractal dimension for random spatial patterns.,Biometrika, 78, pp 463-474.

[10] A. R. Taheriyoun, and K. Shafie. Estimation of fractal dimension of Gaussian fields via Euler characteristic. (to be published).

[11] C. C. Taylor, and S. J. Taylor. (1991). Estimating the dimension of a fractal, J. Roy. Statist. Soc. Ser. B Statist. Methodol., 53, pp 353-364.

[12] Y. Wang. (1997). Fractal function estimation via wavelet shrinkage, J. Roy. Statist. Soc. Ser. B Statist. Methodol., 59, pp 603-613.

The 45th Annual Iranian Mathematics Conference
August 26-29 2014
چهل و پنجمین کنفرانس ریاضی ایران
۴ - ۷ شهریور ۱۳۹۳

A Bayesian approach for the comparison of spectral densities of spatial point processes

A. Saadatjouy , Ph. D. student,
Shahid Beheshti University, saadatjouy.azam@gmail.com
A. R. Taheriyoun*, Faculty of Mathematical Sciences,
Shahid Beheshti University, a_taheriyoun@sbu.ac.ir

Abstract: Spectral analysis prepares an extensive description of both the geometric structure and scales of spatial point patterns. This paper proposes a Bayesian test for comparing the spectral densities of two independent stationary point processes at frequency domain based on the asymptotic distribution of the periodogram by considering a priori distribution for spectral densities.
Keywords: Spatial point process, Spectral analysis, Spectral density function, Periodogram.

1 INTRODUCTION

Simply speaking, a point pattern as a realization of a point process is a set of points in the window where practically the positions and the number of points are random. There are many observations in form of point patterns in nature; for instance, the positions of trees in a forest, and galaxies in the universe. Analysis of spatial point process have been carried out in the spatial and frequency domains via the semivariogram and periodogram, respectively [6]. Just like the time series analysis, the results interpretations in time domain are more objective. This makes a great interest in analysis of spatial point process with the same approach in. The difficulty of Frequency domain analysis is usually originates from application of the Fourier transforms and Hilbert space theory.
On the other hand, the geometric structure of a point pattern has been brought to the interest. This structure is relative to the structure of points aggregation and the effect of points interactions. In the spatial point pattern, the structure of points aggregation and interactions is characterised by intensity and covariance density functions, respectively. Spectral density is the Fourier transform of the complete covariance density function. For stationary point process, spectral density is function of intensity and the Fourier transform of the covariance density functions. Therefore, it is aimed to show that the difference in spectral density functions is resulted from difference in the covariance density functions for two spatial point processes with the same constant intensity functions.
Firstly, [3] suggested spectral techniques for analyzing the spatial point patterns. The technique has been studied and extended by [10] to two-dimensional point patterns. Our method is based on the asymptotic distribution of the ratios of periodograms. For time series problems, [4] initiated a model for spectral ratios and accordingly introduced a nonparametric test for the equality of two

*Corresponding Author

spectra. Reference [7] assumed a regression model for the logarithm of periodogram. Then, they applied generalised likelihood ratio tests to investigate belonging of the spectral density of a stationary time series to a special parametric family. In this paper, a different method is used for testing the equality of the spectral densities of two independent stationary point processes. Accordingly, the theory of spectral analysis of spatial point processes is briefly reviewed in Section 2. In viewpoint of frequentists, a Bayesian testing procedure is presented in Section 3.

2 PRELIMINARIES

It is often possible to discriminate several spatial point processes by comparing their first- and second-order properties. The first order property, namely intensity function, is defined as the expected number of points per unit volume [5] as follows:

$$\lambda_X(\mathbf{a}) = \lim_{|\mathrm{d}\mathbf{a}|\to 0} \frac{E[N_X(\mathrm{d}\mathbf{a})]}{|\mathrm{d}\mathbf{a}|}, \quad \mathbf{a} \in \mathbf{R}^d, \qquad (1)$$

where $d \in \mathbf{N}$, $\mathrm{d}\mathbf{a}$ is the small region around the point $\mathbf{a}$, $|\mathrm{d}\mathbf{a}|$ is the volume of this region, and $N_X(\mathrm{d}\mathbf{a})$ is the cardinality measure of this region.
Furthermore, the second-order properties, often described by means of second-order intensity, characterize the covariance between the number of points at different regions [5] as follows:

$$\lambda_{XX}(\mathbf{a},\mathbf{b}) = \lim_{\substack{|\mathrm{d}\mathbf{a}|\to 0\\ |\mathrm{d}\mathbf{b}|\to 0}} \frac{E[N_X(\mathrm{d}\mathbf{a})N_X(\mathrm{d}\mathbf{b})]}{|\mathrm{d}\mathbf{a}||\mathrm{d}\mathbf{b}|},$$
$$\mathbf{a} \neq \mathbf{b}, \quad \mathbf{a},\mathbf{b} \in \mathbf{R}^d. \qquad (2)$$

Reference [3] proposed a complete covariance density function for description of second-order properties of orderly spatial point process as follows:

$$\kappa_{XX}(\mathbf{a},\mathbf{b}) = \lambda_X(\mathbf{a})\delta(a_1-b_1)\ldots\delta(a_d-b_d)+\gamma_{XX}(\mathbf{a},\mathbf{b}),$$

where $\delta(a)$ is the Dirac delta-function and γ_{XX} is the covariance density function, which in turn can be defined as:

$$\gamma_{XX}(\mathbf{a},\mathbf{b}) = \lambda_{XX}(\mathbf{a},\mathbf{b}) - \lambda_X(\mathbf{a})\lambda_X(\mathbf{b}). \qquad (3)$$

For a stationary point process the intensity function is constant, i.e. $\lambda_X(\mathbf{a}) = \lambda$ and $\lambda_{XX}(\mathbf{a},\mathbf{b}) = \lambda_{XX}(\mathbf{a}-\mathbf{b}) = \lambda_{XX}(\mathbf{h})$. The second equation means that the second-order intensity depends merely on the vector difference, $\mathbf{h}$, among $\mathbf{a}$ and $\mathbf{b}$. Consequently, the complete covariance density function of a stationary point process is reduced to

$$\kappa_{XX}(\mathbf{h}) = \lambda_X\delta(h_1)\ldots\delta(h_d) + \gamma_{XX}(\mathbf{h}). \qquad (4)$$

Reference [3] extended spectral analysis for the case $d = 1$ to spatial case $d = 2$, which has been used by [2].
The characterization of the spectral density function of a point process is carried out similarly to the Fourier transform of the complete covariance density function. It is given by following equation for stationary point process:

$$\mathrm{f}_X(\boldsymbol{\omega}) = \int_{\mathbf{R}^d} \kappa_{XX}(\mathbf{h})e^{-i\langle\boldsymbol{\omega},\mathbf{h}\rangle}\mathrm{d}\mathbf{h} = \int_{\mathbf{R}^d}[\lambda_X\delta(h_1)\ldots\delta(h_d)$$
$$+\gamma_{XX}(\mathbf{h})]e^{-i\langle\boldsymbol{\omega},\mathbf{h}\rangle}\mathrm{d}\mathbf{h} = \lambda_X + \int_{\mathbf{R}^d}\gamma_{XX}(\mathbf{h})e^{-i\langle\boldsymbol{\omega},\mathbf{h}\rangle}\mathrm{d}\mathbf{h}.$$

Assume that an observed pattern contains N_X events in a rectangular region W, in which W has sides of length L_i, where along the i'th coordinate of the Cartesian system for $i = 1,\ldots,d$. Let $\mathbf{z}_j, j = 1,...,N_X$, indicate the positions of the events. The spectral density function for such pattern can be estimated by the discrete Fourier transform, called periodogram, specified at the Fourier frequencies $\boldsymbol{\omega}_{\mathbf{p}} = 2\pi\mathbf{p}$, with $\mathbf{p} = (p_1,\ldots,p_d), p_j = 0,\pm1,\pm2,\ldots$ for $j = 1,\ldots,d$ [6]. Suppose that $\{z_1,\ldots,z_{N_X}\}$ is the observed point pattern in the Euclidean space. The periodogram according to this sample path is defined as follows:

$$I_X(\boldsymbol{\omega}) = F_X(\boldsymbol{\omega})\overline{F_X}(\boldsymbol{\omega})$$
$$= \left(\sum_{j=1}^{N_X} exp\{i\boldsymbol{\omega}^T L^{-1}\mathbf{z}_j\}\right)\overline{\left(\sum_{k=1}^{N_X} exp\{i\boldsymbol{\omega}^T L^{-1}\mathbf{z}_k\}\right)}$$
$$= \sum_{j=1}^{N_X}\sum_{k=1}^{N_X} exp\{i\boldsymbol{\omega}^T L^{-1}(\mathbf{z}_j - \mathbf{z}_k)\} \qquad (5)$$

where $\overline{F_X}(\boldsymbol{\omega})$ is the complex conjugate of $F_X(\boldsymbol{\omega})$, and L is a scaling matrix given by $L = diag(L_1, \ldots, L_d)$.
It is easily verified that by symmetry, $I_X(\boldsymbol{\omega_p}) = I_X(\boldsymbol{\omega_{-p}})$. The periodogram is an unbiased but inconsistent estimate of the spectral density $\mathrm{f}_X(\boldsymbol{\omega})$ [8]. Reference [8] recommended computation of $I_X(\boldsymbol{\omega_p})$ for $\mathbf{p} \in \{0, \pm 1, \ldots, \pm 16\}^d$ when $N_X < 100$. However, in this study, we consider $\mathbf{p} \in \{\pm 1, \ldots, \pm 8\}^d$ for comparing the spectral densities of two point processes.
Reference [9] confirmed that for random fields in two dimensions, when $N_X \to \infty$, spectral estimates have asymptotically χ^2 distribution. Reference [12] extended this problem to the higher dimensions. Similar result satisfied for d-dimensional point processes as follow:

$$\frac{2I_X(\boldsymbol{\omega_p})}{\mathrm{f}_X(\boldsymbol{\omega_p})} \sim \chi^2_{(2)}, \qquad \boldsymbol{\omega_p} \neq \mathbf{0}, \tag{6}$$

and

$$\frac{2\{I_X(\mathbf{0}) - \lambda_{\mathbf{X}}\}}{\mathrm{f}_X(\mathbf{0})} \sim \chi^2_{(1)}. \tag{7}$$

There is restriction for the independency of periodograms at different frequencies and it highly depends on the geometry of the window W. Concerning the random fields, [1] showed that the independency of periodograms can be obtained when W is a hyper-cube in $\mathbf{R}^d$. In our problem we then consider that the assumptions of [1] hold true.

3 TESTING APPROACH

In this section, a test based on periodogram ordinates is proposed for comparing the spectral densities of two independent stationary point processes in $\mathbf{R}^2$. Symmetric property of periodogram allows us to only calculate periodogram values for $p_1 = \pm 1, ..., \pm 8$ and $p_2 = 1, ..., 8$.
Suppose that X and Y are independent stationary point processes and let I_X and f_X and I_Y and f_Y be respectively their corresponding periodogram and spectral density functions.
In this condition, the test can be discussed with respect to the following hypotheses:

$$H_0 : \mathrm{f}_X(\boldsymbol{\omega}) = \mathrm{f}_Y(\boldsymbol{\omega}), \quad \text{versus}$$
$$H_1 : \mathrm{f}_X(\boldsymbol{\omega}) \neq \mathrm{f}_Y(\boldsymbol{\omega}), \quad \boldsymbol{\omega} \in \mathbf{R}^2. \tag{8}$$

We attempt to find a rejection area based on the estimates at only Fourier frequencies. Set $\eta(\mathbf{p}) = \mathrm{f}_X(\boldsymbol{\omega_p})/\mathrm{f}_Y(\boldsymbol{\omega_p})$ and η_j be the restriction of η to $H_j, j = 0, 1$. From (6), and independence of I_X and I_Y, it is clear that

$$\frac{I_X(\boldsymbol{\omega_p})/\mathrm{f}_X(\boldsymbol{\omega_p})}{I_Y(\boldsymbol{\omega_p})/\mathrm{f}_Y(\boldsymbol{\omega_p})} \sim F(2, 2), \tag{9}$$

and hence $T(\mathbf{p}) = I_X(\boldsymbol{\omega_p})/I_Y(\boldsymbol{\omega_p})$ is distributed as

$$f_T(t|\eta) = \frac{\eta}{(\eta + t)^2}, \quad t > 0. \tag{10}$$

Under the null hypothesis, $\eta_0 = 1$ and $T(\mathbf{p}) = I_X(\boldsymbol{\omega_p})/I_Y(\boldsymbol{\omega_p}) \sim F(2, 2)$ with the following density

$$f_T(t|\eta_0) = \frac{1}{(1 + t)^2}, \quad t > 0. \tag{11}$$

3.1 Bayesian approach

Here, the [11] definition was used for calculation of the Bayes factor and comparison of the spectral densities of two stationary point processes.
Reference [11] described the Bayes factor as equation (12) by taking into account the point null hypothesis H_0 and marking by g_1 and m_1, respectively, the prior density and marginal distribution under H_1,

$$B^{\pi}_{01} = \frac{f(t|\eta_0)}{m_1(t)} = \frac{f(t|\eta_0)}{\int_0^{\infty} f(t|\eta) g_1(\eta) d\eta}. \tag{12}$$

For the sake of conciseness, we refer to [11][Section 5.2] for the rejection criterion using the Bayes factor. For convenience, $\mathrm{f}_X(\boldsymbol{\omega_p})$ and $\mathrm{f}_Y(\boldsymbol{\omega_p})$ are abbreviated as f_X and f_Y. Suppose f_X and f_Y have the Log-normal prior distributions with parameters (μ_X, σ^2_X) and (μ_Y, σ^2_Y), respectively, i.e., the density function of f_X is

$$f_{\mathrm{f}_X}(\mathrm{f}) = \frac{1}{\mathrm{f}\sqrt{2\pi\sigma^2_X}} exp\{-\frac{1}{2\sigma^2_X}(\ln \mathrm{f} - \mu_X)^2\},$$
$$\mathrm{f} > 0, \quad \mu_X \in \mathbf{R}, \quad \sigma_X > 0. \tag{13}$$

The prior distribution of $\eta(\mathbf{p})$ under the alternative hypothesis is Log-normal with parameters $(\mu_X - \mu_Y, \sigma_X^2 + \sigma_Y^2)$, and the marginal distribution of T on H_1 is

$$m_1(t) = \int_0^\infty \frac{\eta}{(\eta+t)^2} \frac{1}{\eta\sqrt{2\pi(\sigma_X^2+\sigma_Y^2)}} \times$$
$$exp\{-\frac{1}{2(\sigma_X^2+\sigma_Y^2)}(\ln\eta - (\mu_X-\mu_Y))^2\}d\eta. \quad (14)$$

For this problem, the Bayes factor equals to

$$B_{10}^\pi = \frac{1}{B_{01}^\pi} = (1+t)^2 \int_0^\infty \frac{1}{(\eta+t)^2} \frac{1}{\sqrt{2\pi(\sigma_X^2+\sigma_Y^2)}} \times$$
$$exp\{-\frac{1}{2(\sigma_X^2+\sigma_Y^2)}(\ln\eta - (\mu_X-\mu_Y))^2\}d\eta. \quad (15)$$

A numerical study based on 1000 simulation replicates from Strauss process with parameters $\beta = 5$ (intensity parameter), $\gamma = 0.7$ (interaction parameter) and $R = 0.05$ (interaction radius) on a 10×10 square and assuming values $\mu_X = \mu_Y = 5$ and $\sigma_X^2 = \sigma_Y^2 = 1$ for hyper parameters, shows that this test reject the null hypothesis at level less than 1%. When two point processes with the same constant intensities were compared, rejection of the null hypothesis indicated that their spectral densities were different. Since for a stationary point process the spectral density is a function of constant intensity and covariance density function, so, one can conclude that the difference of two spectral densities are due to the difference of their covariance density functions. But this is not true for Poisson processes, due to the independency of points position in stationary Poisson processes, in which $\lambda_{XX}(\mathbf{a}, \mathbf{b}) = \lambda_{XX}(\mathbf{h}) = \lambda^2$ and $\gamma_{XX}(\mathbf{a}, \mathbf{b}) = \gamma_{XX}(\mathbf{h}) = 0$, so the spectral density function equals to the constant intensity.

References

[1] Bandyopadhyay, S., & Lahiri, S. N. (2009). Asymptotic properties of discrete Fourier transforms for spatial data. Sankhyā: The Indian Journal of Statistics, Series A (2008-), 221-259.

[2] Bartlett, M. S. (1963). The spectral analysis of point processes. Journal of the Royal Statistical Society. Series B (Methodological), 264-296.

[3] Bartlett, M. S. (1964). The spectral analysis of two-dimensional point processes. Biometrika, 51(3/4), 299-311.

[4] Coates, D. S., & Diggle, P. J. (1986). Tests for comparing two estimated spectral densities. Journal of Time Series Analysis, 7(1), 7-20.

[5] Diggle, P. J. (1983). Statistical analysis of spatial point patterns.

[6] Dorai-Raj, S. S. (2001). First-and Second-Order Properties of Spatio-temporal Point Processes in the Space-Time and Frequency Domains (Doctoral dissertation, Virginia Polytechnic Institute and State University).

[7] Fan, J., & Zhang, W. (2004). Generalised likelihood ratio tests for spectral density. Biometrika, 91(1), 195-209.

[8] Mugglestone, M. A., & Renshaw, E. (1996). A practical guide to the spectral analysis of spatial point processes. Computational statistics & data analysis, 21(1), 43-65.

[9] Pagano, M. (1971). Some asymptotic properties of a two-dimensional periodogram. Journal of Applied Probability, 8(4), 841-847.

[10] Renshaw, E., & Ford, E. D. (1984). The description of spatial pattern using two-dimensional spectral analysis. Vegetatio, 56(2), 75-85.

[11] Robert, C. P. (2001). The Bayesian Choice: From Decision-Theoretic Foundations to Computational Implementation (Springer Texts in Statistics) by.

[12] Taheriyoun, A. R. (2012). Testing the covariance function of stationary Gaussian random fields. Statistics & Probability Letters, 82(3), 606-613.

The 45th Annual Iranian Mathematics Conference
August 26-29 2014
چهل و پنجمین
کنفرانس ریاضی ایران
۴ - ۷ شهریور ۱۳۹۳

IBM Word-Alignment Model 1 for Statistical Machine Translation

N. Daneshgar*, Ph.D. student

M. Sarmad, Assistant Proffessor, Faculty of Mathematical Sciences

Ferdowsi University of Mashhad, sarmad@um.ac.ir

Ferdowsi University of Mashhad, ne da978@stu-mail.um.ac.ir

Abstract: This paper discribes the word alignment based on IBM model 1 for an English-Persian Parallel Corpus (Mizan). We only consider 100,000 sentence pairs. It is been obtained that for each sentence, more than 80 percentage of word alignments between English and Persian words are reasonable.

Keywords: word alignment, statistical models, IBM model 1, EM algorithm

1 INTRODUCTION

Machine translation (MT) is the automatic translation of a text in one language into another. Statistical Machine Translation (SMT) is one of the methods of MT that is based on statistical models. [2]

There are two general models to build SMT systems: word-based and phrase-based models, but the latter is currently used [3].

Word alignment is used by phrase-based systems to extract phrase pairs from trining data and build tables of possible translations of phrase [2]. There are several methods for word alignment that are divided into two groups: generative word alignment and discriminative word alignment models. The former is usually based on IBM models. For the first time, Brown et. al [1] introduced IBM models.

In this paper, it is considered IBM model 1 and its parameter estimation via Expectation-Maximization (EM) algorithm. For a number of bilingual parallel corpus it has been done by Giza++ [5]. But we do it by R software for 100,000 sentence pairs for an English-Persian named Mizan [4] and over 80 percentage of word alignments having true link (translation) between English and Persian words.

2 Definitions

Sentence-aligned Parallel Corpus: Large set of bilingual texts such that for each sentence in one natural language there is its translation into another [2].*

Alignment : It is a link between English and Persian words.

Alignment Function: Alignment can be formalized with an alignment function. Mapping an English target word at position j to a Persian source

*Corresponding Author

*In this paper it is considered an English-Persian parallel corpus.

word at position i with a function $a : j \to i$ [3].

In this paper it is assumed that we want to translate Persian sentence into English.

3 IBM Model 1

The fundamental aim of SMT is finding most probable English sentence given a Persian sentence as follows:

$$\widehat{e} = argmax_e P(\mathbf{e}|\mathbf{f}) \quad (1)$$

Probability are determined by training a statistical model using the parallel corpus [2].

Assuming that we have a sentence-aligned parallel corpus (F,E) where F (E) is a set of $\mathbf{f}$ ($\mathbf{e}$) sentences, also is called source (target) language and $f_i(e_j)$ denotes the $i-th$ $(j-th)$ word in $\mathbf{f}$ ($\mathbf{e}$) sentence with length of $l_f(l_e)$.

In IBM model 1, we have

$$P(\mathbf{e}|\mathbf{f}) = \sum_{\mathbf{a}} \mathbf{P}(\mathbf{a}, \mathbf{e}|\mathbf{f}) \quad (2)$$

where $P(a, \mathbf{e}|\mathbf{f})$ is the translation probability for a Persian sentence $\mathbf{f}$ to an English sentence $\mathbf{e}$ with an alignment of each English word e_j to a Persian word f_i according to the alignment function $a : j \to i$ as follows

$$P(a, \mathbf{e}|\mathbf{f}) = \frac{\varepsilon}{(\mathbf{l_f}+\mathbf{1})^{\mathbf{l_e}}} \prod_{\mathbf{j=1}}^{\mathbf{l_e}} \mathbf{t}(\mathbf{e_j}|\mathbf{f_{a(j)}}) \quad (3)$$

where ε is a normalization constant and $t(e_j|f_{a(j)})$ is the (conditional) probability of the words being aligned [3].

Relative frequency estimates can be used to estimate $t(e_j|f_{a(j)})$. The problem is that we do not have word-aligned data. There is a mathematical solution to this problem which is EM algorithm.

3.1 IBM Model 1 and EM algorithm

As we mentioned above, $t(e_j|f_{a(j)})$ is to be estimated via EM algorithm. To do this, we obtain

$$t(e_j|f_i; \mathbf{e}, \mathbf{f}) = \frac{\sum_{\mathbf{e},\mathbf{f}} \mathbf{c}(\mathbf{e_j}|\mathbf{f_i}; \mathbf{e}, \mathbf{f})}{\sum_{\mathbf{e}} \sum_{\mathbf{e},\mathbf{f}} \mathbf{c}(\mathbf{e_j}|\mathbf{f_i}; \mathbf{e}, \mathbf{f})} \quad (4)$$

where $c(e_j|f_i; \mathbf{e}, \mathbf{f})$ is a count function. It collects evidence from a sentence pair ($\mathbf{e}$,$\mathbf{f}$) that a particular source word f_i translates into the target word e_j i.e

$$c(e_j|f_i; \mathbf{e}, \mathbf{f}) = \sum_{\mathbf{a}} \mathbf{P}(\mathbf{a}|\mathbf{e}, \mathbf{f}) \sum_{\mathbf{j=1}}^{\mathbf{l_e}} \delta(\mathbf{e}, \mathbf{e_j})\delta(\mathbf{f}, \mathbf{f_{a(j)}}) \quad (5)$$

where $\delta(x, y)$ is 1 if $x = y$ and 0 otherwise and using (3) $P(a|\mathbf{e}, \mathbf{f})$ can be calculated as bellow

$$P(a|\mathbf{e}, \mathbf{f}) = \frac{P(a, \mathbf{e}|\mathbf{f})}{\sum_a P(a, \mathbf{e}|\mathbf{f})} = \prod_{j=1}^{l_e} \frac{t(e_j|f_{a(j)})}{\sum_{i=0}^{l_e} t(e_j|f_{a(j)})} \quad (6)$$

Applying (5) in (6), it is obtained

$$c(e_j|f_i; \mathbf{e}, \mathbf{f}) = \frac{\mathbf{t}(\mathbf{e_j}|\mathbf{f_i})}{\sum_{\mathbf{i=0}}^{\mathbf{l_f}} \mathbf{t}(\mathbf{e_j}|\mathbf{f_i})} \sum_{\mathbf{j=1}}^{\mathbf{l_e}} \delta(\mathbf{e}, \mathbf{e_j}) \sum_{\mathbf{i=0}}^{\mathbf{l_f}} \delta(\mathbf{f}, \mathbf{f_i}) \quad (7)$$

Thus the EM algorithm works as the following:

1. Initialize value for $t(e_j|f_i)$, (e.g. uniform).
2. Calculate $c(e_j|f_i; \mathbf{e}, \mathbf{f})$ by (7) (Expectation Step).
3. Recalculate $t(e_j|f_i)$ by (4) (Maximization Step).
4. Iterate steps 2 and 3 until convergence [3].

R codes are available in appendix.

4 Word Alignment Based on IBM Models

To use the IBM models for word alignment, it is iterated EM algorithm to find estimation of $t(e_j|f_i)$ and then obtain $P(a, \mathbf{e}|\mathbf{f})$ using (3). It is also called Viterbi alignment.

5 Results

EM algorithm is applied to estimate parameters of IBM model 1 for the English-Persian Parallel Corpus Mizan (only for 100,000 sentence pairs) [4]. Over 80 percentage of parameter estimations are reasonable.

6 Appendix

```
a1=a2=a=b=t=u=bb=c()
a1=as.character(read.delim('D:/Mizan_En1.txt'
      ,sep='\n',nrows=100000,h=F)[[1]])
a2=as.character(read.delim('D:/Mizan_Fa1.txt'
     ,sep='\n',encod='UTF-8',nrows=100000
     ,h=F)[[1]])

a=cbind(a1,a2)
a3=nrow(a)

b=apply(a,1,function(x)cbind(Var1=rep.int
    (strsplit(as.character(x[1]),' ')[[1]],
    length(strsplit(as.character(x[2]),' ')
    [[1]])), Var2=rep(strsplit(as.character
    (x[2]),' ')[[1]], each=length(strsplit
    (as.character(x[1]),' ')[[1]]))))

bb=unlist(b)

cc=sapply(b,length)

kk1=c(1,cumsum(cc[-length(cc)])+1)
kk2=kk1+cc/2-1
pp=eval(parse(text=paste('c(',paste(kk1,kk2
     ,sep=':',collapse=','),')',\\sep='')))

e=bb[-pp]
f=bb[pp]

k=cc/2

g=rep(1:a3,k)
t=as.numeric(rep(1/k,k))

e=tolower(e)
f=tolower(f)

library('AnnotationDbi')

o1=order(paste(g,e,f))

z3=paste(f,e)
jjj=order(z3)
t=t[jjj]
for(iiii in 1:10)
{

count0=by(t,paste(g[jjj],e[jjj]),
            function(x)x/sum(x))

count=by(unlist(count0),paste(f[o1],e[o1]),
          function(x)rep(sum(x),length(x)))

total= by(unlist(count0),f[o1],function(x)
          rep(sum(x),length(x)))

t=unlist2(count)/unlist2(total)
}
```

7 Reference

References

[1] P. Brown, S. Della Pietra, V. Della Pietra, R. Mercer, "The Mathematics of Statistical Machine Translation: Parameter Estimation", Computational Linguistics, 1993.

[2] J. Brunning, "Alignment Models and Algorithms for Statistical machine translation", Cambridge University Engineering Department, Aug. 2010 , pp. 1-10.

[3] P. Koehn, "Statistical machine translation", Cambridge University , New York, 2010, pp. 80-90.

[4] www.dadegan.ir

[5] www.statmt.org

The 45th Annual Iranian Mathematics Conference
August 26-29 2014
چهل و پنجمین کنفرانس ریاضی ایران
۴ - ۷ شهریور ۱۳۹۳

ترتیب های تصادفی TTT و EW و برخی ویژگیهای آنها

مجتبی اصفهانی *، عضو هیأت علمي گروه آمار، دانشگاه ولايت، esfahani۶۴@yahoo.com

محمد امینی، عضو هیأت علمي گروه آمار، دانشگاه فردوسی، m-amini@um.ac.ir

غلامرضا محتشمی برزادران ،عضو هیأت علمي گروه آمار، دانشگاه فردوسی gmb۱۳۳۰@yahoo.com،

چکیده: در این مقاله ترتیب های تصادفی کل زمان آزمون (TTT) و فزونی ثروت (EW) معرفی و برخی از ویژگی های آنها و همچنین ارتباط این دو ترتیب تصادفی با یکدیگر و سایر ترتیب های تصادفی مورد مطالعه قرار می گیرد. در ادامه کاربردهایی از این ترتیب های تصادفی در بیمه و نظریه قابلیت اعتماد ارایه می شود.

کلمات کلیدی: ترتیب تصادفی کل زمان آزمون، ترتیب تصادفی فزونی ثروت، تبدیل TTT مرتبط با F، تبدیل EW مرتبط با F، تبدیل توقف زیان

مقدمه

ترتیب تصادفی TTT بطور جدی توسط بارلو، بارتالومو، برمنر و برانک (۱۹۷۲) مورد بررسی و مطالعه قرار گرفت و پس از آن بارلو و داکسوم (۱۹۷۲) برخی از خواص ترتیب تصادفی TTT و رابطه آن با ترتیب محدب و همچنین کاربردهایی از آن را بیان کردند. تحقیقات زیادی روی این ترتیب تصادفی انجام شده است، به عنوان مثال بارتازویچ (۱۹۸۶)، (۱۹۹۸) و (۱۹۹۵) کاربردهایی از تبدیل TTT را بیان کرد. در سال های اخیر رابطه بین ترتیب های تصادفی TTT و EW مورد توجه محققانی مانند لی و شیکد و کوچار (۲۰۰۲) و لی و شیکد (۲۰۰۴) و (۲۰۰۷) قرار گرفته است.

ترتیب تصادفی

در این بخش پس از معرفی ترتیب های تصادفی TTT و EW سایر ترتیب های تصادفی مورد مطالعه در مقاله معرفی می گردد(شیکد و شانتیکو ۲۰۰۷).

تعریف ۱. فرض کنید X و Y دو متغیر تصادفی نامنفی به ترتیب با توزیع های F و G باشند:

الف) اگر ظ $(0,1) \in p$، $\int_0^{F^{-1}(p)} \bar{F}(x)dx \leq \int_0^{G^{-1}(p)} \bar{G}(x)dx$ آنگاه X را کوچکتر از Y در ترتیب تصادفی TTT گوییم و با نمادهای $X \leq_{ttt} Y$ یا $F \leq_{ttt} G$ نشان می دهیم.

ب) اگر $p \in (0,1)$، $\int_{F^{-1}(p)}^{\infty} \bar{F}(x)dx \leq \int_{G^{-1}(p)}^{\infty} \bar{G}(x)dx$ آنگاه X را کوچکتر از Y در ترتیب EW گوییم و با نمادهای $X \leq_{ew} Y$ یا $F \leq_{ew} G$ نشان می دهیم.

تعریف ۲. فرض کنید X و Y دو متغیر تصادفی با توابع توزیع F و G باشند و F^{-1} و G^{-1} به ترتیب توابع معکوس F و G باشند، اگر F^{-1} و G^{-1} از راست پیوسته باشند و $G^{-1}(p) - F^{-1}(p)$ تابعی صعودی، آنگاه X را کوچکتر از Y در ترتیب پراکندگی نامند و با نماد $X \leq_{disp} Y$ نشان می دهیم.

تعریف ۳. فرض کنید X و Y دو متغیر تصادفی نامنفی با توابع توزیع مطلقا پیوسته F و G باشند که دارای تکیه گاه های $[0,a)$ و $[0,b)$ هستند و a و b ثابتهای متناهی یا نامتناهی اند:

الف)اگر $G^{-1}(F)$ تابعی محدب باشد آنگاه X را کوچکتر از Y در ترتیب تبدیل محدب نامند و با نمادهای $X \leq_c Y$ یا $G \leq_c F$ نشان می دهیم.

ب) اگر $\frac{G^{-1}(F)}{x}$ نسبت به $x > 0$ صعودی باشد آنگاه X را کوچکتر از Y در ترتیب ستاره نامیم و با نماد $X \leq_* Y$ نشان می دهیم.

تعریف ۴. فرض کنید X و Y دو متغیر تصادفی باشند:

الف) اگر برای هر تابع محدب ϕ، $E[\phi(X)] \leq E[\phi(Y)]$ آنگاه X را کوچکتر از Y در ترتیب محدب نامیده می شود و با نمادهای $Y \leq_{cx} X$ یا $G \leq_{cx} F$ نشان می دهیم.

ب) اگر برای هر تابع محدب صعودی ϕ، $E[\phi(X)] \leq E[\phi(Y)]$ آنگاه X را کوچکتر از Y در ترتیب محدب صعودی نامیده می شود و با نمادهای $Y \leq_{icx} X$ یا $G \leq_{icx} F$ نشان می دهیم.

ج) اگر برای هر تابع مقعر صعودی ϕ، $E[\phi(X)] \leq E[\phi(Y)]$ آنگاه X را کوچکتر از Y در ترتیب مقعر صعودی نامیده می شود و با نمادهای $Y \leq_{icv} X$ یا $G \leq_{icv} F$ نشان می دهیم.

ویژگی های ترتیب های تصادفی TTT و EW

ویژگی های ترتیب های تصادفی TTT و EW در این بخش مطالعه می شود.

اگر X متغیری تصادفی با تابع توزیع F باشد آنگاه $p \in (0,1)$، $T_X(p) = \int_0^{F^{-1}(p)} \bar{F}(x)dx$ را تبدیل TTT مرتبط با F می نامیم و $p \in (0,1)$، $W_X(p) = \int_{F^{-1}(p)}^{\infty} \bar{F}(x)dx$ را تبدیل EW مرتبط با F نامیم(کوچار و همکاران۲۰۰۲).

وقتی X یک متغیر تصادفی نامنفی با $\mu < \infty$ باشد مشاهده کل زمان آزمون وقتی X رخ داده را با X_{ttt} نشان می دهیم و به صورت $X_{ttt} = T_X(F(X))$ تعریف می شود و مشاهده فزونی ثروت وقتی X رخ داده است را با X_{ew} نشان می دهیم و به صورت $X_{ew} = W_X(F(X))$ تعریف می شود.

تبصره ۱. چون $F(X)$ دارای توزیع یکنواخت روی بازه $(0,1)$ است و $X_{ttt} = \int_0^x \bar{F}(x)dx$ و همچنین $X_{ew} =_{st} W_X(U)$ و $X_{ew} = \int_x^{\infty} \bar{F}(x)dx$ که در آن $=_{st}$ به معنای برابری در توزیع است.

مثال ۱. اگر X متغیرتصادفی با تابع بقای $x > 0$، $\bar{F}(x) = \frac{1}{1+x}$ باشد آنگاه:

$$X_{ttt} = \int_0^x \frac{1}{1+x}dx = log(1+X),$$

و تابع توزیع X_{ttt} برابر است با $y > 0$، $F_{X_{ttt}}(X) = 1 - e^{-y}$ ولی $X_{ew} = \int_x^{\infty} \frac{1}{1+x}dx$.

قضیه ۱. فرض کنید X متغیر تصادفی نامنفی با تابع توزیع پیوسته F و $\mu = E(X) < \infty$ باشد. در این صورت

الف) تابع توزیع X_{ttt} معکوس تبدیل TTT است.

ب) تابع بقای X_{ew} معکوس تبدیل فزونی ثروت است.

نتیجه ۱. اگر شرایط قضیه ۱ برقرار باشد آنگاه $X_{ew} = \mu - X_{ttt}$.

برای هر $k \geq 1$ گشتاورهای X_{ttt} و X_{ew} به صورت زیر محاسبه می شود:

$$\begin{aligned} E[X_{ttt}]^k &= \int_0^{\infty} [\int_0^x \bar{F}(y)dy]^k dF(x) \\ &= \int_0^{\infty} \int_0^x \ldots \int_0^x [\prod_{i=1}^k \bar{F}(y_i)]dy_1 \ldots dy_k dF(x) \\ &= k! \int .. \int_{0 \leq y_1 \leq .. \leq y_k} \bar{F}(y_1)..\bar{F}(y_{k-1})[\bar{F}(y_k)]^2 dy_1..dy_k \end{aligned}$$

در حالت خاص برای $k = 1$ داریم:

$$E[X_{ttt}] = \int_0^{\infty} \bar{F}^2(x)dx,$$

به طور مشابه برای X_{ew} داریم:

$$E[X_{ew}]^k = k! \int .. \int_{0 \leq y_1 \leq .. \leq y_k} F(y_1)\bar{F}(y_1)..\bar{F}(y_k)dy_1..dy_k$$

و در حالت خاص برای $k = 1$ داریم:

$$E[X_{ew}] = \int_0^{\infty} F(x)\bar{F}(x)dx.$$

ارتباط ترتیب های تصادفی TTT و EW با یکدیگر و سایر ترتیب های تصادفی

در این قسمت ارتباط ترتیب های تصادفی TTT و EW با برخی ترتیب های تصادفی مورد بررسی قرارمی دهیم.

قضیه ۲.اگر X و Y دو متغیر تصادفی نامنفی باشند آنگاه:

الف) $X \leq_{ttt} Y \Longleftrightarrow X_{ttt} \leq_{st} Y_{ttt}$

ب) $X_{ttt} \leq_{ttt} Y_{ttt} \Longrightarrow X \leq_{ttt} Y$ و

$X \leq_{st} Y \Longrightarrow X_{ttt} \leq_{st} Y_{ttt}$

ج) اگر $\mu_y < \infty, \mu_x < \infty$ آنگاه

$X_{ew} \leq_{st} Y_{ew} \Longleftrightarrow X \leq_{ew} Y$

د) برای هر $a > 0$، $(aX)_{ttt} =_{st} aX_{ttt}$ و علاوه بر آن اگر

$\mu < \infty$ آنگاه $(aX)_{ew} =_{st} aX_{ew}$

برخی از ویژگی های ترتیب های تصادفی EW برای متغیرهای تصادفی نامنفی X و Y با میانگین متناهی و تبدیل تبدیل های فزونی ثروت W_X و W_Y توسط شیکد و شانتیکومار(۱۹۹۸) بررسی و اثبات شده است. آنها نشان دادند اگر $W_X \leq W_Y$ آنگاه $X \leq_{ew} Y$ و $E(X) \leq E(Y)$، علاوه بر این ثابت کردند $X_{ew} \leq_{st} Y_{ew} \iff X \leq_{ew} Y$. فرناندز (۱۹۹۸) نیز ثابت کرد اگر $X_{ew} \leq_{st} Y_{ew}$ آنگاه $Var(X) \leq Var(Y)$ و $E|X_1 - X_2| \leq E|Y_1 - Y_2|$. لی و شیکد (۲۰۰۴) نشان دادند اگر $X_1, X_2, ..., X_n$ و $Y_1, Y_2, ..., Y_n$ دو نمونه مستقل از X و Y باشند و $1 \leq r \leq n-1$، $[\min\{X_1, X_2, ..., X_{n-r}\}]_{ew} \leq_{st} [\min\{Y_1, Y_2, ..., Y_{n-r}\}]_{ew}$ آنگاه $E(X_{(r+1)} - X_{(r)}) \leq E(Y_{(r+1)} - Y_{(r)})$ که در آن $X_{(i)}$ همان iامین آماره ترتیبی است.

بلزونس (۱۹۹۹) نشان داد که X دارای خاصیت $NBUE$ است اگر و تنها اگر $X \leq_{ew} Y$ که در آن Y یک متغیر تصادفی نمایی با میانگین $E(X)$ است. از طرفی چون $E(X) = E(Y)$ داریم

$$X \leq_{ew} Y \iff X \geq_{ttt} Y$$

(کوچار و همکاران ۲۰۰۲).

متغیر تصادفی X_{ttt} که معرف مشاهده کل زمان آزمون است در مفاهیم IFR و $IFRA$ نیز صدق می کند اگر Y دارای توزیع پارتو باشد یعنی $\bar{F}(y) = \frac{1}{1+y}, y \geq 0$، به سادگی می توان نشان داد که X_{ttt} دارای خاصیت IFR است اگر و تنها اگر $X \leq_c Y$ و علاوه بر این اگر $X \leq_* Y$ آنگاه قضیه ۱ بارتازوویچ (۱۹۹۵) نتیجه می دهد X_{ttt} دارای خاصیت $IFRA$ است.

لی و شیکد(۲۰۰۴) ثابت کردند

$$X \leq_{dmrl} Y \iff X_{ew} \geq_* Y_{ew} \text{ که در آن } X \leq_{dmrl} Y$$

اگر $\frac{\frac{1}{\mu_Y}\int_{G^{-1}(p)}^{\infty}\bar{G}(x)dx}{\frac{1}{\mu_X}\int_{F^{-1}(p)}^{\infty}\bar{F}(x)dx}$ برای $p \in (0,1)$ صعودی باشد و می گوییم X کوچکتر از Y در ترتیب تصادفی $dmrl$ است.

قضیه ۳. فرض کنید X و Y دو متغیر تصادفی نامنفی باشند. آنگاه

الف) $X \leq_{icv} Y \Longrightarrow X_{ttt} \leq_{icv} Y_{ttt}$

ب) $X \leq_{disp} Y \iff X_{ttt} \leq_{disp} Y_{ttt} \iff X_{ew} \leq_{disp} Y_{ew}$

شیکد شانتیکومار (۱۹۹۴) نشان دادند که متغیر تصادفی نامنفی X دارای خاصیت $HNBUE$ است اگر $X \geq_{icv} Y$ که در آن Y یک متغیر تصادفی نمایی با میانگین $E(X)$ است و $Y_{ttt} = U(0, E(X))$. بنابراین قضیه ۳ نتیجه می دهد که اگر متغیر تصادفی نامنفی X دارای خاصیت $HNBUE$ باشد آنگاه $X_{ttt} \geq_{icv} U(0, E(X))$.

کاربردها

کاربرد در بیمه

اگر متغیر تصادفی با توزیع $F(x)$ میزان ضرر یک بیمه را نشان دهد، آنگاه متوسط ضرر(امید ریاضی ضرر) برابر است با $E(X) = \int_0^\infty \bar{F}(x)dx$. متوسط ضرر، $E(X)$ در پرداخت اضافی برای هر قرارداد بیمه مورد توجه قرار می باشد. به عبارت دیگر وقتی ریسک خطر بیشتر است مقدار ضرر وزن بیشتری می گیرد(برای اطلاع بیشتر به لی وشیکد ۲۰۰۷ مراجعه شود).

تعریف ۵. فرض کنید X و Y دو مخاطره باشند. اگر $\Pi_X(t) \leq \Pi_Y(t), t \geq 0$ آنگاه X را کوچکتر از Y در ترتیب توقف زیان نامیم و با نماد $X \leq_{st} Y$ نشان می دهیم که در آن $\Pi_X(t) = \int_t^\infty \bar{F}_X(x)dx, t \geq 0$ را تبدیل توقف زیان مخاطره X گوییم(مولر ۱۹۹۶).

از تعریف فوق به راحتی نتیجه می شود:

$$\Pi_X(X) = \int_X^\infty \bar{F}_X(x)dx = X_{ew} \qquad (۱)$$

تابع Π_X دارای خواص زیر است:

۱) Π_X نزولی و محدب است.

۲) مشتق سمت راست Π_X یعنی $D^+\Pi_X$ وجود دارد و $-1 \leq D^+\Pi_X \leq 0$.

۳) $lim_{t\to\infty}\Pi_X(t) = 0$.

برای هر تابع $\Pi : R^+ \to R$ که دارای خواص بالا باشد یک مخاطره X وجود دارد بطوریکه Π تبدیل توقف زیان X است. تابع زیان X را می توان به صورت $F_X(t) = D^+\Pi_X(t) + 1$ نوشت و همچنین $\Pi_X(0) = E(X)$.

هورلیمن (۲۰۰۱) مقدار کلی تغییر قیمت برای متغیر تصادفی نامنفی X با تابع توزیع $F(x)$ را به صورت زیر معرفی نمود.

$$P_3(x) = \int_0^\infty (\bar{F}(x))^\rho dx \quad , \quad 0 < \rho < 1$$

علاوه بر این ویژگی های زیر نیز برای $P_3(X)$ برقرار هستند(هورلیمن ۲۰۰۱):

۱) $P_3(X) \geq E(X)$

۲) $P_3(X) \leq sup(X)$

۳) $P_3(X+Y) \leq P_3(X) + P_3(Y)$

۴) اگر $X \leq_{st} Y$ آنگاه $P_3(X) \leq P_3(Y)$

نتیجه ۲. اگر X و Y دو متغیر تصادفی نامنفی باشند

$$X_{ew} \leq Y_{ew} \Longrightarrow P_3(X) \leq P_3(Y).$$

کاربرد در قابلیت اعتماد

در این قسمت کاربردی از ترتیب های تصادفی TTT و EW در نظریه قابلیت اعتماد را مورد بررسی قرار می دهیم. ابتدا به معرفی برخی تعاریف و اصطلاحات در زمینه قابلیت اعتماد به ویژه مبحث سیستم ها می پردازیم.

یک مولفه دو وضعیتی را به صورت زیر تعریف می کنیم:

$X_i = 1$ اگر مولفه iام کار کند و $X_i = 0$ اگر مولفه iام کار نکند برای $i = 1, 2, ..., n$ که n تعداد مولفه های یک سیستم است. به تعداد مولفه های یک سیستم مرتبه آن سیستم گویند.

فرض کنید $\Psi(X)$ تابعی از X است که $X = (X_1, ..., X_n)$ و X_i وضعیت مولفه ها را نشان می دهد. به $\Psi(X)$ تابع ساختار سیستم گوییم اگر به صورت زیر تعریف شود:

$\Psi(X) = 1$ اگر سیستم کار کند و $\Psi(X) = 0$ اگر سیستم کار نکند.

طول عمر یک سیستم سری با n مولفه که طول عمر هر کدام از مولفه ها (X_i) متغیرهای تصادفی مستقل و هم توزیع است برابر است با کوچکترین طول عمر در بین مولفه ها یعنی $\min\{X_i\}$ و طول عمر یک سیستم موازی با n مولفه که طول عمر هر کدام از مولفه ها (X_i) متغیرهای تصادفی مستقل و هم توزیع است برابر است با بزرگترین طول عمر در بین مولفه ها یعنی $\max\{X_i\}$.

قضیه ۴. فرض کنید $(X_1, X_2, ..., X_n)$ و $(Y_1, Y_2, ..., Y_n)$ دو بردار تصادفی با مولفه های مستقل و هم توزیع و مستقل از هم و علاوه بر این، X_i و Y_i و $1 \leq i \leq n$، معرف طول عمر مولفه های دو سیستم باشند

الف) اگر سیستم سری باشد و $Y_i \leq_{ttt} X_i$ آنگاه $\min\{X_1, ..., X_n\} \leq_{ttt} \min\{Y_1, ..., Y_n\}$.

ب) اگر سیستم موازی باشد و $Y_i \leq_{ew} X_i$ آنگاه $\max\{X_1, ..., X_n\} \leq_{ew} \max\{Y_1, ..., Y_n\}$.

نتایج

ویژگی های ترتیب های تصادفی TTT و EW و ارتباط آنها با سایر ترتیب های تصادفی مورد بررسی قرار گرفته است. تحقیق در ارتباط با ترتیب های دیگر که در این نوشته ذکر نشده و بدست آوردن ترتیب تصادفی TTT و EW برای رکوردها در آینده تحقیق مدنظر می باشد.

References

[1] Barlow, R. E. Bartholomev, D.J, Bremner, J.M and Brunk, H.d. (1972). Statistical Inference under order restrictions. Wiley, New York.

[2] Barlow, R. E. and Daksum, K. A. (1972). Isotonic tests for convex ordering. Mathematical Statistics and Probability, vol. i, pp. 293-323.

[3] Barlow, R. E. and Proschan, F. (1975). Statistical Theory of Reliability and Life Testing: Reliability Models.

[4] Bartoszewicz, J. (1986). Dispersiv ordering and the total time on test transformation. Statist. Probab. Lett . 4, 285-288.

[5] Bartoszewicz, J. (1995). Stochastic order relations and the total time on test transform. Statist. Probab. Lett. 22, 103-110.

[6] Bartoszewicz, J. (1998). Application of a general composition theorem to the star order of distributions. Statist. Probab. Lett. 38, 1-9.

[7] Hurlimann, W. (2001). Distribution-free comparison of pricing principles . Insurance Math. Econom. 28, 351-360.

[8] Belzunce, F. (1999). On a characterization of right spread order by the increasing convex order. Stat. Probab.Lett. 45, 103-110.

[9] Kochar, S. C., Li, X. and Shaked, M. (2002). The total time on test transform and the excess wealth stochastic orders of distributions. Adv.inAppl,Probab.34,826-845

[10] Li, X. and Shaked, M. (2004). The observed total time on test and the observed excess wealth. Statist. Probab. Lett. 68, 247-258.

[11] Muller, A. (1996). Ordering of risks : A comparative study via stop -loss transform. Insurance Math. Econom. 17, 215-222.

[12] Shaked, M. and Shanthikumar, J. G. (2007). Stochastic Orders. Springer, New York.

[13] Shaked, M. and Shanthikumar, J. G. (1994). Stochastic Orders and their applications.Academic Press,Inc. New York.

[14] Shaked, M. and Shanthikumar, J. G. (1998). Two variability Orders. Probab. Eng. Inform. Sci. 12, 1-23

The 45th Annual Iranian Mathematics Conference
August 26-29 2014
چهل و پنجمین کنفرانس ریاضی ایران
۴ - ۷ شهریور ۱۳۹۳

مدل اتو رگرسیو آستانه‌ای با دو متغیر آستانه

اسماعیل امیری *، عضو هیأت علمی گروه آمار، دانشگاه بین المللی امام خمینی (ره) ، amiri@SCI.ikiu.ac.ir
زینب یوسفی پور ، دانشجوی کارشناسی ارشد، دانشگاه بین المللی امام خمینی (ره) ، p.yousefipoor@yahoo.com
رامین کاظمی ، عضو هیأت علمی گروه آمار، دانشگاه بین المللی امام خمینی (ره)، r.kazemi@SCI.ikiu.ac.ir

چکیده: مدل اتورگرسیو آستانه‌ای مدلی قطعه‌ای خطی است که برای مدل‌سازی رفتار غیرخطی بسیاری از سری‌های زمانی به‌ویژه سری‌های زمانی مالی کاربردهای فراوانی یافته است. در بسیاری از کاربردها، مدل اتورگرسیو آستانه‌ای با تنها یک متغیر آستانه مورد استفاده قرار گرفته است. یک مدل اتورگرسیو آستانه‌ای با دو متغیر آستانه معرفی و در یک رهیافت تجربی با استفاده از آزمون‌های نسبت درستنمایی و تقریب توزیع حدی آماره‌ها به کمک روش بوت استرپ، امکان چند رژیمی بودن سری زمانی بازده شاخص بورس داوجونز مورد بررسی قرار گرفته است. با انتخاب دو متغیر برونزا به عنوان متغیرهای آستانه، وجود ۴ رژیم در سری زمانی آزمون شده و سپس یک مدل اتورگرسیو آستانه‌ای با دو متغیر آستانه به بازده شاخص بورس داوجونز برازش و پارامترهای آن براورد شده است.

کلمات کلیدی: مدل‌های اتورگرسیو آستانه‌ای، روش بوت استرپ، توزیع حدی، آزمون نسبت درستنمایی.

مقدمه

بسیاری از سری‌های زمانی مالی تحت تاثیر عوامل مختلف، دارای رفتار غیرخطی هستند. برای مدل سازی این‌گونه سری‌های زمانی معمولاً مدل‌های غیرخطی ممکن است گزینه مناسبی باشند. یکی از مدل‌های غیرخطی که بسیار مورد توجه قرار گرفته است مدل اتورگرسیو آستانه‌ای (TAR) است. این مدل به صورت قطعه‌ای خطی است، لذا بسیاری از ایده‌های مربوط به مدل‌های خطی قابل تعمیم به این نوع مدل می‌باشند. یک ویژگی مهم و جالب مدل TAR قابلیت آن در بخاطر سپردن دوره‌های تناوب نامعلوم و نامنظم است. تانگ و لیم (۱۹۸۰) مدلTAR را معرفی نمودند. تانگ (۱۹۹۰) بسیاری از جزئیات دیگر مدل TAR را ارائه و از این مدل برای پیش‌بینی نوسانات قیمت سهام استفاده نمود. بسیاری ازمؤلفان مدل‌های TAR با یک متغیر آستانه را مورد استفاده قرار داده‌اند، در حالی‌که در بسیاری از برنامه‌های کاربردی تجربی، مدل با دو متغیر آستانه یا بیشتر ممکن است مناسب‌تر باشد. در این مقاله یک مدل TAR با دو متغیر آستانه برای مدل‌سازی بازده قیمت سهام شاخص بورس داوجونز مورد استفاده قرار گرفته است و همچنین قیمت‌های گذشته شاخص و حجم معاملات بازار به عنوان متغیرهای آستانه در نظر گرفته شده اند. هدف این است که در یک رهیافت تجربی روش‌های آزمون و براورد پارامترهای یک مدل اتورگرسیو آستانه‌ای با دو متغیر آستانه را معرفی و با استفاده از دو متغیر برونزا به عنوان متغیرهای آستانه

یک مدل اتورگرسیو آستانه‌ای با دومتغیر آستانه را به سری زمانی بازده شاخص بورس داوجونز برازش و پارامترهای مدل را براورد نماییم.

در بخش ۲ مدل اتورگرسیو آستانه‌ای با دو متغیر آستانه و شیوه براورد پارامترها معرفی شده است. در بخش ۳ آزمون نسبت درستنمایی برای تعیین تعداد رژیم‌ها و استفاده از روش بوت استرپ برای ارزیابی روش براورد تشریح شده است. بخش ۴ اختصاص به معرفی داده‌ها، متغیرهای آستانه و مقدارهای براورد شده دارد و در بخش آخر نیز نتیجه گیری بیان شده است.

مدل TAR با دو متغیر آستانه و شیوه براورد پارامترها

با استفاده از مقاله‌ی چِن و همکاران (۲۰۱۱) مدل اتورگرسیو آستانه‌ای با دو متغیر زیر را در نظر بگیرید که مشاهدات y_t در چهار رژیم طبقه‌بندی شده‌اند:

$$y_t=\sum_{j=1}^{4}\psi_t^{(j)}(\gamma^\circ)(\beta_0^{(j)}+\sum_{i=1}^{p_j}\beta_i^{(j)}y_{t-i}+u_t), \qquad (1)$$

که در آن $\psi_t^{(j)}(\gamma_0)$ تابع نشانگر است و برابر یک است اگر در شرط آستانه صدق کند. در غیر این صورت برابر صفر است. به ازای $j=1,2,3,4$ داریم:

$$\begin{aligned}\psi_t^{(1)}(\gamma^\circ) &= I(z_{1t}\le\gamma_1^\circ, z_{2t}\le\gamma_2^\circ),\\ \psi_t^{(2)}(\gamma^\circ) &= I(z_{1t}\le\gamma_1^\circ, z_{2t}>\gamma_2^\circ),\\ \psi_t^{(3)}(\gamma^\circ) &= I(z_{1t}>\gamma_1^\circ, z_{2t}\le\gamma_2^\circ),\\ \psi_t^{(4)}(\gamma^\circ) &= I(z_{1t}>\gamma_1^\circ, z_{2t}>\gamma_2^\circ).\end{aligned}$$

از سویی

$z_t=(z_{1t},z_{2t})$ بردار متغیرهای آستانه‌ای و $\gamma^\circ=(\gamma_1^\circ,\gamma_2^\circ)\in\Omega$ بردار پارامتر آستانه است که باید براورد شود و $\Omega=[\underline{\gamma}_1,\overline{\gamma}_1]\times[\underline{\gamma}_2,\overline{\gamma}_2]$ زیر مجموعه‌ای از تکیه‌گاه z_t است. $p_j(j=1,2,3,4)$ مرتبه‌ی اتو رگرسیو در هر رژیم است. $\beta^{(j)}=(\beta_0^{(j)},\beta_1^{(j)},...,\beta_{p_j}^{(j)})$ پارامترهای ساختاری هستند و برای برخی $i\neq j$ داریم $\beta^{(i)}\neq\beta^{(j)}$.

درون هر رژیم یک مدل اتورگرسیو خطی قرار دارد. متغیرهای آستانه‌ای z_{1t} و z_{2t} می‌توانند متغیرهای برونزا یا توابعی از تاخیرهای y_t باشند. با در نظر گرفتن $\{y_t,z_t\}_{t=1}^{T}$، هدف براورد پارامترهای آستانه γ و پارامترهای ساختاری $\beta^{(j)}$ است. بدون از دست دادن کلیت،

به ازای $j=1,2,3,4$ فرض می‌کنیم $p=\max\{p_j\}$ و برای $\beta_q^{(j)}=0, q>p_j$.

برای ساده شدن محاسبات، مدل (۱) را به می‌توان به شکل ماتریسی

$$Y=\sum_{j=1}^{4}I_j(\gamma^\circ)X\beta^{(j)}+U, \qquad (2)$$

نوشت که در آن

$$X=\begin{pmatrix}1 & y_{T-1} & y_{T-2} & \cdots & y_{T-p}\\ 1 & y_{T-2} & y_{T-3} & \cdots & y_{T-p-1}\\ & & \cdots & & \\ 1 & y_p & y_{p-1} & \cdots & y_1\end{pmatrix}$$

به‌طوری‌که X یک ماتریس $(T-p)\times(p+1)$ و همچنین $I_j(\gamma^\circ)=diag\{\psi_T^{(j)}(\gamma_0),\psi_{T-1}^{(j)}(\gamma_0),...,\psi_{p+1}^{(j)}(\gamma_0)\}$ و

$Y=(y_T,y_{T-1},...,y_{p+1})', U=(u_T,u_{T-1},...,u_{p+1})'$.

مفروضات زیر را در نظر می‌گیریم:

۱) y_t مانای ارگودیک است و $E(y_t^4)<\infty$.

۲) $\{u_t\}$ دنباله‌ای از خطاهای مستقل و هم‌توزیع با توزیع نرمالی با میانگین صفر و واریانس σ^2 هستند.

۳) متغیرهای آستانه z_{1t} و z_{2t} اکیداً مانا هستند و توزیع توام پیوسته $F(\gamma)$ را دارند که نسبت به هردو متغیر مشتق پذیر است. فرض کنید $f(\gamma)$ نشان دهنده‌ی تابع چگالی توأم باشد و $f_i(\gamma)=\frac{\partial F(\gamma)}{\partial\gamma_i}$. فرض می‌کنیم که $0<f_i(\gamma)\le\overline{f_i}<\infty$ برای $i=1,2$. این مفروضات برای سازگاری براورد مقادیر آستانه لازم می‌باشند. با فرض

چهل و پنجمین
کنفرانس ریاضی ایران
۴-۷ شهریور ۱۳۹۳
The 45th Annual Iranian Mathematics Conference
August 26-29 2014

معلوم بودن $\gamma=(\gamma_1,\gamma_2)$، براوردگر حداقل توان‌های دوم (CLS) برای $\beta^{(j)}$ به صورت

$$\hat{\beta}^{(j)}(\gamma)=(X'I_j(\gamma)X)^{-1}X'I_j(\gamma)Y \tag{3}$$

و مجموع توان دوم باقیمانده به شکل

$$RSS_T(\gamma)=||\sum_{j=1}^{4}I_j(\gamma^{\circ})X\beta^{(j)}+U-\sum_{j=1}^{4}I_j(\gamma)X\hat{\beta}^{(j)}(\gamma)||^2$$

است. براوردگر γ° را به عنوان مقداری که $RSS_T(\gamma)$ را مینیمم می‌کند به فرم زیر تعریف می‌شوند:

$$\hat{\gamma}=arg\min_{\gamma\in\Omega}RSS_T(\gamma). \tag{4}$$

براوردگرهای ساختاری مبتنی بر مقادیر آستانه به، به صورت

$$\hat{\beta}^{(j)}(\hat{\gamma})=(X'I_j(\hat{\gamma})X)^{-1}X'I_j(\hat{\gamma})Y. \tag{5}$$

هستند. می‌توان نشان داد که براوردگرهای $(\hat{\gamma}_1,\hat{\beta}^{(j)}(\hat{\gamma}))$ سازگار هستند.

آزمون و براورد آستانه

برای تعیین تعداد رژیم‌ها، ابتدا فرض صفر را بدون اثر آستانه به صورت $H_0:\beta^{(1)}=\beta^{(2)}=\beta^{(3)}=\beta^{(4)}$. در نظر می‌گیریم تحت فرض صفر، تنها یک رژیم وجود دارد. آماره آزمون نسبت درستنمایی به صورت

$$J_T=max_{\gamma\in\Omega}(T-p)\frac{\tilde{\sigma}^2-\hat{\sigma}^2(\gamma)}{\hat{\sigma}^2(\gamma)}. \tag{6}$$

تعریف می‌شود که $(T-p)\tilde{\sigma}^2$ مجموع توان دوم تحت فرض صفر است، در حالیکه $(T-p)\hat{\sigma}^2(\gamma)$ مجموع توان دوم باقیمانده تحت فرض مقابل است. اگر H_0 رد نشود، پس مدل یک مدل اتورگرسیو AR ساده است. رد فرض صفر بیان می‌کند که بیش از یک رژیم در مدل وجود دارد. براوردگر آستانه به صورت $\hat{\gamma}=arg\min\hat{\sigma}^2(\gamma)=arg\max J_T(\gamma)$ تعریف می شود. چون γ تحت فرض صفر نامعلوم است پس توزیع مجانبی $J_T(\hat{\gamma})$ خی‌دو (χ^2) استاندارد نیست. هَنسن (۱۹۹۶) نشان داد که توزیع مجانبی با روش بوت‌استرپ زیر تقریب زده می‌شود.

فرض کنید $u_t^*;(t=1,...,T)$ ها مستقل و هم‌توزیع نرمال استاندارد ($(N(0,1))$) باشند. $y_t^*=u_t^*$ را محاسبه، سپس y_t^* را روی $x_t=(1,y_{t-1}^*,y_{t-2}^*,...,y_{t-p}^*)$ رگرس می‌کنیم و $J_T^*(\gamma)=(T-p)\frac{\tilde{\sigma}^{*2}-\hat{\sigma}^{*2}(\gamma)}{\hat{\sigma}^{*2}(\gamma)}$ را محاسبه و سرانجام $J_T^*=\max_{\gamma\in\Omega}J_T^*(\gamma)$ را بدست می آوریم. توزیع J_T^* تحت فرض صفر، همگرایی ضعیف در احتمال به توزیع J_T دارد. بنابراین می‌توان مقدار بوت‌استرپ J_T^* را برای تقریب توزیع مجانبی (تحت فرض صفر) J_T استفاده کرد. برای تعیین تعداد رژیم‌ها از یک رویکرد کل به جزء استفاده می شود. ابتدا مدل دارای سه رژیم در مقابل‌مدل دارای چهار رژیم آزمون می‌شود. هر کدام از فرض‌های

$$(I)\quad H_0:\beta^{(1)}=\beta^{(2)}\qquad (II)\quad H_0:\beta^{(1)}=\beta^{(3)}$$
$$(III)\quad H_0:\beta^{(1)}=\beta^{(4)}\qquad (IV)\quad H_0:\beta^{(2)}=\beta^{(3)}$$
$$(V)\quad H_0:\beta^{(2)}=\beta^{(4)}\qquad (VI)\quad H_0:\beta^{(3)}=\beta^{(4)}$$

را در مقابل فرض مقابل وجود ۴ رژیم آزمون می کنیم.

برای انجام آزمون فرض‌های دو تایی بالا از آزمون نسبت درستنمایی مشابه (۶) و شیوه بوت استرپ استفاده می‌شود.

در بررسی‌های تجربی، مرتبه اتورگرسیو، مقادیر آستانه و ضرایب مدل با یک الگوریتم ساده به این ترتیب براورد می‌شوند. در گام اول یک مدل TAR مرتبه اول براورد و سپس براورد اولیه برای براورد مقادیر آستانه $(\hat{\gamma}_t)$ مورد استفاده قرار می‌گیرد. در گام دوم به شرط مقادیر آستانه بدست آمده از گام اول، معیار اطلاع آکائیکه (AIC) به منظور انتخاب مرتبه‌ی اتورگرسیو در هر رژیم به کار می‌رود (تی‌سِی،۱۹۹۸). در گام سوم برای تعیین تعداد رژیم‌ها آزمون نسبت درستنمایی ترتیبی(متوالی) اجرا می‌شود. در گام چهارم نتیجه بدست آمده از گام سوم برای اصلاح مقادیر آستانه به کار می‌رود و گام دوم و سوم را تا زمانی که همه‌ی براوردگرها همگرا

شوند تکرار می‌شود.

داده‌ها و براوردها

مدل پیشنهادی، به سری‌های بازده روزانه شاخص داوجونز برازش شده است. دوره نمونه از ۱ ژانویه سال ۲۰۰۲ تا ۳۱ دسامبر سال ۲۰۱۲ می‌باشد. سری‌های بازده به عنوان تفاضل لگاریتمی شاخص داوجونز تعریف و با y_t نشان داده می‌شوند. بیش از ۲۷۰۰ مشاهده در نمونه وجود دارد. دو متغیر آستانه برون زا مشابه با گِرانویل (۱۹۶۳)، لی و سوامیناتان (۲۰۰۰) تعریف شده‌اند. اولین و دومین متغیر آستانه به صورت

$$Z_{1t} = \frac{P20_t}{P250_t}, \quad Z_{2t} = \log(V_{t-1}) - V250_{t-1},$$

تعریف مي‌شوند و

$$P250_t = \frac{\sum_{j=1}^{250} P_{t-j}}{250}, \quad P20_t = \frac{\sum_{j=1}^{20} P_{t-j}}{20}.$$

$P20_t$ و $P250_t$ به ترتیب متوسط قیمت برای ۲۵۰ و ۲۰ روز معاملاتی گذشته هستند.از طرفی $V250_t = \frac{\sum_{j=1}^{250} \log(V_{t-j})}{250}$، که V_t حجم معاملات بازار در زمان t است. مدل برازش شده در زیر نشان داده شده است. رژیم I بازده بالا وباثبات، رژیم II بازده کم و نوسانی و رژیم III رژیم خنثی را نشان می دهد.

$$\begin{array}{lll}
z_{1t} > 0/27,\ z_{2t} < 0/92 & y_t = 0/0003 - 0/066x_{t-1} + 0/020x_{t-2} & I \\
 & -0/024x_{t-3} - 0/002x_{t-4} - 0/064x_{t-5} & \\
z_{1t} < 0/27,\ z_{2t} > 0/92 & y_t = 0/0005 - 0/05x_{t-1} - 0/27x_{t-2} & \\
 & +0/23x_{t-3} & II \\
 & y_t = -0/001 - 0/23x_{t-1} & III
\end{array}$$

نتایج

معمولاً مدل‌های آستانه‌ای متعارف تنها شامل یک متغیر آستانه هستند. یک مدل اتورگرسیو با دو متغیر آستانه مورد بررسی قرار گرفته است. از آزمون نسبت درستنمایی برای تشخیص اثر آستانه استفاده شده است. مدل پیشنهادی برای شناسایی رژیم‌های بازار سهام داوجونز به‌کار رفته است. دو متغیر برون زا به عنوان متغیرهای آستانه مورد استفاده قرار گرفته‌اند. نتایج نشان می‌دهد که بازده شاخص داوجونز را می‌توان در سه رژیم، رژیم با بازده بالا و باثبات، رژیم با بازده پایین و نوسانی و یک رژیم خنثی طبقه‌بندی کرد. می‌توان مبانی نظری و عملی پیش‌بینی با این‌گونه مدل‌ها را در پژوهش‌های آتی مورد مطالعه قرار داد.

مراجع

References

[1] Chen, H. Chong, T. T and Bai, J.(2012). Theory and applications of TAR model with two threshold variables. Econometric Reviews, 31(2):142-170 .

[2] Granville, J. (1963). Granvilles New Key to Stock Market Profits. Upper Saddle River, NJ: Prentice Hall.

[3] Hansen, B. E. (1996). Inference when a nuisance parameter is not identified under the nullhypothesis. Econometrica 64: 413-430.

[4] Lee, M. C., Swaminathan, B. (2000). Price momentum and trading volume. Journal of Finance IV: 2017-2069.

[5] Tong,H. (1990). Non-linear time series:a dynamical system approach.Oxford University Press.

[6] Tong, H. and Lim, K. S. (1980). Threshold autoregression, limit cycle and cyclical data. J. R. Statistic. Soc.42(B), 245-292.

[7] Tsay, R. S. (1998). Testing and modeling multivariate threshold models. J. Amer. Statist. Assoc. 93:1188-1202.

The 45th Annual Iranian Mathematics Conference
August 26-29 2014
چهل و پنجمین
کنفرانس ریاضی ایران
۴ - ۷ شهریور ۱۳۹۳

The Comparison Statistical Estimation for HIV Infected People with Different Noise Terms

P.Nabati*, Academic member of School of Mathematics,

Urmia University of Technology, nabati@iust.ac.ir

"Academic member"

Abstract: This paper provides a stochastic model for HIV (Human Immunodeficiency Virus) infections and diffusion of AIDS in Iran. For this purpose the deterministic model of diffusion process will be transferred to stochastic model and then this model will be solved numerically. The effects of colored noise perturbations on the parameters will be investigated. Finally numerical example is performed by using the Euler-Maruyama method in order to show the accuracy of the present work.

Keywords: Stochastic differential equations, Diffusion process, HIV model, Colored noise.

1 INTRODUCTION

In recent decades, Stochastic Differential Equations (SDEs) have been used to model the systems that are subject to vacillations. Diffusion processes have been utilized in modeling several phenomena, for example noisy tumor growth and interest rates. The modeling of population growth in random environment is one of the most suitable application of SDEs. Modeling is a way of making complex data more easily understood. By creating a model, we can earn an overall picture of an event, known in modeling terms as a system. A mathematical model is produced by a mathematical equation designed to resemble actual facts.

One of the problems in most countries is to prevent the publication of deadly virus. The scope of this paper is to investigate the SDE for HIV (Human Immunodeficiency Virus) in Iran using different noise terms. The first case of HIV was diagnoised in IRAN in 1986. To date several works have been presented on the mathematical modeling of ADIS epidemic. For example Arni and Rao via an excellent paper have been presented modeling of ADIS in India. Adituma has been investigate the mathematical modeling of HIV epidemic in Indonesia in 2008-2014. Pasha and Mostafaee examined the diffusion process of ADIS in Iran. Up to now, to our best knowledge the HIV model with nonwhite noises has not been studied before. Since the path of a Wiener process are nowhere differentiable, a white noise can not be considered a stochastic process in the usual way but it can be approximate by conventional stochastic processes with wide spectral bands which are commonly known as colored noise processes. The most famous example of this sort of noise is the Ornestein-Uhlenbeck process. The outline of this paper is as follows. Section 2 describes the stochastic model for HIV. The SDE of HIV with colored noise is stablished in section 3. Numerical experiments are conducted to verify the accuracy of the proposed method in section 4.

*Corresponding Author

2 Stochastic model with white noise

The simplest population growth in a stochastic crowded environment is the form,

$$\frac{dN_t}{dt} = a(t)N(t), \qquad N(0) = N_o, \tag{1}$$

where $N(t)$ is the size of the population at time t, and $a(t)$ is the relative rate of growth at time t. It might happen that $a(t)$ is not completely known, but subject to random environmental effects, so that we have,

$$a(t) = a + \text{"}noise\text{"} = a + b\xi_t$$

where ξ_t is a white noise process of mean zero and variance one, and a, b are constants. By substitute this in Eq. (1), we get,

$$dN_t = aN_t dt + bN_t\xi_t dt, \tag{2}$$

it is reasonable to consider $\xi_t dt$ by a term dB_t, where B_t is a Wiener process. The Eq. (3) is achieved,

$$dN_t = aN_t dt + bN_t dB_t, \tag{3}$$

Eq. (3) known as Geometric Brownian motion. The explicit solution is,

$$N_t = N_o exp(a - \frac{1}{2}b^2)t + bB_t.$$

Let x_t be the number of persons susceptible to infected with HIV in time t, y_t be the number of HIV infected and z_t, the number of death from HIV all at time t in the population. From Arni and Rao, the relationship between these is as follows,

$$\frac{dx_t}{dt} = -r_1 x_t y_t,$$

$$\frac{dy_t}{dt} = r_1 x_t y_t - r_2 y_t, \tag{4}$$

$$\frac{dz_t}{dt} = r_2 y_t,$$

where r_1 and r_2 are constants of proportionality. For a given positive integer n, let $\Delta t = \frac{T}{n}$ and consider the partitions,

$$\Pi_n = \{0, \Delta t, 2\Delta t, \ldots, (n-1)\Delta t, T\}$$

of the interval $[0, T]$. With a simple Euler forward discritization of the Eq. (4), r_1

and r$_2$ can be obtained easily,

$$\hat{r}_1 = \frac{y_{t+\Delta t} - y_t + z_{t+\Delta t} - z_t}{\Delta t y_t z_t},$$

$$\hat{r}_2 = \frac{z_{t+\Delta t} - z_t}{\Delta t y_t}, \tag{5}$$

such that $\lim_{\Delta t \to 0} \hat{r}_1 = r_1$, $\lim_{\Delta t \to 0} \hat{r}_2 = r_2$. Now let us allow some randomness in the r_1 and r_2, then

$$r_1 = r_1 + b_1\xi_t, \; r_2 = r_2 + b_2\xi_t,$$

the stochastic differential equations describing this situation are,

$$dx_t = -r_1 x_t y_t dt - b_1 x_t y_t dB_t$$

$$dy_t = (r_1 x_t y_t - r_2 y_t)dt + (b_1 x_t y_t - b_2 y_t)dB_t$$

$$dz_t = r_2 y_t dt + b_2 y_t dB_t. \tag{6}$$

The solution of these equations are infection diffusion process and the death process. Using the explicit Euler method, the approximation solution of equations can be acquired.

3 Stochastic model with colored noise

A white noise process can not be physically realized but can be approximated by conventional stochastic processes with wide spectral bands which are commonly known as colored noise processes. The stochastic process $\beta(t)$ is called colored noise if it is an Orstein-Uhlenbeck process that satisfies the linear SDE

$$d\beta(t) = \mu\beta(t)dt + \sigma dB(t), \tag{7}$$

where μ, σ are constants. The explicit solution of Eq. (7) is given by

$$\beta(t) = e^{\mu t}(\beta(0) + \sigma\int_0^t e^{-\mu s}dB(s)).$$

Let us consider the noise term $\xi(t)$ as a colored noise process. Therefore

$$r_1 = r_1 + \text{"}colored\, noise\text{"},\quad r_2 = r_2 + \text{"}colored\, noise\text{"},$$

with substitute $d\beta_t$ into equation 6 instead of dB_t, we have,

$$\begin{aligned} dx_t &= -r_1 x_t y_t dt - b_1 x_t y_t d\beta_t \\ dy_t &= (r_1 x_t y_t - r_2 y_t)dt + (b_1 x_t y_t - b_2 y_t)d\beta_t \\ dz_t &= r_2 y_t dt + b_2 y_t d\beta_t. \end{aligned} \quad (8)$$

By means of the Euler method the approximation solution of infection diffusion is equal to,

$$\begin{aligned} y_{t+\Delta t} &= y_t + (r_1 x_t y_t - r_2 y_t)\Delta t \\ &+ (b_1 x_t y_t - b_2 y_t) \cdot (\mu\beta_t \Delta t + \sigma \Delta B_t), \end{aligned} \quad (9)$$

and the number of death is equal to,

$$z_{t+\Delta t} = z_t + r_2 y_t \Delta t + b_2 y_t(\mu\beta_t \Delta t + \sigma\Delta B_t). \quad (10)$$

The Matlab procedure of these equations are given in section 4.

4 Numerical simulation

In order to make clear whether the colored noise obtained in this paper are accurate, we present a simulation in this section. Using the published statistics about HIV in Iran up to 2004 and the number of death from HIV, we estimate the y_t and z_t with the initial values $x_0 = 65540000$, $y_0 = 3680$, $\hat{r_1} = 3 * 10^-8$, $\hat{r_2} = 0.2656$, $n = 20$ and $h = 1$ by using matlab programming and consider the noise as a colored noise process. The estimated values are shown in table 1. Our results are closer to the recorded values.

Table 1: Numerical estimation for the HIV infected people in Iran.

t	2002	2003	2004	2005	2006	2007	2008	2009	2010
y_t	3680	5422	5611	7032	8904	11390	14725	19157	25082
z_t	364	1341	2542	4032	5900	8265	11291	15202	20290

References

[1] t. Y. Aditama, "Mathematic Model Of HIV Epidemic In Indonesia", Director General of Communicable Disease Control and Environmental Health, Jakarta, 2008.

[2] S. R. Arni and Srinivasa Rao. "Mathematical Modeling of AIDS epidemic in Indiana," Current Science, 84(9,10),2003.

[3] P. E. Kloeden and E. Platen, "Numerical solution of stochastic differential equations", Springer-Verlag, Heidlberg, 1992.

[4] E. Pasha and H. R. Mostafaee, "Diffusion Process of AIDS in IRAN", Mathematical Sciences, 2007.

[5] Y. Hu and H. Long, "On the singularity of least squares estimator for mean reverting α stable motions", Acta Mathematica Scientia, 2009.

[6] D. Verotta, "Models and estimation methods for clinical HIV-1 data", Journal of Computational and Applied Mathematics 184, 275-300, 2005.

The 45th Annual Iranian Mathematics Conference
August 26-29 2014
چهل و پنجمین کنفرانس ریاضی ایران
۴-۷ شهریور ۱۳۹۳

تحلیل بیزی مدل نظریه پاسخ سؤال به‌وسیله تابع پیوند چوله نرمال

مریم آقاسی ، دانشجوی کارشناسی ارشد آمار ریاضی، دانشگاه سمنان، maryamaghasi۶۸۰@yahoo.com

امید کریمی ، عضو هیأت علمی گروه آمار، دانشگاه سمنان، omid.karimi@profs.semnan.ac.ir

چکیده: نظریه پاسخ سؤال یک نظریه جامع آماری برای مطالعه و تعیین ویژگی‌های سؤال است تا به‌توان بر اساس آن درباره وضعیت آزمودنی‌ها و چگونگی سنجش توانایی‌ها قضاوت نمود. نظریه پاسخ سؤال مدل سازی روابط چند متغیره بین پاسخ‌ها و توانایی‌ها از n فرد و k سؤال را بیان می‌کند. این مدل شامل متغیر پنهان (رابطه‌ی توانایی فرد) و مجموعه پارامترهای وابسته به سؤال است. در این مدل‌ها معمولا از توابع پیوند متقارن (مدل پروبیت نرمال) برای مدل سازی احتمال پاسخ درست استفاده می‌شود. اما همیشه پیوندهای متقارن ارائه مناسبی برای بعضی از مجموعه داده‌ها ندارند. تابع پیوند نامتقارن با استفاده از توزیع‌های نامتقارن چوله می‌تواند گزینه‌ای مناسب برای تحلیل این‌گونه داده‌ها باشد. در این مقاله مدل نظریه پاسخ سؤال را با استفاده از یک تابع پیوند چوله مورد مطالعه قرار می‌دهیم. چولگی در این مدل را می‌توان توسط توزیع چوله نرمال مدل‌بندی و به‌عنوان پارامتر جدید سؤال مورد تحلیل قرار داد. رهیافت بیزی مدل برای مجموعه داده‌های وزن انجام شده و مقایسه دو مدل پروبیت چوله و متقارن با معیار اطلاع کیبش صورت گرفته است. علارغم اینکه پارامتر چولگی پیچیدگی‌هایی را در مدل وارد می‌کند اما برازش مناسبی را فراهم می‌سازد.

کلمات کلیدی: مدل پروبیت چوله، نظریه پاسخ سؤال، استنباط بیزی، مدل پروبیت نرمال.

مقدمه

نظریه پاسخ سؤال [1] (IRT) اولین بار توسط لرد(۱۹۸۰) بیان و از آن به‌عنوان روش‌های نوین در حوزه سنجش و اندازه‌گیری استفاده شد. در واقع، این نامگذاری مبتنی بر اصطلاح بنیادی این حوزه یعنی منحنی ویژه سؤال [2] (ICC) است که توسط تاکر(۱۹۴۶) برای نشان دادن رابطه بین احتمال پاسخ درست یک سؤال و یک متغیر مستقل به‌کار رفته است. مدل IRT شامل متغیرپنهان، رابطه توانایی فرد و مجموعه پارامترهای وابسته به سؤال را توضیح می‌دهد. ون‌درلیندن و همبلتون (۱۹۹۷) مدل سازی پاسخ

[1] Item Response Theory
[2] Item Characteristic Curve

دوگانه را از طریق احتمال پاسخ درست به‌صورت $p_{ij} = F(m_{ij})$ تعریف کردند و آن را منحنی ویژه سؤال ICC نامیدند. که در آن $m_{ij} = a_j u_i - b_j$، $i = 1, \ldots, n$ و $j = 1, \ldots, k$، a_j پارامتر وابسته به سؤال هستند که پارامتر تمیز [۳] و b_j پارامتر دشواری [۴] را معین می‌کند. u_i متغیر پنهان وابسته به توانایی فرد i ام می‌باشد و دارای توزیع نرمال استاندارد است.

به‌طورکلی معادله خطی $F^{-1}(\cdot)$ تابع پیوند نامیده می‌شود. دو حالت خاص آن به‌صورت $F(\cdot) = \Phi(\cdot)$ و $F(\cdot) = L(\cdot)$ است که $\Phi(\cdot)$ تابع توزیع تجمعی نرمال استاندارد و $L(\cdot)$ تابع توزیع تجمعی لجستیک استاندارد است. مدل پروبیت IRT توسط البرت و قوش (۲۰۰۰) و مدل لجستیک IRT توسط بیرن بایوم (۱۹۶۸) بیان شده است. از ویژگی‌های خاص هر دو مدل طبیعت متقارن پروبیت و تابع خطی لجستیک و ICCهای مربوطه است. همان‌طور که چن و همکاران (۱۹۹۹) تاکید کردند همیشه پیوندهای متقارن ارائه خوبی برای برخی از مجموعه داده‌ها نمی‌دهند.

چولگی مدل پروبیت در این مقاله با استفاده از خانواده توزیع‌های چوله نرمال مدل‌بندی می‌شود. همچنین تحلیل بیزی این مدل و تعمیم آن به حالت کلی‌تر توزیع چوله نرمال بسته بیان می‌گردد. به‌علاوه مقایسه مدل‌ها با معیار اطلاع کیبش [۵] (DIC) انجام و نهایتا مدل پیشنهادی در یک مثال واقعی ارزیابی می‌شود.

توزیع چوله نرمال

یکی از راه‌های مدل بندی پروبیت نامتقارن استفاده از توزیع‌های چوله است که از مهمترین آن می‌توان به توزیع چوله نرمال اشاره کرد. تابع چگالی احتمال (pdf) توزیع چوله نرمال به‌صورت $\phi_{SN}(z;\lambda) = 2\Phi(z)\Phi(\lambda z)$ می‌باشد که در آن ϕ و Φ به‌ترتیب pdf و تابع توزیع نرمال استاندارد (cdf) می‌باشند و با $z \sim SN(\lambda)$ نشان می‌دهیم، پارامتر λ میزان چولگی را کنترل می‌کند، وقتی $\lambda > 0$ چولگی مثبت و $\lambda < 0$ چولگی منفی است. در توزیع نرمال $\lambda = 0$ است. تابع توزیع تجمعی $Z \sim SN(\lambda)$ به‌صورت:

$$\Phi_{SN}(z;\lambda) = 2\Phi_2\left(\begin{pmatrix} z \\ 0 \end{pmatrix}; \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & -\delta \\ -\delta & 1 \end{pmatrix}\right)$$

است که در آن Φ_2، cdf توزیع نرمال استاندارد دو متغیره با ضریب همبستگی $-\delta$ است و برای ساده‌گی به‌صورت $\Phi_2((z,0)^T, -\delta)$ نمایش می‌دهیم، که در آن $\delta = \frac{\lambda}{\sqrt{1+\lambda^2}}$ است.

مدل پروبیت چولهIRT

مدل پروبیت چوله IRT به‌صورت

$$\begin{aligned} Y_{ij}|u_i, a_j, b_j, \lambda_j &\sim Bern(p_{ij}), \qquad i = 1, \ldots, n \\ p_{ij} &= p[Y_i = 1|u_i, a_j, b_j, \lambda_j], \qquad j = 1, \ldots, k \\ &= \Phi_{SN}[m_{ij}; \lambda_j] \\ &= 2\Phi_2[(m_{ij}, 0)^T; -\delta_j], \\ m_{ij} &= a_j u_i - b_j \end{aligned} \tag{1}$$

تعریف می‌شود. که در آن Y_{ij} مستقل شرطی از u_i است. یعنی برای سؤال‌های متفاوت پاسخ‌ها مستقل هستند. $a = (a_1, \ldots, a_k)^T$، $b = (b_1, \ldots, b_k)^T$، $\lambda = (\lambda_1, \ldots, \lambda_k)^T$، $y = (y_{11}, \ldots, y_{kn})^T$ قرارمی‌دهیم، اگر D_{obs} را داده‌های مشاهده شده تعریف کنیم، تابع درستنمایی برای مدل پروبیت چوله IRT به‌صورت

$$L(u, a, b, \lambda|D_{obs}) = \Pi_{i=1}^{n}\Pi_{j=1}^{k}[\Phi_{SN}(m_{ij}, \lambda_j)]^{y_{ij}} \times [1 - \Phi_{SN}(m_{ij}, \lambda_j)]^{1-y_{ij}}$$

است. که شامل n+3k پارامتر نامعلوم است، با افزایش تعداد آزمون‌شونده‌ها و تعدادسؤال‌ها تعداد پارامترها افزایش می‌یابد، بنابراین با پارامترهای نامعلوم زیادی

[۳] Discrimination Parameter
[۴] Difficulty Parameter
[۵] Deviance Information Criterion

مواجه هستیم. p_{ij} احتمال شرطی جواب صحیح برای سؤال زام با متغیر پنهان u_i و متناظر با iامین آزمون است که به آن پروبیت چوله ICC گفته می‌شود. اگر در معادله (۱) $\lambda = 0$ باشد $p_{ij} = \Phi(m_{ij})$ می‌شود که همان مدل پروبیت نرمال است. بنابراین λ به‌عنوان پارامتر چوله در نظر گرفته می‌شود.

مدل IRT چوله نرمال بسته

در این بخش مدل IRT چوله را به حالت مدل IRT چوله نرمال بسته [۶] (CSN) تعمیم می‌دهیم. توزیع CSN اولین بار توسط دامینگوس و همکاران (۲۰۰۳) معرفی شد. چگالی توزیع به‌صورت زیر تعریف می‌شود:

$$\begin{aligned} CSN_{n,q}(\mu,\Sigma,\Gamma,\nu,\Delta) &= [\Phi_q(0;\nu,\Delta+\Gamma\Sigma\Gamma')]^{-1} \\ &\times\Phi_q(\Gamma(x-\mu);\nu,\Delta) \\ &\times\phi_n(x;\mu,\Sigma) \end{aligned}$$

که در آن $\phi_n = (\cdot;\mu,\Sigma)$ و $\Phi(\cdot,\nu,\Delta)$ به‌ترتیب تابع چگالی n بعدی و تابع توزیع q بعدی نرمال چند متغیره با میانگین μ و ν و ماتریس کوواریانس Σ و Δ هستند، Γ پارامتر چولگی نامیده می‌شود. مدل IRT شامل k سؤال و n آزمون شونده معادل تعریف $Y_{ij} = \begin{cases} 1 & Z_{ij} > 0 \\ 0 & Z_{ij} \le 0 \end{cases}$ است که در آن $Z_{ij} \sim CSN_{1,1}(m_{ij},1,-\lambda_j,0,1)$ می‌باشد. در حالت خاص که $\lambda_j = 0$، $j = 1,\dots,n$ نتایج مشابه با مدل پروبیت IRT متقارن می‌باشد. برای چشم پوشی از به‌کارگیری درستنمایی برنولی متغیر پنهان Z_{ij} معرفی شده است و $z = (z_{11},\dots,z_{kn})^T$، $D_1 = (z^T,y^T)^T$ قرار می‌دهیم. تابع درستنمایی مدل پروبیت IRT به‌صورت

$$L(u,a,b,\lambda|D_1) \propto \Pi_{i=1}^n \Pi_{j=1}^k \phi_{CSN}(Z_{ij};m_{ij},1,-\lambda_j,0,1)p(y_{ij}|z_{ij})$$

به‌دست می‌آید، که در آن $p(y_{ij}|z_{ij}) = I(z_{ij} > 0)I(y_{ij} = 1) + I(z_{ij} \le 0)I(y_{ij} = 0)$ است و I تابع نشانگر می‌باشد.

تحلیل بیزی مدل

با استفاده از استقلال پارامترها، مدل پیشین به‌صورت زیر در نظر گرفته شده است:

$\pi(u,a,b,\lambda) = \Pi_i^n \pi(u_i,0,1)\Pi_j^k \pi_1(a_j)\pi_2(b_j)\pi_3(\lambda_j)$

معمولا توزیع‌های آگاهی بخش برای a_j استفاده می‌شود. بنزن و همکاران (۲۰۰۶)، از توزیع نیمه نرمال [۷] $N(\mu_a,\sigma_a^2)I(a_j > 0)$ با مقادیرمشخص μ_a و σ_a^2 استفاده کرده‌اند. به‌طور کلی نیازمند ساختار پیشین برای توسعه اطلاعات و بررسی پارامترها هستیم. برای پارامتری کردن دلتا توزیع پیشین را $\delta_j = \frac{\lambda_j}{\sqrt{1+\lambda_j^2}}$ در نظر می‌گیریم. مقادیر در فاصله $[-1,1]$ در نظر گرفته می‌شود. در آن صورت پیشین به‌صورت $\delta_j \sim U(-1,1)$ است، معادل $\lambda_j \sim T(0,0.5,2)$ که در آن $T(\mu,\sigma^2,v)$ به معنی توزیع t استیودنت با پارامتر مکان μ و مقیاس σ^2 و v درجه آزادی است. برای انجام برآورد بیزی می‌توان از درستنمایی برنولی استفاده کرد. محاسبه انتگرال حاشیه‌ای خیلی دشوار است و از دو رهیافت داده افزایی استفاده می‌شود که این رهیافت‌ها از روش MCMC به‌دست می‌آید. مدل درستنمایی کامل سلسله مراتبی برای پارامتری کردن دلتا به‌صورت زیر است:

$$\begin{aligned} &Z^*_{ij}|v_{ij},y_{ij},a_j,b_j,\delta_j, \\ &\sim N(a_ju_i-b_j-\delta_jv_{ij},1-\delta_j^2)I(z^*_{ij},y_{ij}); \\ &V_{ij} \sim HN(0,1); U_i \sim N(0,1); \\ &a_j \sim \pi_1(\mu_a,\sigma_a^2); b_j \sim \pi_2(\mu_b,\sigma_b^2); \delta_j \sim \pi_3(\cdot) \end{aligned}$$

[۶] Closed Skew Normal
[۷] Half-normal Distribution

مجموعه داده‌های وزن

نتایج مطالعات مجموعه داده ۲۱۴۱ دختر نوجوان با ۱۵ سؤال گزارش شده است. که مجموعه داده‌ها مربوط به همه گیرشناسی اختلال رژیم غذایی متروپلیس لیما(Per) می‌باشد. برای تحلیل بیزی این مجموعه داده‌ها توسط مدل پروبیت و پروبیت چوله IRT پیشین‌های $a_j \sim N(۱, ۰/۵)I(a_j > ۰)$, $b_j \sim N(۰, ۲), \delta_j \sim U(-۱, ۱)$ به‌کار رفته است. برای استنباط بیزی و مقایسه‌ی مدل‌های پیشنهاد شده از ۲۲۰۰۰ تکرار ۲۰۰۰ تای اول آن را دور می‌اندازیم. مقدار DIC برای مدل پروبیت نرمال و پروبیت چوله نرمال به‌ترتیب، ۳۱۰۲۸/۸ و ۲۸۳۱۵/۵ است که نشان می‌دهد مدل پروبیت چوله IRT نسبت به مدل معرفی شده بهتر برازش شده است. از این‌رو انتظار داریم برآورد ICC مدل پروبیت نامتقارن دقت بیشتری داشته باشد.

نتایج

در این مقاله مدل IRT چوله نرمال معرفی شد و همچنین تعمیم آن به کلاس بزرگتری از خانواده توزیع‌های چوله نرمال بسته بیان گردید. رهیافت داده‌افزایی برای برآورد بیزی با روش MCMC در مدل پروبیت چوله پیشنهاد شد. به‌خاطر احتمال خودهمبستگی زیاد از داده افزایی استفاده می‌کنیم که لازمه آن افزایش تکرارها در برآورد پارامترها برای پارامتر چولگی است. مقایسه‌ی مدل‌های پروبیت نرمال و پروبیت چوله نرمال IRT با استفاده از معیار DIC انجام شد و نشان داد مدل چوله نرمال IRT مناسب‌تر است.

مراجع

[1] Albert, J. H. and Ghosh, M. (2000). *Item response modeling*. In D. Dey, S. G. and Mallick, Generalized Linear Models: A Baye. Perspe., 173 - 193 . New York: Addison Wesley.

[2] Bazan, J.L, Branco, M.D., Bolfarine, H., (2006). *A Skew Item Response Model* . Beysian Analysis, 861-892.

[3] Birnbaum, A. (1968). *Some latent trait models and their use in inferring an examinee'sability*. In Lord, F. M. and Novick, M. R. (eds.), Statis. Theori. of Ment. Tes. Scor., 395,479. Massachusetts: Addison Wesl

[4] Chen, M.-H., Dey, D. K., and Shao, Q. M. (1999) *A New Skewed Link Model for Dichotomous Quantal Response Data*. JASA, 94: 1172-1186 .

[5] Domingues-Molina,J., Gonzalez-Farias, G. and Gupta,A.(2003). *The Multivariate Closed Skew Normal Distribution*. Tech. Rep. 03-12, Department of Mathemmatics and Statistics, Bowling Green State University.

[6] Lord, F.M., (1980). Applications of Item Response Theory to Practical Testing Problems, first edition. Lawrence Erlbaum Associates, Hillsdale.

[7] Tucher,L.R.(1946) *Maximum Validity of a test with equivaled items,* Psychometrika,11: 1-13.

[8] vanderLinden, W. J. and Hambleton, R. K. (1997). *Handbook of Modern Item response Theory*. New York: Springer-Verlag.

The 45th Annual Iranian Mathematics Conference
August 26-29 2014
چهل و پنجمین
کنفرانس ریاضی ایران
۴ - ۷ شهریور ۱۳۹۳

Lindley Logarithmic Distribution: Model, Properties and Applications

E. Mahmoudi*
Academic member of School of Mathematics, Yazd University
emahmoudi@yazd.ac.ir
S. Abolhosseini
M. Sc. student of School of Mathematics, Yazd University
s-abolhoseini@yahoo.com

Abstract: In this paper we propose a new distribution with increasing and bathtub shaped failure rate, called as the Lindley logarithmic (LL) distribution. We obtain several properties of the new distribution such as its probability density function, its reliablity and failure rate functions, quantiles and moments. Rényi and Shannon entropies are presented in this paper.

Keywords: Hazard rate function, Power series, Rényi entropy, Shannon entropy, Survival function.

1 INTRODUCTION

Recently, attempts have been made to define new families of probability distributions that extend well-known families of distributions and at the same time provide great flexibility in modeling data in practice. The exponential-geometric (EG), exponential-Poisson (EP), exponential-logarithmic (EL), exponential-power series (EPS), Weibull-geometric (WG) and Weibull-power series (WPS) distributions were introduced and studied by Adamidis and Loukas [1], Kus [6], Tahmasbi and Rezaei [13], Chahkandi and Ganjali [5], Barreto-Souza et al. [3] and Morais and Barreto-Souza et al. [12], respectively.

Barreto-Souza and Cribari-Neto [2] and Louzada et al. [7] introduced the exponentiated exponential-Poisson (EEP) and the complementary exponential-geometric (CEG) distributions where the EEP is the generalization of the EP distribution and the CEG is complementary to the EG model proposed by Adamidis and Loukas [1]. Recently, Cancho et al. [4] introduced the two-parameter Poisson-exponential (PE) lifetime distribution with increasing failure rate. Mahmoudi and Jafari [8] introduced the generalized exponential-power series (GEPS) distribution by compounding the generalized exponential (GE) distribution with the power series distribution. Also exponentiated Weibull-logarithmic (EWL), exponentiated Weibull-geometric (EWG) and exponentiated Weibull-power series (EWP) distributions has been introduced and analyzed by Mahmoudi and Sepahdar [9] and Mahmoudi and Shiran [10, 11].

In this paper, we propose a new two-parameters distribution, referred to as the Lindley logarithmic (LL) distribution, which contains as special sub-models the Lindley and exponential logarithmic distributions. The main reasons for introducing the LL distribution are: (i) This distribution due to its flexibility in accommodating different forms of the risk function is an important

*Corresponding Author

The 45th Annual Iranian Mathematics Conference
August 26-29 2014
چهل و پنجمین کنفرانس ریاضی ایران
۴ - ۷ شهریور ۱۳۹۳

model that can be used in a variety of problems in modeling lifetime data. (ii) It provides a reasonable parametric fit to skewed data that cannot be properly fitted by other distributions and is a suitable model in several areas such as public health, actuarial science, biomedical studies, demography and industrial reliability.

The paper is organized as follows. In Section 2, we review the Lindley distribution and its properties. In Section 3, we define the LL distribution. The density, survival and hazard rate functions and some of their properties are given in this section. We derive quantiles and moments of the LL distribution in Section 4. Rényi and Shannon entropies are provided in Section 5.

2 LINDLEY DISTRIBUTION: A BRIEF REVIEW

The random variable X has Lindley distribution if its cumulative distribution function (cdf) takes the form

$$F_X(x) = 1 - (1 + \frac{\gamma x}{\gamma+1})e^{-\gamma x}, \quad x > 0, \qquad (1)$$

where $\gamma > 0$. The corresponding probability density function (pdf) is

$$f_X(x) = \frac{\gamma^2}{\gamma+1}(1+x)e^{-\gamma x}, \quad x > 0. \qquad (2)$$

The survival and hazard rate functions of the Lindley distribution are

$$S(x) = (1 + \frac{\gamma x}{\gamma+1})e^{-\gamma x},$$

and

$$h(x) = \frac{\frac{\gamma^2}{\gamma+1}(1+x)}{(1+\frac{\gamma x}{\gamma+1})},$$

respectively. The kth moment about zero of the Lindley distribution is given by

$$E(X^k) = \frac{\gamma^2}{\gamma+1}\left[\frac{\Gamma(k+1)}{(\gamma)^{k+1}} + \frac{\Gamma(k+2)}{(\gamma)^{k+2}}\right].$$

3 LINDLEY LOGARITHMIC DISTRIBUTION

Suppose that the random variable X has the Lindley distribution where its cdf and pdf are given in (1) and (2). Given N, let $X_1, \cdots, X_N$ be independent and identify distributed random variables from Lindley distribution. Let N is distributed according to logarithmic distribution with pdf

$$P(N=n) = \frac{-\theta^n}{n\log(1-\theta)}, \; n = 1, 2, \cdots, \; 0 < \theta < 1.$$

Let $Y = \min(X_1, \cdots, X_N)$, then the cdf of $Y|N = n$ is given by

$$F_{Y|N=n}(y) = 1 - ((1 + \frac{\gamma y}{\gamma+1})e^{-\gamma y})^n,$$

The Lindley logarithmic distribution, denote by LL(θ, γ), is defined by the marginal cdf of Y, i.e.

$$F(y) = 1 - \frac{\log(1-\theta(1+\frac{\gamma y}{\gamma+1})e^{-\gamma y})}{\log(1-\theta)}. \qquad (3)$$

The pdf of LL distribution is given by

$$f(y) = \frac{\theta\frac{\gamma^2}{\gamma+1}e^{-\gamma y}(1+y)}{(\theta(1+\frac{\gamma y}{\gamma+1})e^{-\gamma y}-1)\log(1-\theta)}, \qquad (4)$$

where $0 < \theta < 1, \gamma > 0$.
The survival and hazard rate functions of LL distribution are given, respectively, by

$$S(y) = \frac{\log(1-\theta(1+\frac{\gamma y}{\gamma+1})e^{-\gamma y})}{\log(1-\theta)}, \qquad (5)$$

and

$$h(y) = \frac{\theta\frac{\gamma^2}{\gamma+1}e^{-\gamma y}(1+y)}{(\theta(1+\frac{\gamma y}{\gamma+1})e^{-\gamma y}-1)\log(1-\theta(1+\frac{\gamma y}{\gamma+1})e^{-\gamma y})}. \qquad (6)$$

Proposition 3.1. *The limiting distribution of LL (θ, γ) where $\theta \to 0^+$ is*

$$\lim_{\theta\to 0^+} F(y) = 1 - (1 + \frac{\gamma x}{\gamma+1})e^{-\gamma x},$$

which is the cdf of Lindley distribution.

Proposition 3.2. *The limiting behavior of hazard function of LL distribution in (6) is*

$\lim_{y\to 0} h(y) = \frac{\theta\frac{\gamma^2}{\gamma+1}}{(\theta-1)\log(1-\theta)}$ *and* $\lim_{y\to\infty} h(y) = 0$

The 45th Annual Iranian Mathematics Conference
August 26-29 2014
چهل و پنجمین کنفرانس ریاضی ایران
۴ - ۷ شهریور ۱۳۹۳

Theorem 3.3. *Considering the LL distribution with the probability density function (4), we have the following properties:*

(i) As $\theta \to 0$, then $LL(\theta, \gamma)$ leads to Lindley distribution with parameter γ.

(ii) If $\gamma \geq \sqrt{1-\theta}$, $f(y)$ is decreasing in y. If $0 < \gamma < \sqrt{1-\theta}$, $f(y)$ is a unimodal function with mode at y_0, where y_0 is the solution of the equation $\gamma(y+1) + (A-1) - A\frac{\gamma(y+1)}{\gamma(y+1)+1} = 0$, with $A = \theta(1 + \frac{\gamma y}{\gamma+1})e^{-\gamma y}$.

Theorem 3.4. *considering the hazard function of the LL distribution (6),, we have the following properties:*

(i) If $(\theta-1)(\gamma-1) > (\frac{\theta\gamma^2}{\gamma+1})$ and the equation $\left[\gamma(A-1) + \theta\frac{\gamma^2}{\gamma+1}e^{-\gamma y}(1+y)(1-\gamma(1+y))\right](A-1)(1+y)$ $+ [(A-1) - \theta\frac{\gamma^2}{\gamma+1}e^{-\gamma y}(1+y)^2](1-\gamma(1+y))(A-1)$ $+ \theta\frac{\gamma^2}{\gamma+1}e^{-\gamma y}(1+y)^2[(\gamma(1+y)-2)(A-1) + (A-1) - \theta\frac{\gamma^2}{\gamma+1}e^{-\gamma y}(1+y)^2] = 0$ has no real roots, then the hazard function is increasing.

(ii) If $(\theta-1)(\gamma-1) < (\frac{\theta\gamma^2}{\gamma+1})$ and the equation $\left[\gamma(A-1) + \theta\frac{\gamma^2}{\gamma+1}e^{-\gamma y}(1+y)(1-\gamma(1+y))\right](A-1)(1+y)$ $+ [(A-1) - \theta\frac{\gamma^2}{\gamma+1}e^{-\gamma y}(1+y)^2](1-\gamma(1+y))(A-1)$ $+ \theta\frac{\gamma^2}{\gamma+1}e^{-\gamma y}(1+y)^2[(\gamma(1+y)-2)(A-1) + (A-1) - \theta\frac{\gamma^2}{\gamma+1}e^{-\gamma y}(1+y)^2] = 0$ has one real roots, then the hazard function is bathtub shaped.

4 QUANTILES AND MOMENTS OF LL DISTRIBUTION

Some of the most important features and characteristics of a distribution can be studied through its moments and quantiles such as tending, dispersion, skewness and kurtosis. Also, the quantiles of a distribution can be used in data generation from a distribution. The pth quantile of the Lindley logarithmic distribution is given by

$$x_p = \frac{\gamma+1}{\gamma}(\frac{e^{\gamma x}}{\theta}(1-(1-\theta)^{1-p})-1), \quad (7)$$

which is used for data generation from the LL distribution.

Now we obtain the moment generating function of the LL distribution. Suppose that $Y \sim LL(\theta, \gamma)$ and $X_{(1)} = \min(X_1, \cdots, X_n)$, where $X_i \sim L(\gamma)$ for $i = 1, 2, \cdots, n$, then

$$\begin{aligned} M_X(t) &= \textstyle\sum_{n=1}^{\infty} P(N=n) M_{X_{(1)}}(t) \\ &= \textstyle\sum_{n=1}^{\infty} P(N=n) \sum_{i=0}^{n-1} \binom{n-1}{i} (\frac{\gamma}{\gamma+1})^{n-i+1} n\gamma \\ &\times \left[\frac{\Gamma(n-i)}{(n\gamma-t)^{n-i}} + \frac{\Gamma(n-i+1)}{(n\gamma-t)^{n-i+1}}\right] \\ &= \frac{-\theta^n}{\log(1-\theta)} \textstyle\sum_{n=1}^{\infty} \sum_{i=0}^{n-1} \binom{n-1}{i} \frac{\gamma^{n-i+2}}{(\gamma+1)^{n-i+1}} \\ &\times \left[\frac{\Gamma(n-i)}{(n\gamma-t)^{n-i}} + \frac{\Gamma(n-i+1)}{(n\gamma-t)^{n-i+1}}\right]. \end{aligned} \quad (8)$$

One can use $M_X(t)$ to obtain the kth moment about zero of the LL distribution. We have

$$\begin{aligned} E(Y^k) &= \textstyle\sum_{n=1}^{\infty} P(N=n) E(X_{(1)}^k) \\ &= \textstyle\sum_{n=1}^{\infty} \sum_{i=0}^{n-1} \binom{n-1}{i} \frac{-\theta^n \gamma^{i+2}}{(\gamma+1)^{i+1}\log(1-\theta)} \\ &\times \left[\frac{\Gamma(k+i+2)}{(n\gamma)^{k+i+2}} + \frac{\Gamma(k+i+1)}{(n\gamma)^{k+i+1}}\right]. \end{aligned} \quad (9)$$

The mean and variance of the LL distribution are given, respectively, by

$$\begin{aligned} E(Y) &= \textstyle\sum_{n=1}^{\infty} \sum_{i=0}^{n-1} \binom{n-1}{i} \frac{-\theta^n \gamma^{i+2}}{(\gamma+1)^{i+1}\log(1-\theta)} \\ &\times \left[\frac{\Gamma(i+3)}{(n\gamma)^{i+3}} + \frac{\Gamma(i+2)}{(n\gamma)^{i+2}}\right], \end{aligned} \quad (10)$$

and

$$\begin{aligned} Var(Y) &= \textstyle\sum_{n=1}^{\infty} \sum_{i=0}^{n-1} \binom{n-1}{i} \frac{-\theta^n \gamma^{i+2}}{(\gamma+1)^{i+1}\log(1-\theta)} \\ &\times \left[\frac{\Gamma(i+4)}{(n\gamma)^{i+4}} + \frac{\Gamma(i+3)}{(n\gamma)^{i+3}}\right] - E^2(Y), \end{aligned} \quad (11)$$

where $E(Y)$ is given in Eq. (10).

5 Rényi and Shannon entropies

If X is a random variable having an absolutely continuous cumulative distribution function $F(x)$ and probability distribution function $f(x)$, then the basic uncertainty measure for distribution F (called the entropy of F) is defined as $H(x) = E[-\log(f(X))]$. Statistical entropy is a probabilistic measure of uncertainty or ignorance about the outcome of a random experiment, and is a measure of a reduction in that uncertainty. Numerous entropy and information indices, among them the Rényi entropy, have been developed and used in various disciplines and contexts. Information theoretic principles and methods have become integral parts of probability and statistics and have been applied in various branches of statistics and related fields.

Entropy has been used in various situations in science and engineering. The entropy of a random variable Y is a measure of variation of the uncertainty. For a random variable with the pdf f, the Rényi entropy is defined by $I_R(r) = \frac{1}{1-r}\log\{\int_{\mathbb{R}} f^r(y)dy\}$, for $r > 0$ and $r \neq 1$. Using the power series expansion

$$(1-z)^{-k} = \sum_{j=0}^{\infty} \frac{\Gamma(k+j)}{\Gamma(k)j!} z^j$$

gives

$$\begin{aligned}\int_0^\infty f^r(y)dy &= \left(\tfrac{\theta\gamma^2}{(\gamma+1)\log(1-\theta)}\right)^r \sum_{i=0}^r \binom{r}{i}(-1)^r \\ &\times \int_0^\infty e^{-\gamma r y} y^i (1-\theta(1+\tfrac{\gamma y}{\gamma+1})e^{-\gamma y})^{-r} dy.\end{aligned}$$

Thus, we have

$$\begin{aligned}\int_0^\infty f^r(y)dy &= \left(\tfrac{\theta\gamma^2}{(\gamma+1)\log(1-\theta)}\right)^r \sum_{i=0}^r \sum_{j=0}^\infty \sum_{k=0}^j \\ &\times \binom{r}{i}\binom{j}{k}(-1)^r \tfrac{\Gamma(r+j)}{\Gamma(r)j!}(\tfrac{\theta\gamma}{\gamma+1})^j \tfrac{\Gamma(i+j+1)}{(\gamma(j+r))^{i+j+1}}.\end{aligned}$$

According to the definition of Rényi entropy we have

$$\begin{aligned}I_R(r) &= \tfrac{1}{1-r}\log[(\tfrac{\theta\gamma^2}{(\gamma+1)\log(1-\theta)})^r \sum_{i=0}^r \sum_{j=0}^\infty \sum_{k=0}^j \\ &\times \binom{r}{i}\binom{j}{k}(-1)^r \tfrac{\Gamma(r+j)}{\Gamma(r)j!}(\tfrac{\theta\gamma}{\gamma+1})^j \tfrac{\Gamma(i+j+1)}{(\gamma(j+r))^{i+j+1}}].\end{aligned}$$

The Shannon entropy is defined by $E[-\log[f(Y)]]$. This is a special case derived from $\lim_{r\to 1} I_R(r)$.

References

[1] K. Adamidis, and S. Loukas, "A lifetime distribution with decreasing failure rate", Statistics and Probability Letters, vol. 39, pp. 35-42, 1998.

[2] W. Barreto-Souza, and F. Cribari-Neto, "A generalization of the exponential-Poisson distribution", Statistics and Probability Letters, vol. 79, pp. 2493-2500, 2009.

[3] W. Barreto-Souza, A. L. Morais, and G. M. Cordeiro, "The Weibull-geometric distribution", Journal of Statistical Computation and Simulation, vol. 81, pp. 645-657, 2011.

[4] V. G. Cancho, F. Louzada-Neto, and G. D. C. Barriga, "The Poisson-exponential lifetime distribution", Computational Statistics and Data Analysis, vol. 55, pp. 677-686, 2011.

[5] M. Chahkandi, and M. Ganjali, "On some lifetime distributions with decreasing failure rate", Computational Statistics and Data Analysis, vol. 53, pp. 4433-4440, 2009.

[6] C. Kus, "A new lifetime distribution", Computational Statistics and Data Analysis, vol. 51, pp. 4497-4509, 2007.

[7] F. Louzada, M. Roman, and V. G. Cancho, "The complementary exponential geometric distribution: model, properties, and comparison with its counter part", Computational Statistics and Data Analysis, vol. 55, pp. 2516-2524, 2011.

[8] E. Mahmoudi, and A. A. Jafari, "Generalized exponential-power series distributions", Computational Statistics and Data Analysis, vol. 56, pp. 4047-4066, 2012.

[9] E. Mahmoudi and A. Sepahdar, "Exponentiated Weibull-logarithmic distribution", submitted for publication.

[10] E. Mahmoudi, and M. Shiran, "Exponentiated Weibull-geometric distribution and its applications", submitted for publication, http://arxiv.org/abs/1206.4008v1.

[11] E. Mahmoudi, and M. Shiran, "Exponentiated Weibull-power series distribution", submitted for publication.

[12] A. L. Morais, and W. Barreto-Souza, "A compound class of Weibull and power series distributions", Computational Statistics and Data Analysis, vol. 55, pp. 1410-1425, 2011.

[13] R. Tahmasbi and S. Rezaei, "A two-parameter lifetime distribution with decreasing failure rate", Computational Statistics and Data Analysis, vol. 52, pp. 3889-3901, 2008.

The 45th Annual Iranian Mathematics Conference
August 26-29 2014
چهل و پنجمین کنفرانس ریاضی ایران
۴ - ۷ شهریور ۱۳۹۳

The Marshall-Olkin Extended Rayleigh Distribution

S.M.T.K. MirMostafaee, Academic member of Faculty of Mathematical Sciences, University of Mazandaran, m.mirmostafaee@umz.ac.ir

M. Naghizadeh Qomi, Academic member of Faculty of Mathematical Sciences, University of Mazandaran, m.naghizadeh@umz.ac.ir

Sara Ziari*, M. Sc. student of Faculty of Mathematical Sciences, University of Mazandaran, ziarisara@yahoo.com

Abstract: In this paper, we propose a new extension of the Rayleigh distribution. Several mathematical properties of the new model, called the Marshall-Olkin extended Rayleigh distribution, are derived. We also discuss the estimation of the model parameters by maximum likelihood and obtain the observed information matrix. We provide an application to real data which illustrates the usefulness of the model.
Keywords: Marshall-Olkin's scheme, Maximum likelihood estimation, Rayleigh distribution.

1 INTRODUCTION

Reference [1] introduced a new family of distributions, called the Marshall-Olkin extended (MOE) family, by adding a new shape parameter to the baseline distribution. Let $G(x)$ be a cumulative distribution function (cdf) of a continuous random variable X. Suppose that X has survival function (SF) $\bar{G}(x) = 1 - G(x)$. The new survival function takes the form

$$\bar{F}(x) = \frac{\alpha \bar{G}(x)}{1 - \bar{\alpha}\bar{G}(x)}, \qquad -\infty < x < \infty. \quad (1)$$

This procedure of generalization was called Marshall-Olkin's (MO) scheme later by statisticians. Note that for $\alpha = 1$, $F(x) = G(x)$ and therefore $G(x)$ is a basic exemplar of (1). Reference [1] studied two special cases of (1) by considering $G(x)$ to be the exponential and Weibull distributions, which are called MOE exponential and MOE Weibull distributions, respectively. Since then many authors introduced new extended distributions using the MO scheme. Examples include the MOE Lomax distribution [2], the MOE normal distribution [3] and the MOE Fréchet distribution [4] among others.

In this paper, we apply the MO scheme to generalize the Rayleigh distribution. The Rayleigh distribution is widely applied in several areas of statistics, partly because of its linear and increasing failure rate, which makes it an appropriate distribution for modeling the lifetime distribution of components which age rapidly with time. This distribution is a special case of the two-parameter Weibull distribution with the shape parameter equal to 2. The SF of the Rayleigh distribution is given by

$$\bar{G}_R(x) = e^{-\lambda x^2}, \quad x > 0, \lambda > 0.$$

Taking $\bar{G}(x)$ in (1) to be the SF of the Rayleigh

*Corresponding Author

distribution, we obtain

$$\bar{F}(x) = \frac{\alpha e^{-\lambda x^2}}{1-\bar{\alpha} e^{-\lambda x^2}}, \qquad x > 0. \qquad (2)$$

Then the corresponding probability density function (pdf) of this new model is given by

$$f(x) = \frac{2\alpha\lambda x e^{-\lambda x^2}}{\left(1-\bar{\alpha} e^{-\lambda x^2}\right)^2}, \quad x > 0, \quad \alpha, \lambda > 0. \quad (3)$$

We shall refer to the new model with pdf given in (3) as the Marshall-Olkin extended Rayleigh (MOER) distribution and we denote a random variable X with this density function (3) by $X \sim \text{MOER}(\alpha, \lambda)$. It may be noted that when $\alpha = 1$, (3) reduces to the pdf of the Rayleigh distribution. Plots of MOER densities for some selected values of α when $\lambda = 1$ are shown in Fig. 1. We derive some mathematical properties of the MOER distribution in Section 2. Data application is discussed in Section 3.

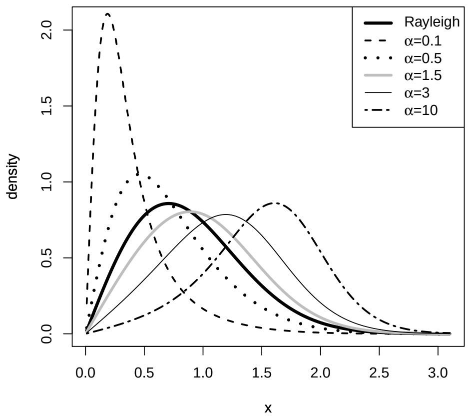

Figure 2: MOER pdfs for selected values of α when $\lambda = 1$.

2 GENERAL PROPERTIES

In this section, we discuss some general properties as well as maximum likelihood estimation of the parameters of the MOER distribution. The cdf of the MOER distribution can be written as

$$F(x) = \frac{1-e^{-\lambda x^2}}{1-\bar{\alpha} e^{-\lambda x^2}}, \qquad x > 0.$$

Inverting the cdf function, we readily obtain the quantile function as

$$x = Q(u) = \sqrt{\frac{1}{2}\log\left(\frac{1-\bar{\alpha}u}{1-u}\right)}. \qquad (4)$$

Simulation of an MOER random variable follows directly from (4), i.e. if U is a simulated random variable from uniform distribution on $(0,1)$, then $X = Q(U) \sim \text{MOER}(\alpha, \lambda)$.

Using (2) and (3), the hazard rate function (hrf) of the MOER model is found to be

$$r(x) = \frac{2\lambda x}{1-\bar{\alpha} e^{-\lambda x^2}}, \qquad x > 0. \qquad (5)$$

It is clear that $r(0) = 0$ and $\lim_{x\to\infty} r(x) = \infty$. The first derivative of $r(x)$ with respect to (w.r.t) x is

$$r'(x) = \frac{2\lambda\eta(\lambda x^2)}{\left(1-\bar{\alpha} e^{-\lambda x^2}\right)^2},$$

where $\eta(y) = 1 - \bar{\alpha}(1+2y)e^{-y}$. Take $y = \lambda x^2$, if $\alpha > 1$, then $\eta(y) > 0$ and hence $r(x)$ is increasing w.r.t. x. For $\alpha \leq 1$, consider the first derivative of $\eta(y)$ w.r.t. y which equals $\eta'(y) = \bar{\alpha}(2y-1)e^{-y}$ implying that $\eta(y)$ has a unique critical point $y_* = 1/2$. Since $\eta''(1/2) = \bar{\alpha}e^{-1/2} \geq 0$, $\eta(y)$ has an absolute minimum at y_*. The absolute minimum value of $\eta(y)$ is $\eta(1/2) = 1 - 2\bar{\alpha}e^{-1/2}$. Note that $\eta(0) = \alpha > 0$ and $\lim_{y\to\infty} \eta(y) = 1 > 0$. If $\eta(y_*) \geq 0$ or equivalently $\alpha \geq 1 - \sqrt{e}/2$, then $\eta(y) \geq 0$ for all $y \geq 0$ and hence $r'(x) \geq 0$ for all $x > 0$. If $\eta(y_*) < 0$, then $\eta(y)$ has exactly two roots $x_1 = \sqrt{y_1/\lambda} < x_2 = \sqrt{y_2/\lambda}$. Since $r(0) = 0$ and $\lim_{x\to\infty} r(x) = \infty$ and $r(x) > 0$, then x_1 (x_2) must be a point of local maximum (minimum) for $r(x)$. Summing up, if $\alpha \geq 1 - \sqrt{e}/2$, then $r(x)$ is increasing w.r.t. x and if $0 < \alpha < 1 - \sqrt{e}/2 \simeq 0.17564$, then $r(x)$ is increasing-decreasing-increasing w.r.t. x. Plots of MOER hrfs for some selected values of α when $\lambda = 1$ are shown in Fig. 2.

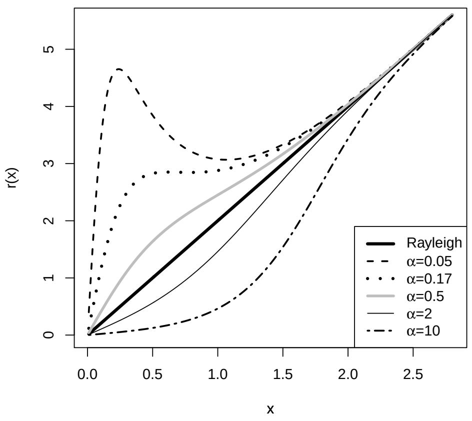

Figure 2: MOER hrfs for selected values of α when $\lambda = 1$.

Next, we provide useful expansions for the pdf and cdf of the MOER model. Consider the following well-known series expansion

$$(1-z)^{-s} = \sum_{j=0}^{\infty} \binom{s+j-1}{s-1} z^j, \qquad (6)$$

for $|z| < 1$ and $s > 0$. Since $1-\alpha < 1$, we consider two cases: if $|1-\alpha| < 1$, or equivalently $0 < \alpha < 2$, then from (3) and (6) we can write

$$f(x) = 2\alpha\lambda x \sum_{j=0}^{\infty} (j+1)\bar{\alpha}^j e^{-(j+1)\lambda x^2},$$

and the cdf can be expanded as

$$F(x) = (1-e^{-\lambda x^2}) \sum_{j=0}^{\infty} \bar{\alpha}^j e^{-j\lambda x^2}.$$

Now, if $1-\alpha < 1/2$ or $\alpha > 1/2$, then $|\bar{\alpha}|/\alpha < 1$ and we use the following equality

$$1-\bar{\alpha}e^{-\lambda x^2} = \alpha\left[1+\frac{\bar{\alpha}}{\alpha}\left(1-e^{-\lambda x^2}\right)\right].$$

Therefore, from (6) and the binomial expansion, we can write

$$\begin{aligned} f(x) &= 2\lambda x e^{-\lambda x^2} \sum_{j=0}^{\infty} \frac{(j+1)(-\bar{\alpha})^j}{\alpha^{j+1}} \left(1-e^{-j\lambda x^2}\right)^j \\ &= 2\lambda x \sum_{j=0}^{\infty}\sum_{i=0}^{j} \binom{j}{i} \frac{(j+1)(-1)^{i+j}\bar{\alpha}^j}{\alpha^{j+1}} e^{-(i+1)\lambda x^2}, \end{aligned}$$

and

$$F(x) = \sum_{j=0}^{\infty}\sum_{i=0}^{j+1} \binom{j+1}{i} \frac{(-1)^{i+j}\bar{\alpha}^j}{\alpha^{j+1}} e^{-i\lambda x^2}.$$

Using the series expansions for the MOER pdf, we can find the k-th ($k > 0$) moment of this new distributions in two cases: if $0 < \alpha < 2$, then

$$E(X^k) = \frac{\alpha}{\lambda^{k/2}} \sum_{j=0}^{\infty} \frac{\bar{\alpha}^j}{(j+1)^{k/2}} \Gamma(k/2+1),$$

where $\Gamma(\cdot)$ is the complete gamma function. If $\alpha > 1/2$, then the k-th moment can be written as

$$E(X^k) = \frac{1}{\lambda^{k/2}} \sum_{j=0}^{\infty}\sum_{i=0}^{j} \binom{j}{i} \frac{(j+1)(-1)^{i+j}\bar{\alpha}^j}{\alpha^{j+1}(i+1)^{k/2+1}} \Gamma(k/2+1).$$

Now, we consider the estimation of the parameters of the MOER distribution using maximum likelihood method. Suppose that $x_1, \ldots, x_n$ are an observed random sample of size n from MOER distribution with unknown parameters α and λ, then the log-likelihood function for (α, λ) becomes

$$\begin{aligned} \ell &= \ell(\alpha,\lambda) = n\log 2\alpha\lambda + \sum_{i=1}^{n}\log x_i - \lambda\sum_{i=1}^{n} x_i^2 \\ &\quad -2\sum_{i=1}^{n}\log\left(1-\bar{\alpha}e^{-\lambda x_i^2}\right). \qquad (7) \end{aligned}$$

The maximum likelihood estimates of the unknown parameters are obtained by maximizing the log-likelihood function $\ell(\alpha,\lambda)$ with respect to (α,λ). The maximum likelihood estimators (MLEs) of α and λ say $\widehat{\alpha}$ and $\widehat{\lambda}$, respectively, can be obtained as the solutions of the following non-linear equations

$$\begin{aligned} \frac{\partial\ell}{\partial\alpha} &= \frac{n}{\alpha} - 2\sum_{i=1}^{n} \frac{e^{-\lambda x_i^2}}{1-\bar{\alpha}e^{-\lambda x_i^2}} = 0, \\ \frac{\partial\ell}{\partial\lambda} &= \frac{n}{\lambda} - \sum_{i=1}^{n} x_i^2 - 2\bar{\alpha}\sum_{i=1}^{n} \frac{x_i^2 e^{-\lambda x_i^2}}{1-\bar{\alpha}e^{-\lambda x_i^2}} = 0. \end{aligned}$$

These equations can not be solved analytically and statistical software can be used to solve them numerically using iterative methods such as the NewtonRaphson type algorithms. For interval estimation and hypothesis testing of α and λ, we may use the asymptotic joint distribution of $(\widehat{\alpha}, \widehat{\lambda})$. Under certain regularity conditions that are stated in [5], pages 461-463, that are fulfilled for the parameters in the interior of the parameter space, we have that $\sqrt{n}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \overset{a}{\sim} N_2\left(\mathbf{0}, \mathbf{I}_{\boldsymbol{\theta}}^{-1}\right)$, where $\overset{a}{\sim}$ means approximately distributed, $\boldsymbol{\theta} = (\alpha,\lambda)^T$ is the vector of unknown parameters, $\widehat{\boldsymbol{\theta}} = (\widehat{\alpha}, \widehat{\lambda})^T$, $\mathbf{I}_{\boldsymbol{\theta}}$ is the expected information matrix and $\mathbf{I}_{\boldsymbol{\theta}}^{-1}$ is the inverse of $\mathbf{I}_{\boldsymbol{\theta}}$. The asymptotic behavior remains valid if $\mathbf{I}_{\boldsymbol{\theta}}$ is replaced by $\frac{1}{n} J_n(\boldsymbol{\theta})$ where $J_n(\boldsymbol{\theta})$ is the observed information matrix defined as

$$J_n(\boldsymbol{\theta}) = -\begin{bmatrix} J_{\alpha\alpha} & J_{\alpha\lambda} \\ J_{\lambda\alpha} & J_{\lambda\lambda} \end{bmatrix} = -\begin{bmatrix} \frac{\partial^2}{\partial\alpha^2}\ell(\alpha,\lambda) & \frac{\partial^2}{\partial\alpha\partial\lambda}\ell(\alpha,\lambda) \\ \frac{\partial^2}{\partial\lambda\partial\alpha}\ell(\alpha,\lambda) & \frac{\partial^2}{\partial\lambda^2}\ell(\alpha,\lambda) \end{bmatrix}$$

The elements of $J_n(\boldsymbol{\theta})$ can be obtained as

$$\begin{aligned} J_{\alpha\alpha} &= -\frac{n}{\alpha^2} + 2\sum_{i=1}^{n} \frac{e^{-2\lambda x_i^2}}{\left(1-\bar{\alpha}e^{-\lambda x_i^2}\right)^2}, \\ J_{\lambda\lambda} &= -\frac{n}{\lambda^2} + 2\bar{\alpha}\sum_{i=1}^{n} \frac{x_i^4 e^{-\lambda x_i^2}}{\left(1-\bar{\alpha}e^{-\lambda x_i^2}\right)^2}, \\ J_{\alpha\lambda} &= J_{\lambda\alpha} = 2\sum_{i=1}^{n} \frac{x_i^2 e^{-\lambda x_i^2}}{\left(1-\bar{\alpha}e^{-\lambda x_i^2}\right)^2}. \end{aligned}$$

Simply we have

$$\begin{aligned} J_n^{-1}(\boldsymbol{\theta}) &= \frac{-1}{J_{\alpha\alpha}J_{\lambda\lambda} - J_{\alpha\lambda}^2} \begin{bmatrix} J_{\lambda\lambda} & -J_{\alpha\lambda} \\ -J_{\alpha\lambda} & J_{\alpha\alpha} \end{bmatrix} \\ &= \begin{bmatrix} \widehat{Var}(\widehat{\alpha}) & \widehat{Cov}(\widehat{\alpha},\widehat{\lambda}) \\ \widehat{Cov}(\widehat{\alpha},\widehat{\lambda}) & \widehat{Var}(\widehat{\lambda}) \end{bmatrix} \end{aligned}$$

The unknown parameters in the elements of $J_n^{-1}(\boldsymbol{\theta})$ can be replaced by their corresponding MLEs. The asymptotic equal tailed $100(1-p)$ percent confidence intervals for the parameters α and λ are $\widehat{\alpha} \pm z_{p/2}\sqrt{\widehat{Var}(\widehat{\alpha})}$ and $\widehat{\lambda} \pm z_{p/2}\sqrt{\widehat{Var}(\widehat{\lambda})}$ respectively, where z_a denotes the $100a$ percentile of the standard normal random variable.

3 DATA APPLICATION

In this section, we analyze a real data set and compare the MOER distribution with Rayleigh distribution. The data set is taken from [6] p. 105 and shows the relief times of 20 patients receiving an analgesic in hours. These data are as follows

1.1, 1.4, 1.3, 1.7, 1.9, 1.8, 1.6, 2.2, 1.7, 2.7,
4.1, 1.8, 1.5, 1.2, 1.4, 3, 1.7, 2.3, 1.6, 2.

We have fitted both Rayleigh and MOER distributions to the above data using the Kolmogorov-Smirnov (K-S) test. The values of K-S statistics and their corresponding p-values are reported in Table 1. In addition, the Akaike information criterion (AIC) and the Bayesian information criterion (BIC) are utilized to compare the candidate models. These measure are also presented in Table 1. One can easily conclude from Table 1 that MOER distribution fits the data better than the Rayleigh distribution.

TABLE 1
MLEs, AIC, BIC, K-S statistics and p-value for the data set.

Model	MLEs	AIC	BIC	K-S	p-value
Rayleigh	$\widehat{\lambda} = 0.245$	46.958	47.953	0.257	0.1436
MOER	$\widehat{\lambda} = 0.463$ $\widehat{\alpha} = 3.643$	15.393	17.385	0.171	0.6025

References

[1] A. W. Marshall and I. Olkin, "A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families," Biometrika, vol. 84, pp. 641-652, 1997.

[2] M. E. Ghitany, F. A. Al-Awadhi, and L. A. Alkhalfan, "Marshall-Olkin extended Lomax distribution and its application to censored data," Commun. Stat. Theory Meth., vol. 36, pp. 1855-1866, 2007.

[3] V. J. Garcia, E. Gómez-Déniz, and F. J. Vázquez-Polo, "A new skew generalization of the normal distribution: Properties and applications," Comput. Stat. Data Anal., vol. 54, pp. 2021-2034, 2010.

[4] E. Krishna, K. K. Jose, T. Alice, and M. M. Ristić, "The Marshall-Olkin Fréchet distribution," Commun. Stat. Theory Meth., vol. 42, pp. 4091-4107, 2013.

[5] L. E. Lehmann and G. Casella, Theory of Point Estimation, 2nd ed., Springer, New York, 1998.

[6] A. J. Gross and V. A. Clark, Survival Distributions: Reliability Applications in the Biomedical Sciences, Willey, New York, 1975.

The 45th Annual Iranian Mathematics Conference
August 26-29 2014
چهل و پنجمین کنفرانس ریاضی ایران
۴ - ۷ شهریور ۱۳۹۳

Stochastic Ordering of Medians in Equi-Correlated Trivariate Normal Vectors Based on the Correlation Coefficient

M. Madadi*, Academic member Mathematics and Computer Shahid Bahonar University, madadi@uk.ac.ir

A. Jamalizadeh, Academic member Mathematics and Computer Shahid Bahonar University, ajamalizadeh@uk.ac.ir

Abstract: In this paper, we establish the stochastic ordering of median from an equi-correlated trivariate normal vector based on the strength of the correlation coefficient. Specifically, by considering two equi-correlated trivariate normal vectors with different correlation coefficients, we show that the absolute value of the median in the vector with smaller correlation coefficient is stochastically smaller than the absolute value of the median in the vector with larger correlation coefficient. We prove this result by utilizing skew-normal distributions.

Keywords: Stochastic ordering, equi-correlated trivariate normal distribution, order statistics, median, skew-normal distribution, correlation coefficient

1 INTRODUCTION

Let $\mathbf{X} = (X_1, \cdots, X_n)^T$ and $\mathbf{Y} = (Y_1, \cdots, Y_n)^T$ be two n-dimensional equi-correlated normal random vectors as

$$\begin{aligned} \mathbf{X} &\sim N_n\left(\mathbf{0}_n, (1-\rho_X)\mathbf{I}_n + \rho_X \mathbf{1}\mathbf{1}^T\right), \\ \mathbf{Y} &\sim N_n\left(\mathbf{0}_n, (1-\rho_Y)\mathbf{I}_n + \rho_Y \mathbf{1}\mathbf{1}^T\right), \qquad (1) \\ &-\frac{1}{n-1} < \rho_X, \rho_Y < 1, \end{aligned}$$

where $\mathbf{0}_n$ and $\mathbf{1}_n$ denote the vectors of zeros and ones of dimension n, respectively, and $\mathbf{I}_n$ denotes the identity matrix of dimension n. Let the cumultative distribution functions (CDFs) of $\mathbf{X}$ and $\mathbf{Y}$ be denoted by $F_{\mathbf{X}}(\mathbf{x}; \rho_X) = P(X_1 \le x_1, \cdots, X_n \le x_n)$ and $G_{\mathbf{Y}}(\mathbf{x}; \rho_Y) = P(Y_1 \le x_1, \cdots, Y_n \le x_n)$, respectively, for $\mathbf{x} = (x_1, \cdots, x_n)^T \in \mathbb{R}^n$. Further, let $X_{1:n} \le \cdots \le X_{n:n}$ and $Y_{1:n} \le \cdots \le Y_{n:n}$ denote the order statistics arising by arranging in ascending order the components of $\mathbf{X}$ and $\mathbf{Y}$, respectively, and $F_{r:n}(\cdot; \rho_X)$ and $G_{r:n}(\cdot; \rho_Y)$ denote the CDFs of $X_{r:n}$ and $Y_{r:n}$, respectively.

A random variable X is said to be smaller than Y in the usual stochastic order, denoted by $X \le_{\text{st}} Y$, if and only if $P(X > x) \le P(Y > x)$ (or $F_X(x) \ge F_Y(x)$), for all $x \in \mathbb{R}$ [?]. By using the result in Theorem 5.1 of Das Gupta et al. [?], if $\rho_X < \rho_Y$, then $F_{\mathbf{X}}(\mathbf{x}; \rho_X) \le G_{\mathbf{Y}}(\mathbf{x}; \rho_Y)$, for $\mathbf{x} \in \mathbb{R}^n$. If we set $\mathbf{x} = x\mathbf{1}_n$, then if $\rho_X < \rho_Y$, we

*Corresponding Author

deduce

$$\begin{aligned} F_{n:n}(x;\rho_X) &= F_{\mathbf{X}}(x\mathbf{1}_n;\rho_X) \leq G_{\mathbf{Y}}(x\mathbf{1}_n;\rho_Y) \\ &= G_{n:n}(x;\rho_Y), \quad x \in \mathbb{R}, \end{aligned}$$

which shows

$$Y_{n:n} \leq_{\text{st}} X_{n:n}. \tag{2}$$

Since, $X_{1:n} \stackrel{d}{=} -X_{n:n}$ and $Y_{1:n} \stackrel{d}{=} -Y_{n:n}$, if $\rho_X < \rho_Y$, we also readily obtain

$$X_{1:n} \leq_{\text{st}} Y_{1:n}. \tag{3}$$

A natural question is about the stochastic ordering of other order statistics. In this regard, we derive some results by considering the trivariate case. When $n = 3$, from (**??**) and (**??**), we have, for $-\frac{1}{2} < \rho_X < \rho_Y < 1$,

$$Y_{3:3} \leq_{\text{st}} X_{3:3} \quad \text{and} \quad X_{1:3} \leq_{\text{st}} Y_{1:3}. \tag{4}$$

In this note, we establish that, for $-\frac{1}{2} < \rho_X < \rho_Y < 1$,

$$|X_{2:3}| \leq_{\text{st}} |Y_{2:3}|. \tag{5}$$

A random variable Z_λ is said to have the skew-normal distribution with shape parameter $\lambda \in \mathbb{R}$, denoted by $Z_\lambda \sim SN(\lambda)$, if its probability density function (PDF) is [**?**],

$$f_{SN}(z;\lambda) = 2\phi(z)\,\Phi(\lambda z), \quad z \in \mathbb{R}, \tag{6}$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ denote the PDF and CDF of the standard normal distribution, respectively. By using Proposition 6 in [**?**], if $\lambda_1 < \lambda_2$, then

$$Z_{\lambda_1} \leq_{\text{st}} Z_{\lambda_2}. \tag{7}$$

Also, it is known that [**?**],

$$X_{2:2} \sim SN\left(\sqrt{\frac{1-\rho_X}{1+\rho_X}}\right), Y_{2:2} \sim SN\left(\sqrt{\frac{1-\rho_Y}{1+\rho_Y}}\right). \tag{8}$$

Since $\sqrt{\frac{1-\rho_Y}{1+\rho_Y}} < \sqrt{\frac{1-\rho_X}{1+\rho_X}}$ for $\rho_X < \rho_Y$, the stochastic ordering in (**??**) yields the results in (**??**) and (**??**) for $n = 2$.

In this paper, by using normal skew-normal and generalized normal distributions, we establish the result in (**??**). in Section 2, we discuss some stochastic ordering results for normal skew-normal and generalized normal distributions. Then, by using these results, we establish the main result in (**??**) in Section 3 and discussed some implications briefly.

2 Stochastic orderings for normal skew-normal (NSN) and generalized normal (GN) distributions

Azzalini and Regoli [**?**] discussed stochastic ordering a general class of skewed distributions. In this section, we discuss briefly some stochastic orderings for two distributions that we will use in the sequel. Recently, Gomez et al. [**?**] proposed a normal skew-normal (NSN) distribution. A random variable is said to have a NSN distribution with parameters $\alpha, \beta \in \mathbb{R}$ if its PDF is

$$c(\alpha,\beta)\,\phi(z)\,F_{SN}(\alpha z;\beta), \quad z \in \mathbb{R},$$

where $c(\alpha,\beta) = \left\{\frac{1}{2} - \frac{1}{\pi}\arctan\left(\frac{\beta}{\sqrt{1+\alpha^2(1+\beta^2)}}\right)\right\}^{-1}$ and $F_{SN}(\cdot;\beta)$ denotes the CDF of $SN(\beta)$. This NSN distribution is, in fact, a special case of a skew-normal distribution discussed by Loperfido et al. [**?**]. Here, we consider a special case of this distribution when $\alpha = \lambda$ and $\beta = \sqrt{\frac{1+\lambda^2}{3-\lambda^2}}$, for $|\lambda| < 3$. If a random variable Z^*_λ has this specific NSN distribution, then we denote it by $Z^*_\lambda \sim NSN(\lambda)$, and its PDF becomes

$$f_{NSN}(z;\lambda) = 3\phi(z)\,F_{SN}\left(\lambda z; \sqrt{\frac{1+\lambda^2}{3-\lambda^2}}\right), \quad z \in \mathbb{R}. \tag{9}$$

We denote the corresponding CDF by $F_{NSN}(\cdot;\lambda)$. Many of the properties of the NSN distribution in (**??**) are similar to those of the SN distribution in (**??**). For example, it reduces to the standard normal distribution when $\lambda = 0$ and it is strongly unimodal for all values of $\lambda \in \mathbb{R}$ and Loperfido et al.

[?]. Moreover, just like the SN distribution, the NSN distribution in (??) is stochastically ordered with respect to λ. This property can not be obtained by the result in Proposition 6 of [?]. To prove this property, we need the following lemma. In the lemma, we assume that $a(\lambda)$ is a real-valued function with domain $D_a \subseteq \mathbb{R}$, and also that the first derivative of $a(\lambda)$ exists for $\lambda \in D_a$, and is denoted by $a'(\lambda)$.

Lemma 1 For $z \in \mathbb{R}$ and $\lambda \in D_a$, if $\left(\frac{\lambda a(\lambda)}{1+\lambda^2} - \frac{a'(\lambda)}{1+a^2(\lambda)}\right) \leq 0$, then

$$\frac{\partial}{\partial \lambda} \int_{-\infty}^{z} \phi(t) F_{SN}(\lambda t; a(\lambda))\, dt < 0.$$

By using the result in Lemma 1, we can easily prove that the NSN distribution in (??) is stochastically ordered with respect to λ, as done in the following Corollary.

Corollary 1 If $Z^*_{\lambda_1} \sim NSN(\lambda_1)$ and $Z^*_{\lambda_2} \sim NSN(\lambda_2)$, and $-\sqrt{3} < \lambda_1 < \lambda_2 < \sqrt{3}$, then $Z^*_{\lambda_1} \leq_{\text{st}} Z^*_{\lambda_2}$. A random variable U_λ is said to have a GN distribution with parameter $-\sqrt{3} < \lambda < \sqrt{3}$, denoted by $U_\lambda \sim GN(\lambda)$, if its PDF is

$$f_{GN}(u;\lambda) = 3f_{SN}(u;\lambda) - 2f_{NSN}(u;\lambda), \quad u \in \mathbb{R}. \tag{10}$$

To see this is a valid density, we first note that $f_{GN}(u;\lambda) = 6\phi(u)\left(\Phi(\lambda u) - F_{SN}\left(\lambda u; \sqrt{\frac{1+\lambda^2}{3-\lambda^2}}\right)\right)$. Because the CDF of the SN distribution in (??) is decreasing with respect to λ and $\sqrt{\frac{1+\lambda^2}{3-\lambda^2}} > 0$, we have $F_{SN}\left(\lambda u; \sqrt{\frac{1+\lambda^2}{3-\lambda^2}}\right) \leq F_{SN}(\lambda u; 0) = \Phi(\lambda u)$, which implies that $f_{GN}(u;\lambda) \geq 0$. Next, it is clear that $\int_{-\infty}^{+\infty} f_{GN}(u;\lambda) = 1$ which means $f_{GN}(u;\lambda)$ is a valid PDF. If $\lambda = 0$, then it reduces to the PDF of the standard normal distribution. The CDF of the GN density in (??) is clearly

$$F_{GN}(u;\lambda) = 3F_{SN}(u;\lambda) - 2F_{NSN}(u;\lambda), \quad u \in \mathbb{R}. \tag{11}$$

So the GN distribution in (??) is symmetric about 0. In the following lemma, we prove the stochastic ordering of $|U_\lambda|$, which will be used to establish the main result of this paper in the next section.

Lemma 2 For $0 \leq \lambda_1 < \lambda_2 < \sqrt{3}$, if $U_{\lambda_1} \sim GN(\lambda_1)$ and $U_{\lambda_2} \sim GN(\lambda_2)$, then

$$|U_{\lambda_2}| \leq_{\text{st}} |U_{\lambda_1}|. \tag{12}$$

3 Stochastic ordering of medians in two equi-correlated trivariate normal vectors

In this section, we prove the main result in (??) and then mention an extension of this result. In the following lemma, we express the distributions of order statistics in terms of NSN and GN distributions. Although all these representations are identical, there are two advantages in the last representation. First, both NSN and GN distributions have just one parameter, and in Loperfido et al. [?] and Jamalizadeh and Balakrishnan [?], we deal with two and three parameters. The representation given here enables us to study several properties of order statistics easily. The second advantage is that in trivariate case, the result is very similar to that of the bivariate case (see Eq. (??) and Lemma 3).

Lemma 3 We have

$$X_{1:3} \sim NSN\left(-\sqrt{\frac{1-\rho_X}{1+\rho_X}}\right),$$

$$X_{2:3} \sim GN\left(\sqrt{\frac{1-\rho_X}{1+\rho_X}}\right),$$

$$X_{3:3} \sim NSN\left(\sqrt{\frac{1-\rho_X}{1+\rho_X}}\right).$$

Corollary 2 For $-\frac{1}{2} < \rho_X < \rho_Y < 1$, we have

$$|X_{2:3}| \leq_{\text{st}} |Y_{2:3}|.$$

The stochastic ordering of medians in Lemma 3 translates immediately into a set of impli-

The 45th Annual Iranian Mathematics Conference
August 26-29 2014
چهل و پنجمین کنفرانس ریاضی ایران
۴ - ۷ شهریور ۱۳۹۳

cations about the ordering of moments and quantiles of these medians. Specifically, for $-\frac{1}{2} < \rho_X < \rho_Y < 1$, we have the following:

(i) If $Q_{|X_{2:3}|}(p)$ and $Q_{|Y_{2:3}|}(p)$ denote the pth quantiles of $|X_{2:3}|$ and $|Y_{2:3}|$ for any $0 < p < 1$, then

$$Q_{|X_{2:3}|}(p) \leq Q_{|Y_{2:3}|}(p)\,; \tag{13}$$

(ii) For any non-decreasing function $t(\cdot)$ for which the involved expectations exist, we have

$$E\left(t\left(|X_{2:3}|\right)\right) \leq E\left(t\left(|Y_{2:3}|\right)\right). \tag{14}$$

Moreover, from proof of lemma 2, both inequalities in (**??**) and (**??**) are also strict.

The result in Corollary can be extended to scale mixture of normal distributions. More specifically, n-dimensional random vectors $\mathbf{X}^* = (X_1^*, \cdots, X_n^*)^T$ and $\mathbf{Y}^* = (Y_1^*, \cdots, Y_n^*)^T$are said to be scale mixures of vectors of $\mathbf{X}$ and $\mathbf{Y}$ in (**??**) if there is a positive random vaiable W, with CDF H, such that

$$\mathbf{X}^* \stackrel{d}{=} \sqrt{W}\mathbf{X} \quad \text{and} \quad \mathbf{Y}^* \stackrel{d}{=} \sqrt{W}\mathbf{Y},$$

where $\stackrel{d}{=}$ means equality in distribution. Some prominent multivariate distributions, such as Student-t snd Laplace, are scale mixtures of normal distribution. If $X_{r:n}^*$ and $Y_{r:n}^*$ denote the r-th order statistics from $\mathbf{X}^*$ and $\mathbf{Y}^*$, then from (**??**) and (**??**), we have $Y_{n:n}^* \leq_{\text{st}} X_{n:n}^*$ and $X_{1:n}^* \leq_{\text{st}} Y_{1:n}^*$, for $-\frac{1}{n-1} < \rho_X < \rho_Y < 1$. Although this latter result can also be obtained directly from Theorem 5.1 of [?], in the trivariate case, by Corollary 2, we also have, for $-\frac{1}{2} < \rho_X < \rho_Y < 1$,

$$|X_{2:3}^*| \leq_{\text{st}} |Y_{2:3}^*|$$

in this general case.

References

[1] A. Azzalini, "A class of distributions which includes the normal ones," Scandinavian Journal of Statistics, vol. 12, pp. 171-178, 1985.

[2] A. Azzalini, G. Regoli, "Some properties of skew-symmetric distributions," Annals of the Institute of Statistical Mathematics, vol. 64,pp. 857-879, 2012.

[3] S. Das Gupta, M. L. Eaton, I. Olkin, M. Perlman, L. J. Savage, M. Sobel, "Inequalities on the probability content of convex regions for elliptically contoured distributions," In:Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, Vol. II: Probability Theory, pp. 241-265. University of California Press, Berkeley, California, 1972.

[4] H. W. Gomez, H. Varela, L. Vidal, "A new class of skew-symetric distributions and related families," Statistics, vol. 47, pp. 411-421, 2013.

[5] A. Jamalizadeh, N. Balakrishnan, "Order statistics from trivariate normal and t_ν-distributions in terms of generalized skew-normal and skew-t_ν distributions," Journal of Statistical Planning and Inference, vol. 139, pp. 3799-3819, 2009.

[6] N. Loperfido, "Quadratic forms of skew-normal random vectors," Statistics & Probability Letters, vol. 54, pp. 381-387, 2001.

[7] N. Loperfido, J. Navarro, J. M. Ruiz, C. J. Sandoval, "Some relationships between skew-normal distributions and order statistics from exchangeable normal random vectors," Communication in Statistics – Theory and Methods, vol. 36, pp. 1719-1733, 2007.

[8] M. Shaked, J. Shanthikumar, "Stochastic Orders," New York, Springer, 2007.

The 45th Annual Iranian Mathematics Conference
August 26-29 2014
چهل و پنجمین
کنفرانس ریاضی ایران
۴ - ۷ شهریور ۱۳۹۳

Stochastic comparison of the residual and past lifetimes of two $(n-k+1)$-out-of-n systems with non-identical and dependent components

V. Amirzadeh[*], Academic member of Department of statistics, Shahid Bahonar University of Kerman, Kerman, Iran (PhD), v_amirzadeh@uk.ac.ir

M. Rezapour, Academic member of Department of statistics, Shahid Bahonar University of Kerman, Kerman, Iran (PhD), mohsenrzp@uk.ac.ir

Abstract: In this paper, we compare the residual and past lifetimes of two $(n-k+1)$-out-of-n systems with non-identical and dependent components according to a multivariate log-normal distribution. We also show that two univariate doubly truncated normal distributions with same variance and different means are stochastically ordered with respect to hazard rate stochastic ordering.

Keywords: Stochastic ordering, hazard rate ordering, doubly truncated, multivariate log-normal, $(n-k+1)$-out-of-n system.

1 Introduction

The normal distribution has been used to model the potential returns on an investment in capital budgeting and portfolio analysis (see [10], and references therein). Normal distributions have often been investigated in the context of stochastic dominance (e.g., [11, 3, 14, 2]). However, there are three reasons why investigation of the truncation of the distributions may be appropriate. First, the fact that the normal is unbounded indicates that it is only to be used as an approximation to an actual return distribution. In particular, the left tail of a return distribution must be zero below the maximum loss possible for an investment. In many investment situations, the maximum loss possible is limited to the amount of the investment. In any case, an investment must have a limited liability no matter how large that amount may be. Second, Ben-Horim [2] censored his sample data before testing for stochastic dominance. The motivation was to eliminate outliers and thereby increase the chances that the sample data would then agree with the dominance relation existing between the two populations (i.e., reducing estimation risk). Third, when X is the lifetime of a device, $X_t = [X - t \mid X > t]$ is the residual lifetime of the device at time t, given that the device is alive at time t. As a dual notion to the residual life, the past lifetime (or the inactivity time) measures the time elapsed since the system hazard. Past lifetime of a system is also a truncated random variable.

Consider one degrading item which operates in a baseline environment (regime) and denote the corresponding distribution function of time to hazard X by $F(t)$. By degrading we mean that the

[*]Corresponding Author

hazard occurs due to some degradation processes, e.g., as a result of wear accumulation. Let another statistically item described by the lifetime Y with the distribution function $G(t)$ be operating in a more severe environment. Assume for simplicity that environments are not varying with time and that distributions are absolutely continuous and let $\lambda_X(t)$ and $\lambda_Y(t)$ denote the corresponding hazard rates. It is reasonable to assume that degradation in the second regime is more intensive and, therefore, the time for accumulating the same amount of degradation or wear is smaller than the baseline one. Therefore, assume that the corresponding lifetimes are ordered in terms of the (usual) stochastic ordering (see e.g. [13]), i.e.,

$$\bar{F}(t) = 1 - F(t) \leq \bar{G}(t) = 1 - G(t), t \in \mathcal{R}, \quad (1)$$

in which case, we say X is smaller than Y in (usual) stochastic order and denote it by $X \prec_{st} Y$. This general relationship naturally models the impact of a more severe environment as compared with the baseline one. Note that, the corresponding hazard rate ordering denoted by $X \prec_{hr} Y$ is defined as

$$\lambda_X(t) \geq \lambda_Y(t), t \in \mathcal{R}. \quad (2)$$

which is also very popular in reliability applications and it is more stronger one such that (2) leads to (1).

A $(n-k+1)$-out-of-n system is a system consisting of n components that works if and only if at least $n-k+1$ of its components are operating ($k \leq n$). Thus, this system fails if k or more of its components fail (see [7] for a thorough discussion of these systems). It is indeed a very popular and commonly studied reliability structure. If we denote the lifetimes of the individual components by $X_1, \ldots, X_n$, then the lifetime of the $(n-k+1)$-out-of-n system is simply the kth order statistic $X_{k:n}$. The concept of the past lifetime of the components of a parallel system (at the system level), under the condition that the system has failed by time t, has been introduced in [1] as

$$(t - X_{l:n} | X_{n:n} \leq t), \quad l = 1, 2, \ldots, n.$$

Khaledi and Shaked [6] considered the residual life of a coherent system, given that at least $(n-k+1)$ components of the system are working, and gave some stochastic comparison results for this system. Li and Zhao [9] carried out a stochastic comparison on residual and past lifetimes of two $(n-k+1)$-out-of-n systems and generalized the results in [6].

Usually, it is assumed that the n lifetimes $X_1, X_2, \ldots, X_n$ of the components of the system are independent and identically distributed. In particular, however, there may be a structural dependence among the components of the system. Such a dependency may be due to a common shock affecting system's components. The objective of the present work is to model the structural dependence of the system using multivariate normal distribution and then establish some stochastic ordering results.

The rest of this paper is organized as follows. In Section 2, we extend the main result of Müller [12] to the case of multivariate truncated normal distribution, and then present some conditions under which we can compare the residual and past lifetimes of two $(n-k+1)$-out-of-n systems whose component lifetimes are distributed as multivariate log-normal. In Section 3, we consider stochastic hazard rate ordering of two doubly univariate truncated normal distributions.

2 Stochastic Comparisons of Truncated Normal Distributions

The following multivariate generalizations of the usual stochastic order are well-known in the literature; see Shaked and Shanthikumar [13] for related properties, equivalent definitions and applications. Consider two multivariate random vectors $\mathbf{X}$ and $\mathbf{Y}$. We say that $\mathbf{X}$ is smaller than $\mathbf{Y}$ in the *usual stochastic order* (denoted by $\mathbf{X} \prec_{st} \mathbf{Y}$) if, and only if, $E[h(\mathbf{X})] \leq E[h(\mathbf{Y})]$ for every increasing function

The 45th Annual Iranian Mathematics Conference
August 26-29 2014
چهل و پنجمین
کنفرانس ریاضی ایران
۴ - ۷ شهریور ۱۳۹۳

$h : \mathcal{R}e^d \to \mathcal{R}e$ provided that both expectations exist. In this section, we consider stochastic orderings of truncated normal distributions. Let us assume $\mathbf{X}_i \sim N_n(\boldsymbol{\mu}_i, \boldsymbol{\Sigma})$, $i = 1, 2$, $\mathbf{Y}_i \sim \mathbf{X}_i | (\mathbf{c} \le \mathbf{X}_i \le \mathbf{d})$, where $\mathbf{c} = (c_1, \ldots, c_n)$ and $\mathbf{d} = (d_1, \ldots, d_n)$ are vectors of constants. Here, we consider joint density function of a random vector $\mathbf{Z}$ as

$$\phi_\lambda(\mathbf{x}) = \frac{\exp\{-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}(\lambda))'\boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu}(\lambda))\}}{\int_{\mathbf{m}(\lambda)}^{\mathbf{n}(\lambda)} \exp\{-\frac{1}{2}\mathbf{u}'\boldsymbol{\Sigma}^{-1}\mathbf{u}\}d\mathbf{u}}, \quad (3)$$

where $\mathbf{m}(\lambda) = \mathbf{c} - \boldsymbol{\mu}(\lambda)$, $\mathbf{n}(\lambda) = \mathbf{d} - \boldsymbol{\mu}(\lambda)$, and $\boldsymbol{\mu}(\lambda) = \lambda\boldsymbol{\mu}_1 + (1 - \lambda)\boldsymbol{\mu}_2$. Let us introduce the following notation:

$$\boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_{-l} \\ \mu_l \end{pmatrix}, \quad \boldsymbol{\Sigma}^{-1} = \begin{pmatrix} \boldsymbol{\Sigma}_{-l}^{(-1)} & \boldsymbol{\sigma}_l^{(-1)} \\ (\boldsymbol{\sigma}_l^{(-1)})' & \sigma_{ll}^{(-1)} \end{pmatrix},$$

$$\mathbf{m}(\lambda) = \begin{pmatrix} \mathbf{m}_{-l}(\lambda) \\ m_l(\lambda) \end{pmatrix}, \quad \mathbf{n}(\lambda) = \begin{pmatrix} \mathbf{n}_{-l}(\lambda) \\ n_l(\lambda) \end{pmatrix},$$

where $l = 1, \ldots, n$.

Now, we present a result that can be used for comparing two multivariate truncated normal random variables, with same truncation and covariance matrices, with respective to multivariate usual stochastic ordering. This result is an extension of the main theorem in [12]. Since the proof of this result is long and involved, we omit the proof of the Theorem.

Theorem 2.1. *Suppose $f(\mathbf{x}) : \mathcal{R}^n \to \mathcal{R}$ is continuously differentiable and vanishes outside $(\mathbf{c}, \mathbf{d})$. Then,*

$$\begin{aligned} Ef(\mathbf{Y}_1) - Ef(\mathbf{Y}_2) = {} & \sum_{l=1}^{n} \mu_{12l} \\ & \int_0^1 \Bigg(\int_{\mathbf{c}}^{\mathbf{d}} \frac{\partial}{\partial x_l} f(\mathbf{x})\phi_\lambda(\mathbf{x})d\mathbf{x} + \\ & \int_{\mathbf{c}}^{\mathbf{d}} f(\mathbf{x})\phi_\lambda(\mathbf{x})\Big(I_{d_l}(x_l) - I_{c_l}(x_l)\Big)d\mathbf{x} \\ & + \int_{\mathbf{c}}^{\mathbf{d}} \phi_\lambda(\mathbf{x})\Big(I_{d_l}(x_l) - I_{c_l}(x_l)\Big)d\mathbf{x} \\ & \int_{\mathbf{c}}^{\mathbf{d}} f(\mathbf{x})\phi_\lambda(\mathbf{x})d\mathbf{x}\Bigg)d\lambda, \end{aligned} \quad (4)$$

where $I_z(x)$ equals one when $x = z$ and zero otherwise, and $\mu_{ijl} = \mu_l^{(i)} - \mu_l^{(j)}, i, j = 1, 2, l = 1, \ldots, n$, with $\boldsymbol{\mu}_i = (\mu_1^{(i)}, \ldots, \mu_n^{(i)})$.

Corollary 2.2. *If $\boldsymbol{\mu}_1 \le \boldsymbol{\mu}_2$, then $\mathbf{Y}_1 \prec_{st} \mathbf{Y}_2$.*

Since the the residual [past] lifetime of a $(n - k + 1)$-out-of-n system whose component lifetimes are distributed as multivariate normal, $(X_{k:n}|X_{1:n} \ge t)$ $[(X_{k:n}|X_{n:n} \le t)]$, can be considered as the order statistics from a multivariate truncated normal distribution, we can compare the residual and past lifetimes of two $(n-k+1)$-out-of-n system whose component lifetimes are distributed as multivariate normal, by Theorem 2.1 and Theorem 6.B.23 in [13] as follows.

Corollary 2.3. *Assume $\mathbf{X}$ and $\mathbf{Y}$ are two multivariate normal random vectors with means $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$, respectively, and same covariance matrix. Moreover, if $\boldsymbol{\mu}_1 \le \boldsymbol{\mu}_2$ and $X_{k:n}$ and $Y_{k:n}$, $k = 1, \ldots, n$, are the order statistics arising from random vectors $\mathbf{X}$ and $\mathbf{Y}$, respectively, then $(X_{k:n}|X_{1:n} \ge t) \prec_{st} (Y_{k:n}|Y_{1:n} \ge t)$ and $(X_{k:n}|X_{n:n} \le t) \prec_{st} (Y_{k:n}|Y_{n:n} \le t)$ for any constant t.*

Now, we can present a result regarding the comparison of residual and past lifetimes of two $(n - k + 1)$-out-of-n systems with respect to the usual stochastic ordering. In this case, we assume the components of the systems to be jointly distributed as multivariate log-normal.

Corollary 2.4. *Assume $\mathbf{X}$ and $\mathbf{Y}$ are two multivariate log-normal random vectors with the corresponding normal distributions having means $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$, respectively, and same covariance matrix. Moreover, if $\boldsymbol{\mu}_1 \le \boldsymbol{\mu}_2$ and $X_{k:n}$ and $Y_{k:n}$, $k = 1, \ldots, n$, are the order statistics arising from random vectors $\mathbf{X}$ and $\mathbf{Y}$, respectively, then $(X_{k:n}|X_{1:n} \ge t) \prec_{st} (Y_{k:n}|Y_{1:n} \ge t)$ and $(X_{k:n}|X_{n:n} \le t) \prec_{st} (Y_{k:n}|Y_{n:n} \le t)$ for any constant t.*

3 Univariate doubly truncated normal distribution

In this section, we consider stochastic hazard rate orderings of univariate truncated normal distributions. Let us assume $X_i \sim N(\mu_i, \Sigma)$, $i = 1, 2$, $Y_i \sim X_i|(c \le X_i \le d)$. Here, we consider the density function of a random variable Z as

$$\phi_\lambda(x) = \frac{\exp\{-\frac{(x-\mu(\lambda))^2}{2\sigma^2}\}}{\int_{m(\lambda)}^{n(\lambda)} \exp\{-\frac{u^2}{2\sigma^2}\}du}, \tag{5}$$

where $m(\lambda) = c - \mu(\lambda)$, $n(\lambda) = d - \mu(\lambda)$, and $\mu(\lambda) = \lambda\mu_1 + (1-\lambda)\mu_2$. With an argument similar to the one in the proof of Theorem 2.1, we can obtain the same results for the univariate doubly truncated normal distribution as follows.

Theorem 3.1. *Suppose $f(x) : \mathcal{R} \to \mathcal{R}$ is continuously differentiable and vanishes outside (c, d) and $g(\lambda) = \int_c^d f(x)\phi_\lambda(x)dx$. Then,*

$$\begin{aligned} Ef(Y_1) - Ef(Y_2) = \int_0^1 g'(\lambda)d\lambda = (\mu_1 - \mu_2) \\ \int_0^1 \Big(- \int_c^d f(x)\frac{\partial}{\partial x}\phi_\lambda(x)dx \\ + \int_c^d f(x)\phi_\lambda(x)\Big(I_d(x) - I_c(x)\Big)dx \\ + \int_c^d \phi_\lambda(x)\Big(I_d(x) - I_c(x)\Big)dx \\ \int_c^d f(x)\phi_\lambda(x)dx\Big)d\lambda, \end{aligned} \tag{6}$$

where $I_z(x)$ equals one when $x = z$ and zero otherwise.

We now consider a result which compares hazard rate functions of two doubly truncated normal distribution functions with same variance in terms of their means.

Corollary 3.2. *If $\mu_1 \le \mu_2$, then $Y_1 \prec_h Y_2$.*

References

[1] M. Asadi (2006). On the mean past lifetime of the components of a parallel system, *Journal of Statistcal Planning and Inference* **136**, 1197-1206.

[2] M. Ben-Horim (1990). Stochastic dominance and truncated sample data, *Journal of Financial Research* **13**, 105-116.

[3] Y. Kroll, H. Levy (1980). Sampling errors and portfolio efficient analysis, *Journal of Financial and Quantitative Analysis* **15**, 655-688.

[4] H.A. David, H.N. Nagaraja (2003). *Order Statistics*, Third edition, John Wiley & Sons, Hoboken, New Jersey.

[5] W.C. Horrace (2005). Some results on the multivariate truncated normal distribution, *Journal of Multivariate Analysis* **94**, 209-221.

[6] B. Khaledi, M. Shaked (2007). Ordering conditional lifetimes of coherent systems, *Journal of Statistical Plannning and Inference* **137**, 1173-1184.

[7] W. Kuo, M. J. Zuo (2003). *Optimal Reliability Modeling: Principles and Applications,* John Wiley & Sons, Hoboken, New Jersey.

[8] H. Lèvy (1982). Stochastic dominance rules for truncated normal distributions: A note, *Journal of Finance* **37**, 1299-1303.

[9] X. Li, P. Zhao (2008). Stochastic comparison on general inactivity time and general residual life of $(n-k+1)$-out-of-n systems, *Communications in Statisttics-Simulation and Computation* **37**, 1005-1019.

[10] R. Norgaard, T. Killeen (1980). Expected utility and the truncated normal distribution, *Management Science* **26**, 901-909.

[11] R. Meyer (1989). Two-moment decision models and ex- pected utility maximization: Reply, *American Economic Review* **79**, 603.

[12] A. Müller (2000). Stochastic ordering of multivariate normal distributions, *Annalas of innstitie of statistical mathematics* **53**, 567-575.

[13] M. Shaked, J. G. Shanthikumar (2007). *Stochastic Orders and Their Applications.* Springer, New York.

[14] W.E. Stein, R.C. Pfaffenberger (1986). Estimation of error probabilities in stochastic dominance, *Journal of Statistical Computation and Simulation* **25**, 137-165.

The 45th Annual Iranian Mathematics Conference
August 26-29 2014
چهل و پنجمین
کنفرانس ریاضی ایران
۴ - ۷ شهریور ۱۳۹۳

مدل بندی داده‌های بقا طولانی ـ مدت با استفاده از تابع مفصل کلایتون

مرضیه روزبهانی* ، دانشجوی کارشناسی ارشد آمار، دانشگاه شهید چمران اهواز، m-ruzbahani@mscstu.scu.ac.ir
محمد رضا آخوند، عضو هیأت علمی گروه آمار، دانشگاه شهید چمران اهواز، mra.biostat@gmail.com
سلیمان خیری، عضو هیأت علمی گروه آمار و اپیدومولوژی، دانشگاه علوم پزشکی شهرکرد، kheiri@hbi.ir.ac.ir

چکیده: در حالت کلی داده‌های بقا چند متغیره به هم وابسته هستند و یکی از راه‌های در نظر گرفتن وابستگی بین داده‌ها استفاده از تابع مفصل می‌باشد. در این مقاله قصد داریم داده‌های بقا طولانی مدت را با استفاده از مفصل کلایتون مدل‌سازی کنیم. برای این منظور برآورد پارامترها را به کمک روش بیزی به دست می‌آوریم. از آن‌جایی که توزیع‌های پسین دارای فرم بسته‌ای نمی‌باشند، برآورد مشخصات توزیع‌های پسین را با به کارگیری روش مونت کارلوی زنجیر مارکوفی به دست خواهیم آورد.

کلمات کلیدی:تابع مفصل، بقا طولانی مدت، روش بیزی، روش‌های مونت کارلوی زنجیر مارکوفی

مقدمه

در بسیاری از زمینه‌های کاربردی با داده‌های زمان شکستی برخورد می‌کنیم که ساختاری دو یا چند متغیره دارند، در این حالت نمی‌توان زمان های شکست را مستقل فرض نمود. داده‌های بقا چند متغیره در حالت کلی به هم وابسته هستند و مطالعه‌ای از وابستگی بین متغیرها معمولا مورد توجه محققین است. از جمله افرادی که بر روی مدل‌سازی داده‌های بقا چند متغیره کار کرده‌اند می‌توان به اسلانید و همکاران[۱]، رومئو و همکاران[۲] و هنگل[۳] اشاره نمود. برای در نظر گرفتن ساختار وابستگی میان چنین داده‌هایی روش‌های مختلفی مورد استفاده قرار می‌گیرد. که رایج‌ترین آن مدل شکنندگی است که یک یا چند اثر تصادفی به منظور نشان دادن وابستگی بین مشاهدات در مدل وارد می‌شود و زمان‌های بقا به طور شرطی با در نظر گرفتن اثر تصادفی از هم مستقل در نظر گرفته می‌شوند. وایپل و همکاران[۴] برای اولین بار مدل‌های شکنندگی را مطرح کردند. در سال‌های اخیر، توابع مفصل به عنوان یک ابزار مناسب برای بیان وابستگی بین متغیرها مطرح شده‌اند و در بیشتر زمینه‌های پزشکی، مالی و بیمه مورد استفاده قرار گرفته‌اند. مفصل‌ها توابعی هستند که تابع توزیع چند متغیره را به تابع توزیع حاشیه‌ای یک متغیره متصل می‌کنند.

در آنالیز بقا به طور کلی فرض می‌شود که همه افراد در معرض خطر هستند و پیشامد مورد نظر را تا

پایان مطالعه تجربه می‌کنند. اگر قسمتی از جامعه پیشامد مورد علاقه را تجربه نکنند این‌گونه از افراد تحت عنوان افراد با بقا طولانی مدت یا شفایافته نامیده می‌شوند. رایج‌ترین نوع از مدل‌های شفایافته، مدل آمیخته است که توسط بوگ[۵] ارائه شده است. لئوزادا و همکاران[۶] از تابع مفصل FGM برای مدل‌بندی داده‌های بقا دو متغیره استفاده کردند. خیری و همکاران[۷] از مدل‌های شکنندگی برای مدل‌بندی داده‌های قوز قرنیه استفاده کردند. در این مقاله قصد داریم به مدل‌سازی داده‌های بقا طولانی مدت با استفاده از توابع مفصل بپردازیم و در پایان از این مدل‌ها برای مدسازی داده‌های پیوند قوز قرنیه استفاده کنیم.

مدل آماری

مفصل‌ها توابع چند متغیره‌ای هستند که توابع توزیع حاشیه‌ای آن‌ها به صورت یکنواخت روی بازه‌ی $(0,1)$ توزیع شده‌اند. در حالت دو متغیره فرض کنید که (T_1, T_2) متغیرهای تصادفی پیوسته با توابع بقا حاشیه‌ای به ترتیب (S_1, S_2) باشند. در این صورت تابع بقا توام را با استفاده از مفصل C_ϕ می‌توان به صورت زیر نوشت:

$$S(t_1, t_2) = C_\phi\left(S_1(t_1), S_2(t_2)\right)$$

که با استفاده از تابع مفصل کلایتون تابع بقا توام به صورت زیر به دست می‌آید:

$$S(t_1, t_2) = \left(S_1(t_1)^{-\alpha} + S_2(t_2)^{-\alpha} - 1\right)^{-\frac{1}{\alpha}}$$

که در آن $S_1(t_1)$ و $S_2(t_2)$ توابع بقا حاشیه‌ای برای T_1 و T_2 به ترتیب و α پارامتر وابستگی است. اگر $\alpha \to 0$ حاشیه‌ها مستقل خواهند بود یعنی $S(t_1, t_2) = S(t_1) \times S(t_2)$. با مشتق‌گیری از تابع بقا توام نسبت به t_1 و t_2 می‌توان تابع چگالی توام را به دست آورد. با استفاده از ضریب همبستگی تاوکندال می‌توان میزان همبستگی بین متغیرها را بهتر درک کرد که در تابع مفصل کلایتون از فرمول زیر به دست می‌آید:

$$\tau_\alpha(T_1, T_2) = \frac{\alpha}{\alpha+2}$$

تابع بقا شفایافته

در مدل‌های شفایافته آمیخته فرض بر این است که جامعه به صورت ناهمگن به دو زیر جامعه از افراد مصون یا شفایافته و افراد در معرض خطر تقسیم شده است. در این مدل‌ها فرض می‌شود که هر فرد با احتمال p شفایافته و با احتمال $1-p$ در معرض خطر است. در مدل‌های آمیخته تابع بقا برای کل افراد جامعه به صورت زیر است:

$$S_j(t_j) = p_j + (1-p_j)\, S_0(t_j)$$

که p درصدی از افراد شفایافته است و $S_0(t_j)$ تابع بقا برای افراد شفانیافته است. که معمولا برای آن توزیع نمایی، وایبل و گامبرتز در نظر گرفته می‌شود.

تشکیل تابع درستنمایی

فرض کنید که زمان‌های طول عمر (T_{i1}, T_{i2}) و زمان سانسور شدن (C_{i1}, C_{i2}) برای $i = 1,..,n$ از یکدیگر مستقل هستند. در نتیجه $t_{ij} = min(T_{ij}, C_{ij})$ طول عمر مشاهده شده برای فرد i ام و $\delta_{ij} = I(t_{ij} = T_{ij})$ نشانگر سانسور برای $j = 1,2$ است. در نتیجه تابع درستنمایی برای داده‌ها به صورت زیر به دست می‌آید:

$$L(\theta) = \prod_{i=1}^{n} f(t_{i1}, t_{i2})^{\delta_{i1}\delta_{i2}} \frac{\partial S(t_{i1}, t_{i2}|\theta)^{\delta_{i1}(1-\delta_{i2})}}{\partial t_{i1}} \times$$

$$\frac{\partial S(t_{i1}, t_{i2}|\theta)^{\delta_{i2}(1-\delta_{i1})}}{\partial t_{i2}} \times S(t_{i1}, t_{i2})^{(1-\delta_{i1})(1-\delta_{i2})}$$

با در نظر گرفتن توزیع وایبل برای حاشیه‌ها $T_j \sim weibull(r_j, \lambda_j)$ متغیرهای کمکی از طریق پارامتر $\lambda_{ij} = (\beta_j + \beta_{ij} X_{ij})$ وارد مدل می‌شوند و تابع درستنمایی با در نظر گرفتن فرضیات بالا به صورت

The 45th Annual Iranian Mathematics Conference
August 26-29 2014
چهل و پنجمین کنفرانس ریاضی ایران
۴-۷ شهریور ۱۳۹۳
دانشگاه سمنان

زیر به دست می‌آید
:

$$L=\prod_{i=1}^{n}(\alpha+1)^{\delta_{i1}\delta_{i2}}\left(f_1\left(t_{i1}\right)S_{i1}(t_{i1})^{-\alpha-1}\right)^{\delta_{i1}}$$
$$\left(f_{i2}\left(t_{i2}\right)S_{i2}(t_{i2})^{-\alpha-1}\right)^{\delta_{i2}}$$
$$\left(S_{i1}(t_{i1})^{-\alpha}+S_{i2}(t_{i2})^{-\alpha}-1\right)^{-\delta_{i1}-\delta_{i2}-\frac{1}{\alpha}}$$

که

$$S_j\left(t_j\right)=p_j+\left(1-p_j\right)\exp\left(-\lambda_j t_j^{r_j}\right)$$

و

$$f_j\left(t_j\right)=\left(1-p_j\right)\lambda_j r_j t_j^{r_j-1}\exp\left(-\lambda_j t_j{}^{r_j}\right)$$

به ترتیب تابع بقا شفایافته آمیخته وایبل و چگالی آمیخته وایبل برای $j=1,2$ است.

چگالی پیشین و پسین

پس از مدل‌سازی داده‌های بقا طولانی مدت می‌توان از روش‌های مختلفی برای برآورد پارامترها استفاده نمود، که از جمله آنها می‌توان روش‌های بیزی را نام برد. از مزیت‌های روش بیزی نسبت به روش کلاسیک این است که در صورتی که اطلاعات اضافی در مورد پارامترها در اختیار باشد می‌توان آن را از طریق توزیع پیشین وارد مدل نمود و در صورتی که چنین اطلاعاتی وجود نداشته باشد از توزیع پیشین ناآگاهی بخش که هیچ‌گونه اطلاعات اضافی وارد مدل نمی‌کنند، می‌توان استفاده کرد. با در نظر گرفتن توزیع‌های پیشین برای پارامترهای مدل به صورت $r_j\sim Gamma(a_j,b_j)$ ، $\beta_{ij}\sim N(\mu_{ij},\sigma_{ij})$ ، $p_j\sim Beta(v_j,u_j)$ و $\alpha\sim Gamma\left(c,d\right)$ توزیع پسین توام پارامترها به صورت زیر به دست می‌آید:

$$\pi\left(\alpha,r_1,r_2,\beta_1,\beta_2|X,\delta_1,\delta_2\right)\propto\prod_{i=1}^{n}(\alpha+1)^{\delta_{i1}\delta_{i2}}$$
$$\left(\left(1-p_1\right)e^{(\beta_{01}+\beta_{k1})}t_1^{r_1-1}\left(p_1+\left(1-p_1\right)e^{-\exp(\beta_{01}+\beta_{k1})t_1^{r_1}}\right)^{-\alpha-1}\right)^{\delta_{i1}}$$
$$\left(\left(1-p_2\right)e^{(\beta_{02}+\beta_{k2})}t_2^{r_2-1}\left(p_2+\left(1-p_2\right)e^{-\exp(\beta_{02}+\beta_{k2})t_2^{r_2}}\right)^{-\alpha-1}\right)^{\delta_{i2}}$$
$$\left(\left(p_1+\left(1-p_1\right)e^{-\lambda_1 t_1^{r_1}}\right)^{-\alpha}+\left(\left(p_2+\left(1-p_2\right)e^{-\lambda_2 t_2^{r_2}}\right)\right)^{-\alpha}-1\right)^{-\left(\frac{1}{\alpha}+\delta_{i1}+\delta_{i2}\right)}$$
$$a_1^{1}\,{r_1}^{a_1-1}e^{-b_1 r_1}\,{b_2}^{a_2}\,{r_2}^{a_1-1}e^{-b_2 r_2}\,c^{d}\,\alpha^{c-1}e^{-d\alpha}\,{p_2}^{v_2-1}\left(1-p_2\right)^{u_2-1}\mathrm{b}$$
$$\left(-\frac{1}{2\sigma_{01}^2}\left(\beta_{01}-\mu_{01}\right)^2\right)e^{\left(-\frac{1}{2\sigma_{02}^2}\left(\beta_{02}-\mu_{02}\right)^2\right)}\left(-\frac{1}{2\sigma_{k1}^2}\left(\beta_{k1}-\mu_{k1}\right)^2\right)e^{\left(-\frac{1}{2\sigma_{k2}^2}\left(\beta_{k2}-\mu_{k2}\right)^2\right.}$$

به دلیل پیچیدگی و ابعاد گسترده توزیع پسین توام به دست آمده، امکان محاسبه توزیع پسین پارامترها به روش تحلیلی وجود ندارد. برای تقریب توزیع پسین پارامترها از روش مونت کارلوی زنجیر مارکوفی استفاده می‌گردد. در این روش نمونه‌هایی تصادفی از توزیع پسین تولید می‌شود و بر اساس نمونه‌ها استنباط در مورد پارامترها صورت می‌گیرد.

مثال کاربردی

از آن‌جایی که مهمترین علت شکست پیوند، دفع عضو عضو پیوندی می‌باشد. بررسی و تعیین دقیق عوامل موثر بر دفع پیوند قرنیه از اهمیت ویژه‌ای برخوردار است. در این مقاله اطلاعات ۱۱۹ بیمار که طی سال‌های ۸۰-۶۵ تحت عمل پیوند قرنیه دو طرفه قرار گرفته‌اند،مورد تجزیه وتحلیل قرار گرفته است. متغیرهای کمکی موجود در این داده‌ها عبارتند از: جنس، سن، قطر قرنیه‌دهنده، قطر بستر گیرنده، تازگی قرنیه، ورم ملتحمه بهاره، واسکولاریزیشن و پیوند مجدد هستند که در مدل وارد شده‌اند.

برای انجام تحلیل بیزی از توزیع پیشین ناآگاهی بخش برای پارامترها به صورت $r_j\sim\beta_{ij}\sim N\left(0,100^2\right)$ ، $Gamma(0/1,0/001)$ ، $p_j\sim Beta(1,1)$ و $\alpha\sim Gamma\left(1,0/001\right)$ استفاده گردید.

برای انجام این تحلیل از نرم افزار OpenBugs استفاده شده است و نتایج حاصل براساس برآوردهای پسین مدل، که شامل میانگین، انحراف معیار و فاصله باورمند ۹۵درصد است در جدول ۱آمده است. تمام شبیه سازی‌ها براساس دو زنجیر به طول ۵۰ هزار که ۱۰ هزار نمونه اول به عنوان دوره داغیدن در نظر گرفته شده است و از هر ۵ مشاهده یک مشاهده انتخاب شده است بدست آمده است.

نتایج

تحلیل چندگانه داده‌ها با استفاده از مفصل کلایتون نشان داد که سن بیمار در زمان پیوند، قطر قرنیه دهنده، قطر بستر گیرنده، تازگی قرنیه، ورم ملتحمه بهاره و واسکولاریزیشن عوامل معنی‌داری در زمان دفع پیوند بوده‌اند. اما عواملی چون جنس بیمار و پیوند مجدد تاثیری بر زمان بقای پیوند نداشته‌اند. رحیم زاده و همکاران [۸]در مطالعه خود که با استفاده از مدل های شکنندگی انجام شده است نشان دادند که متغیرهای سن و واسکولاریزیشن دارای اثری معنی‌دار بر روی دفع پیوند می‌باشد.

جدول ۱: میانگین پسین، انحراف استاندارد و بازه باورمند ۰/۹۵ برای پیوند قوز قرنیه

پارامتر	میانگین	انحراف استاندارد	بازه باورمند ۰/۹۵
جنس	-0.369	0.583	(-1.539 , 0.775)
سن	0.077	0.032	(0.011 , 0.137)
قطر قرنیه دهنده	-3.244	0.830	(-4.489 , -1.67)
قطر بستر گیرنده	4.119	1.493	(2.274 , 6.675)
تازگی قرنیه	1.265	0.645	(0.052 , 2.562)
ورم ملتحمه بهاره	3.049	0.874	(1.259 , 4.652)
واسکولاریزیشن	3.36	0.758	(1.824 , 4.793)
پیوند مجدد	-2.302	1.525	(-5.664 , 0.381)
نسب شفایافتگی	0.545	0.073	(0.39 , 0.679)
میزان همبستگی	0.498	0.45	(0.016, 1.649)

مراجع

[1] H. Aslanidou, D. K. Dey, and D. sinha. (2008). ”Bayesian analysis of multivariate survival data using monte carlo methods,” Canadian Journal of stastistics. 1, 33-48

[2] J. S. Romeo, N. I. Tanaka and A. C. Pedroso de Lima.(2006). ”Bivariate survival odeling a Bayesian approach based on copulas,” Lifetime Data Anal, 12: 205-222.

[3] D. Hanagal.(2011).” Modeling survival data using frailty models,” Chapman and Hall, Boca Raton.

[4] J. W. Vaupel and K. G. Manton.(1979). ”The impact of heterogeneity in individual frailty on the dynamics of mortality,” Demography, 16, 439-454.

[5] J. W. Boag.(1949). ”Maximum likelihood estimates of the proportion of patients cured by cancer therapy,” Journal of the Royal Statistical Society, 11, 15-53

[6] F.Louzada, A. K. Suzuki. and V. G. Cancho. (2013). ”The FGM long-term bivariate survival copula model: Modeling bayesian estimation and case influence diagnostics,” Communications in Statistics – Theory and Methods, 42:4,673-691.

[7] S. Kheiri, A. Kimber and M. R. Meshkani(2007). ”Bayesian analysis of an inverse Gaussian correlated frailty model,” Computational Statistics and Data Analysis, 51, 5317 – 5326.

[8] M. Rahimzadeh, E. Hajizadeh and F. Eskandari(2011). ”Non-mixture cure correlated frailty models in Bayesian approach,” Journal of Applied Statistics, 38:8, 1651-1663.