Probability and Statistics

43rd Annual Iranian Mathematics Conference,
27-30 August 2012
University of Tabriz

Prediction of α-stable processes with incomplete past

Ali Abolhasani[1]
Amirkabir University of Technology

Mina Aminghafari
Amirkabir University of Technology

Abstract
Mohammadi and Mohammadpour (2009), present a method for the best linear prediction in the case of α-stable random processes with complete past. In this paper, we extend their method to find best linear prediction for α-stable random processes with incomplete past i.e. we have missing data. Simulation is presented for comparing new method and minimum dispersion method. The simulations proves good performance of proposed method.
Keywords: α-stable random processes, missing value, time series.

Mathematics Subject Classification: 37M10, 60G52, 60G25

1 Introduction

Cheng and Pourahmadi (1997), generalize innovation algorithm to solve the problems of prediction of future values based on incomplete past and interpolate missing values of stationary time series. Their method is useful when there are several missing values with arbitrary patterns, but the method is for finite variance processes.

Bondon (2002), obtains an explicit formula for the prediction error of future value of a stationary process when the infinite past is altered by some missing observations with an arbitrary pattern. The autoregressive representation of the predictor of a future value is proposed in Bondon (2002). In this method the variance of process is also finite. One method for predicting of infinite variance processes is minimum dispersion method which is presented by Brockwell and Davis (1991) in Chapter 13. Another method is the method of Mohammadi and Mohammadpour (2009), which finds best linear prediction for infinite variance processes. They introduce new norm and define weakly stationarity for processes with infinite variance.

For estimation missing value, a method was presented by Pourahmadi (2001, Chapter 8). In this method the space of all observed data is divided into two orthogonal subspaces and then projection of missing data on to each subspaces is found and by adding up those projections, prediction for missing value is obtained. Bondon (2002) extended the method proposed by Pourahmadi (2001, Chapter 8), for prediction of time series with incomplete past. In this paper we will use the idea of Bondon (2002), Cheng and Pourahmadi (1997), Pourahmadi (2001, Chapter 8) and the method of Mohammadi and Mohammadpour (2009) to predict time series with infinite variance and incomplete past. Forsake of simplicity we consider just one missing value, and we present a recursive algorithm to find coefficients in linear prediction equation.

2 Main Result

We know that any α-stable random processes with countable index set which centered with sequence of constants, has an integral representation with bounded integrands. For more detail see

[1]Department of Mathematics, Faculty of Science, Azarbayjan University of Tarbiat Moallem, Tabriz, Iran

the proof of Theorem 13.1.2 of Samordnitsky and Taqqu (1994). Based on this theorem for any real α-stable random processes $\{X(t), t \in \mathbb{N}\}$, there is a sequence of bounded measurable function f_t on (0,1) and real function $\mu(t)$, $t \in \mathbb{N}$, such that

$$\{X(t)+\mu(t), t \in \mathbb{N}\} \stackrel{d}{=} \Big\{ \int_{(0,1)} f_t(x) dM_\alpha(x), t \in \mathbb{N} \Big\} \tag{1}$$

where M_α is an α-stable random measure on measurable space $\big((0,1), \mathfrak{B}\big)$ with Lebesgue control measure, and skewness intensity $\beta \equiv 1$, also, $\mathfrak{B}$ is Borel set. The symbol $\stackrel{d}{=}$ means identical in distribution.

Consider the following class of α-stable random variables

$$\mathcal{C}_\alpha = \Big\{ \int_E f_t(x) dM_\alpha(x), \quad f_t \in L^2(\Omega, \mathcal{F}, P), t \in T \Big\} \tag{2}$$

For any $Y_1, Y_2 \in \mathcal{C}_\alpha$, stable covariation is defined as

$$SC(Y_1, Y_2) = \int_E \text{(Integrand of 1st stochastic integral) (Integrand of 2nd stochastic integral)} dp \tag{3}$$

With considering $SC(Y_1, Y_2)$ as an inner product, Mohammadi and Mohammadpour (2009) show that $\mathcal{C}_\alpha$ is a complete Hilbert space. Based on the definition of stable covariation, they extend projection theorem (Brockwell and Davis (1991) Chapter 5) for infinite variance case. They solve h-step prediction in infinite variance time series problem leading to a generalized Yule-Walker equation. To find coefficients of best linear prediction, it is just enough to solve $\Gamma \mathbf{a} = \mathbf{b}$, such that Γ is stable covariation matrix, $\mathbf{a}$ is coefficients in best linear prediction and $\mathbf{b}$ is stable covariation vector. Also they define weakly stationary for processes with infinite variance. For more details see Mohammadi and Mohammadpour (2009). We extend their method to predict α-stable processes with incomplete past. For simplicity we apply the notation $< Y_i, Y_j >$ or $SC(i,j)$ instead of $SC(Y_i, Y_j)$. In this paper, first we solve the simple case, then the results can be stated in general case.

Our strategy is as follows. Suppose $\mathcal{H} = \{X_1, X_2, X_3, X_4, X_6, X_7, X_8\}$ and we want to predict X_9 in presence of missing value X_5. We divide $\mathcal{H}$ into $\mathcal{H}_1 = \{X_1, X_2, X_3, X_4\}$ and $\mathcal{H}_2 = \{X_6 - \hat{X}_6, X_7 - \hat{X}_7, X_8 - \hat{X}_8\}$ such that $\hat{X}_i, i = 6,7,8$ be the projection of X_i onto $\mathcal{H}_1$, Then $X_i - \hat{X}_i$ is orthogonal to $\mathcal{H}_1$. We want to write elements of $\mathcal{H}_2$ such that $\mathcal{H}_2$ become an orthogonal space. So, we obtain a orthogonal set such that it is orthogonal to $\mathcal{H}_1$ too i.e.

$$\mathcal{H} = \overline{sp}\{X_1, X_2, X_3, X_4\} \oplus \overline{sp}\{X_6 - \hat{X}_6, X_7 - \hat{X}_7, X_8 - \hat{X}_8\} = \mathcal{H}_1 \oplus \mathcal{H}_2.$$

For finding $\hat{X}_i, i = 6,7,8$, one can use the method proposed by Mohammadi and Mohammadpour (2009). In fact via their method we find 2-step, 3-step, 4-step predictions based on $\{X_1, X_2, X_3, X_4\}$ and we have

$$\hat{X}_6 = \textstyle\sum_{j=1}^4 a_{6j} X_j, \qquad \hat{X}_7 = \sum_{j=1}^4 a_{7j} X_j, \qquad \hat{X}_8 = \sum_{j=1}^4 a_{8j} X_j$$

such that coefficient a_{ij}, $i = 6,7,8, j = 1,2,3,4$ are obtained from method of Mohammadi and Mohammadpour (2009). Now we convert $\mathcal{H}_2$ to orthogonal space in following way. Consider $Y_i = X_i - \hat{X}_i$ and $\hat{Y}_i = P_{\{Y_j - \hat{Y}_j, j<i\}} Y_i$, then

$$\mathcal{H}_2 = \overline{sp}\{X_6 - \hat{X}_6, X_7 - \hat{X}_7, X_8 - \hat{X}_8\} \equiv \overline{sp}\{Y_6, Y_7 - \hat{Y}_7, Y_8 - \hat{Y}_8\}$$

such that

$$\hat{Y}_7 = P_{\{Y_6\}} Y_7 = b_{76} Y_6 = b_{76}(X_6 - \hat{X}_6) = b_{76}(X_6 - a_{61}X_1 - a_{62}X_2 - a_{63}X_3 - a_{64}X_4)$$
$$\hat{Y}_8 = P_{\{Y_6, Y_7 - \hat{Y}_7\}} Y_8 = b_{86} Y_6 + b_{87}(Y_7 - \hat{Y}_7) = (b_{86} - b_{76}) Y_6 + b_{87} Y_7$$

If $\tilde{X}_9$ be orthogonal projection onto $\mathcal{H}$, then because of orthogonality of $\mathcal{H}_1$ and $\mathcal{H}_2$, we can write

$$\tilde{X}_9 = P_{\mathcal{H}_1}X_9 + P_{\mathcal{H}_2}X_9 = \hat{X}_9 + \hat{X'}_9$$

Finding $\hat{X}_9$ using the method presented by Mohammadi and Mohammadpour (2009), is an easy task. In fact we need to find 5-step prediction based on $\{X_1, X_2, X_3, X_4\}$. On the other hand, we know that $\hat{X'}_9$ is projection of X_9 onto $\mathcal{H}_2$, then $X_9 - \hat{X'}_9$ is orthogonal to all elements of $\mathcal{H}_2$. Now suppose $\hat{X'}_9$ as a linear combination of elements of $\mathcal{H}_2$ as follows

$$\hat{X'}_9 = \theta_{86}Y_6 + \theta_{87}(Y_7 - \hat{Y}_7) + \theta_{88}(Y_8 - \hat{Y}_8)$$

To find θ_{8j}, $j = 6, 7, 8$, we use inner product in both side of above equation. We know that any element in $\mathcal{H}_2$ is orthogonal to other ones and also we know

$$X_9 - \hat{X}_9 \perp (Y_6, Y_7 - \hat{Y}_7, Y_8 - \hat{Y}_8)$$

then we can find the θ_{ij}. After finding θ_{ij}, we are able to find $\tilde{X}_9$.

We can generalize the strategy of the proposed example to general case. We compare proposed method with minimum dispersion method by simulation. The results of simulation show that our method performs better.

References

[1] Bondon, P., 2002. Prediction with Incomplete Past of a Stationary Processes. *Stochastic Processes and Their Applications,* 98, 67-76.

[2] Brockwell , P. J., Davis, R. A., 1991. *Time Series : Theory and Methods.* 2nd ed. Springer-Verlag, New York.

[3] Cheng, R. and Pourahmadi, M., 1997. Prediction with Incomplete Past and Interpolation of Missing values. *Statistics and Probaility Letters,* 33, 341-346.

[4] Mohammadi, M., Mohammadpour, A., 2009. Best Linear Prediction for α-stable Random Processes. *Statistics and Probaility Letters,* 79, 2266-2272.

[5] Pourahmadi, M., 2001. *Foundations of Time Series Analysis and Prediction Theory.* Wiley Series in Probability and Statistics.

[6] Samordnitsky, G., Taqqu, M. S., 1994. *Stable non-Gussian Random Processes.* 3rd ed. Chapman and Hall, NewYork 5-50.

Email:abolhasani_alii@yahoo.com

Email:aminghafari@aut.ac.ir

43^{rd} Annual Iranian Mathematics Conference,
27-30 August 2012
University of Tabriz

Reliability for two Log-normal populations with common mean

K. Abdollahnezhad
Golestan University

F. Yeghmai
Golestan University

A. A. Jafari
Yazd University

S. Aghadoust
Golestan University

Abstract
A computational approach test is proposed for testing about reliability of two log-normal populations when the means are same. This approach does not require the explicit knowledge of the sampling distribution of the test statistic. Simulation studies demonstrate that the proposed approach can perform hypothesis testing with satisfying actual size even at small sample sizes and also is better than the score test.

Keywords: Log-normal distribution, Hypothesis test, Actual size, Reliability

Mathematics Subject Classification: 62-04, 62F99, 62F03, 62E17

1 Introduction

The log-normal distribution has many special characteristics and features together with its relation with the normal distribution that allowed it to be used as a model in various real life applications. In particular it is importance in modeling lifetimes of products and individuals (Lawless, 1982). Also as stated in Cheng (1997), reliability studies indicate that many semi-conductor devices follow lifetime distributions, which are well represented by the log-normal. Various other motivations and applications of the log-normal distribution can be found in Crow and Shimizu (1988).

Let $X_1 \sim LN(\mu_1, \sigma_1^2)$ independently of $X_2 \sim LN(\mu_2, \sigma_2^2)$, where μ_1, μ_2, σ_1^2, σ_2^2 are unknown. We know that $Y_1 = \log(X_1) \sim N(\mu_1, \sigma_1^2)$ and $Y_2 = \log(X_2) \sim N(\mu_2, \sigma_2^2)$. Then the reliability parameter under the assumption the equality of means of two log-normal populations can be expressed as

$$R = P(X_1 > X_2) = P(Y_1 > Y_2) = \Phi(\frac{\mu_1 - \mu_2}{\sqrt{\sigma_1^2 + \sigma_2^2}}) = \Phi(\frac{\sigma_1^2 - \sigma_2^2}{2\sqrt{\sigma_1^2 + \sigma_2^2}}), \quad (1)$$

where $\Phi(.)$ is the standard normal cumulative distribution function.

The problem of testing for the mean of the log-normal distribution has received considerable attention in the literature; see Zhou et al. (1997), Zhou and Tu (1999). Exact confidence intervals and tests for the ratio and difference of the means of two log-normal distributions using the generalized variable and generalized p-value is constructed by Krishnamoorthy and Mathew (2003). Score tests for testing the equality of the means of two independent log-normal populations, and for testing the reliability when the means are equal are proposed by Gupta and Li (2005).

The computational approach test (CAT) is introduced by Pal et al.(2007) for testing hypothesis. An advantage of this method is that it does not require the explicit knowledge of the sampling distribution of the test statistic and is easy to compute and implement. An application of CAT is provided for construct test hypothesis in strength-stress model in reliability of two log-normal populations with equal means. Our detailed study based on a comprehensive simulations indicate the actual sizes of our CAT are smaller than nominal level and better than the actual sizes of other existing approach.

2 Approaches

Let $X_{i1}, X_{i2}, ..., X_{in_i}$, $i = 1, 2$, be two independent random samples from log-normal population, i.e. $X_{ij} \sim LN(\mu_i, \sigma_i^2)$, $i = 1, 2$, $j = 1, ..., n_i$, where the parameters $(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2)$ are unknown. We know that $Y_{ij} = \ln(X_{ij}) \sim N(\mu_i, \sigma_i^2)$. Let us consider the problem of testing

$$H_0 : R = R_0 \qquad vs. \qquad H_a : R \neq (< or >)R_0, \tag{2}$$

where R_0 is a specified probability which is equivalent to

$$H_0^* : \gamma = 0 \qquad vs. \qquad H_a^* : \gamma \neq (< or >)0, \tag{3}$$

where $\gamma = \sigma_1^2 - \sigma_2^2 - 2r_0\sqrt{\sigma_1^2 + \sigma_2^2}$ and r_0 is the R_0-th quantile of the standard normal distribution. In the following, we present a score test method, introduced by Gupta and Li (2005), and derive the CAT for the parameter R in (1).

2.1 Score Test

Gupta and Li (2005) proposed a score test about reliability when the means are equal. When $\mu_1 + \frac{1}{2}\sigma_1^2 = \mu_2 + \frac{1}{2}\sigma_2^2 = \nu$, the MLE's of three parameters $(\nu, \sigma_1^2, \sigma_2^2)$ are obtained using solving the following equations

$$\tilde{\nu} = \frac{1}{\frac{n_1}{\tilde{\sigma}_1^2} + \frac{n_2}{\tilde{\sigma}_2^2}}\{\frac{n_1 + n_2}{2} + \frac{t_1}{\tilde{\sigma}_1^2} + \frac{t_2}{\tilde{\sigma}_2^2}\} \tag{4}$$

$$\tilde{\sigma}_i^2 = -2 + 2\sqrt{\tilde{\nu}^2 - 2\frac{t_i}{n_i}\tilde{\nu} + \frac{t_{i+2}}{n_i} + 1}, \qquad i = 1, 2. \tag{5}$$

where

$$(t_1, t_2, t_3, t_4) = \left(\sum_{j=1}^{n_1}\log(x_{1j}), \sum_{j=1}^{n_2}\log(x_{2j}), \sum_{j=1}^{n_1}(\log(x_{1j}))^2, \sum_{j=1}^{n_2}(\log(x_{2j}))^2\right).$$

A point estimate for R is

$$\tilde{R} = \Phi(\frac{\tilde{\sigma}_1^2 - \tilde{\sigma}_2^2}{2\sqrt{\tilde{\sigma}_1^2 + \tilde{\sigma}_2^2}}), \tag{6}$$

where $\tilde{\sigma}_i^2$, $i = 1, 2$ are obtained by using solving the equations in (4) and (5). Gupta and Li (2005) showed that $\frac{\tilde{r} - r}{\sqrt{\tilde{V}ar(\tilde{r})}} \to N(0, 1)$, where $\tilde{r} = \frac{\tilde{\sigma}_1^2 - \tilde{\sigma}_2^2}{2\sqrt{\tilde{\sigma}_1^2 + \tilde{\sigma}_2^2}}$ and $r = \frac{\sigma_1^2 - \sigma_2^2}{2\sqrt{\sigma_1^2 + \sigma_2^2}}$. Therefore, at level α, the test statistic $Z_0 = \frac{\tilde{r} - r_0}{\sqrt{\tilde{V}ar(\tilde{r})}}$, reject H_0^* for testing $H_0^* : \gamma = 0$ vs. $H_a^* : \gamma \neq 0$ iff $|Z_0| > Z_{\alpha/2}$, and reject H_0^* for testing $H_0^* : \gamma = 0$ vs. $H_a^* : \gamma > 0 (\gamma < 0)$ iff $Z_0 > Z_\alpha (Z_0 < -Z_\alpha)$.

2.2 CAT

As we know, the CAT finds its critical values automatically. In the following steps, we implement this method for testing (3):

step 1. First obtain $\tilde{\sigma}_i^2$, $i = 1, 2$ by using solving the equations in (4) and (5) and then obtain $\tilde{\gamma} = \tilde{\sigma}_1^2 - \tilde{\sigma}_2^2 - 2r_0\sqrt{\tilde{\sigma}_1^2 + \tilde{\sigma}_2^2}$.

step 2. Assume that $H_0^*:\gamma=0$ is true, i.e. $\sigma_1^2-\sigma_2^2-2r_0\sqrt{\sigma_1^2+\sigma_2^2}=0$ (or $\sigma_1^2=\sigma_2^2+2r_0(r_0+\sqrt{r_0^2+2\sigma_2^2})$). Under this restricted model, we have

$$\begin{aligned} Y_{1j} &= \log(X_{1j})\sim N(\mu-\frac{1}{2}\rho,\rho), \qquad j=1,...,n_1,\\ Y_{2j} &= \log(X_{2j})\sim N(\mu-\frac{1}{2}\sigma_2^2,\sigma_2^2), \qquad j=1,...,n_2. \end{aligned}$$

where $\rho=\sigma_2^2+2r_0(r_0+\sqrt{r_0^2+2\sigma_2^2})$. Therefore, the joint log-likelihood function can be written as

$$\begin{aligned} \log(L(\mu,\sigma_2^2)) &= c-\frac{n_1}{2}\log(\rho)-\frac{n_2}{2}\log(\sigma_2^2)-\frac{1}{2}(t_1+t_2)+\frac{\nu}{\rho}t_1-\frac{1}{2\rho}t_3\\ &\quad -\frac{(\nu-0.5\rho)^2}{2\rho}n_1+\frac{\nu}{\sigma_2^2}t_2-\frac{1}{2\sigma_2^2}t_4-\frac{(\nu-0.5\ \sigma_2^2)^2}{2\sigma_2^2}n_2, \end{aligned}$$

Call the MLE of parameters ν and σ_2^2 as the "restricted MLE (RML)", and denote by $\hat{\hat{\nu}}$ and $\hat{\hat{\sigma}}_2^2$. Then, they are obtained by solving the following equations:

$$\hat{\hat{\nu}}=\frac{2t_1\hat{\hat{\sigma}}_2^2+n_1\hat{\hat{\sigma}}_2^2\hat{\hat{\rho}}+2t_2\hat{\hat{\rho}}+n_2\hat{\hat{\rho}}\hat{\hat{\sigma}}_2^2}{2(n_1\hat{\hat{\sigma}}_2^2+n_2\hat{\hat{\rho}})}, \tag{7}$$

and

$$\begin{aligned} &\left(1+\frac{2r_0}{\sqrt{r_0^2+2\hat{\hat{\sigma}}_2^2}}\right)\times\\ &\left(-n_1\hat{\hat{\rho}}\hat{\hat{\sigma}}_2^4-2\hat{\hat{\nu}}t_1\hat{\hat{\sigma}}_2^4+t_3\hat{\hat{\sigma}}_2^4+\hat{\hat{\rho}}n_1\hat{\hat{\sigma}}_2^4\left(\hat{\hat{\nu}}-0.5\hat{\hat{\rho}}\right)+n_1\hat{\hat{\sigma}}_2^4\left(\hat{\hat{\nu}}-0.5\hat{\hat{\rho}}\right)^2\right)\\ &-\hat{\hat{\rho}}^2\hat{\hat{\sigma}}_2^2n_2-2\hat{\hat{\rho}}^2\hat{\hat{\nu}}t_2+\hat{\hat{\rho}}^2t_4+\hat{\hat{\rho}}^2\hat{\hat{\sigma}}_2^2n_2\left(\hat{\hat{\nu}}-0.5\hat{\hat{\sigma}}_2^2\right)+\hat{\hat{\rho}}^2n_2\left(\hat{\hat{\nu}}-0.5\hat{\hat{\sigma}}_2^2\right)^2=0, \end{aligned}$$

where $\hat{\hat{\rho}}=\hat{\hat{\sigma}}_2^2+2r_0(r_0+\sqrt{r_0^2+2\hat{\hat{\sigma}}_2^2})$.
step 3. Generate artificial samples a large number of times (Say, M times)

$$Y_{1j}\sim N(\hat{\hat{\nu}}-\frac{1}{2}\hat{\hat{\rho}},\hat{\hat{\rho}}), \quad Y_{2j}\sim N(\hat{\hat{\nu}}-\frac{1}{2}\hat{\hat{\sigma}}_2^2,\hat{\hat{\sigma}}_2^2), \qquad j=1,...,n_2,$$

and for each of these replicated samples, recalculate the MLE of γ based on step 1. Thus we will have $\tilde{\gamma}_{01},\tilde{\gamma}_{02},...,\tilde{\gamma}_{0M}$.
step 4. Let $\hat{\gamma}_{0(1)}\le\hat{\gamma}_{0(2)}\le...\le\hat{\gamma}_{0(M)}$ be the ordered values of $\hat{\gamma}_{0l}$, $l=1,...,M$.
step 5. (i) For testing $H_0^*:\gamma=0$ vs. $H_a^*:\gamma<0$, define the critical value $\tilde{\gamma}_L$ as $\tilde{\gamma}_L=\tilde{\gamma}_{0(\alpha M)}$. If $\tilde{\gamma}<\tilde{\gamma}_L$, then H_0^* is rejected. Alternatively the p-value can be defined as

$$p_1=\frac{1}{M}\sum_{l=1}^{M}I(\tilde{\gamma}_{0(l)}<\tilde{\gamma}). \tag{8}$$

(ii) For testing $H_0^*:\gamma=0$ vs. $H_a^*:\gamma>0$, define $\tilde{\gamma}_U=\tilde{\gamma}_{0((1-\alpha)M)}$. If $\tilde{\gamma}>\tilde{\gamma}_L$, then H_0^* is rejected. Alternatively the p-value can be defined as

$$p_2=\frac{1}{M}\sum_{l=1}^{M}I(\tilde{\gamma}_{0(l)}>\tilde{\gamma}). \tag{9}$$

(iii) For testing $H_0^*:\gamma=0$ vs. $H_a^*:\gamma\neq 0$, define $\tilde{\gamma}_L=\tilde{\gamma}_{0((\alpha/2)M)}$ and $\tilde{\gamma}_U=\tilde{\gamma}_{0((1-\alpha/2)M)}$. If $\tilde{\gamma}$, is either greater than $\tilde{\gamma}_U$ or less than $\tilde{\gamma}_L$, then H_0^* is rejected. Alternatively the p-value can be defined as

$$p=2\min(p_1,p_2). \tag{10}$$

Table 1: Sizes of tests for reliability at 5% significant level when $\nu = 2$.

		$(\sigma_1^2,\ \sigma_2^2)$			
(n_1, n_2)	Method	(0.4, 0.2)	(1.5, 0.5)	(2.4, 0.2)	(3.5, 0.5)
(5, 5)	Score test	0.141	0.119	0.115	0.087
	CAT	0.058	0.057	0.052	0.051
(10, 10)	Score test	0.083	0.063	0.056	0.052
	CAT	0.069	0.058	0.044	0.043
(20, 10)	Score test	0.098	0.094	0.068	0.072
	CAT	0.069	0.065	0.063	0.062
(10, 20)	Score test	0.095	0.075	0.065	0.057
	CAT	0.048	0.054	0.048	0.057
(20, 20)	Score test	0.058	0.053	0.046	0.058
	CAT	0.045	0.043	0.044	0.045
(30, 35)	Score test	0.076	0.054	0.055	0.054
	CAT	0.055	0.047	0.049	0.041
(50, 50)	Score test	0.054	0.061	0.057	0.053
	CAT	0.056	0.054	0.052	0.046
(100, 50)	Score test	0.057	0.061	0.054	0.058
	CAT	0.045	0.052	0.051	0.051
(50, 100)	Score test	0.061	0.055	0.054	0.053
	CAT	0.053	0.051	0.052	0.054

3 Simulation Study

A simulation study is performed to evaluate the actual size in comparison to the score test. We note that the data are generated such that $\nu = 2$. For this propose, random samples with size n_i, $i = 1, 2$, are generated from log-normal population with parameters μ_i and σ_i^2, $i = 1, 2$. For the CAT, we consider $M = 10000$ and compute the p-value based on given steps in section 2. For the score test also we obtain the p-value. We note that $\mu_2 = 0$ is considered in all cases. We repeat this system for $N = 10000$ and computing the p-values for the two approaches. The actual size of each test calculated using the number cases that the p-values are smaller than nominal level, $\alpha = 0.05$. Numerical results on the actual sizes are given in Table 1, and we can conclude that: i) the actual sizes of score test are close to nominal level when the sample sizes are large and are very liberal when the sample sizes are small. ii) the actual sizes of the CAT are close to the nominal level in all cases and are smaller than the actual sizes of the score test.

References

[1] Cheng, S. S. (1977). Optimal replacement rate of devices with lognormal failure distribution. *IEEE Trans. Reliability* **R-26**, 174-178.

[2] Crow, E. L. and Shimizu, K. (1988). *Lognormal Distributions*, Marcel Dekker: New York.

[3] Gill, P. S. (2004). Small-sample inference for the comparison of means of log-normal distributions, *Biometrics*, **60**, 525-527.

[4] Gupta, R. C. and Li, X. (2006). Statistical inference for the common mean of two log-normal distributions and some applications in reliability, *Computational Statistics and Data Analysis*, **50**, 3141-3164.

[5] Krishnamoorthy, K. and Mathew, T. (2003). Inferences on the means of lognormal distributions using generalized p-values and generalized confidence intervals, *Journal of statistical planning and inference*, **115**, 103-121.

[6] Lawless, J. F. (1982). *Statistical Models and Methods for Lifetime Data*, New York, John Wiley.

[7] Pal, N., Lim, W. K. and Ling, C. H. (2007). A computational approach to statistical inferences, *Journal of Applied Probability and Statistics.*, **2**, 13-35.

[8] Zhou, X. H., Gao, S., and Hui, S. L. (1997). Methods for comparing the means of two independent log-normal samples, *Biometrics*, **53**, 1129-1135.

[9] Zhou, X. H. and Tu, W. (1999). Comparison of Several Independent Population Means when their Samples Contain Lognormal and Possibly Zero Observations, *Biometrics*, **55**, 645-651.

Email:k.abdollahnezhad@gu.ac.ir

Email:aajafari@yazduni.ac.ir

43rd Annual Iranian Mathematics Conference,
27-30 August 2012
University of Tabriz

Minimum χ^2-divergence joint probability density function given prior density function and moments

Solmaz Seifollahi Kosehlar
University of Tabriz

Hossein Bevrani
University of Tabriz

Abstract
One of applied discussion in statistics is the estimation ofprobability density function. One of the issues relating to this subject in recent years is estimate probability density function through initial density function and information about moments. In this paper, we introduce minimum chi-square divergence principle for bivariate manner, in addition to we denote the application of minimum chi-square divergence for determination joint density function given prior density function and retail information about moments. In final consequences are consider in detail, followed by a numerical illustration.
Keywords: Bivariate chi-square divergence, Probability density function, Exponential distribution, Moment

Mathematics Subject Classification: 62G07

1 Introduction

One of applied discussion in statistics is the estimation of probability density function . One of the issues relating to this subject in recent years is estimate probability density function through initial density function and information about moments. For this purpose, expect estimation procedures like maximum likelihood and least square error, maximum entropy (Kagan et al. [4], Kapur [5], Kawamura and Iwas [6], Kesavan and Kapur [7]) are used for estimating of density function.

Guiasu [3] has analyzed the weighted deviations subject to given mean value of random variable and determined the probability distribution by considering measure of deviations such as the Pearson's chi-square [10], Neyman's reduced Chi-square [2], Kullback-Leibler [8], and Kolmogorov's index [1]. Kumar [9] has considered the minimum Chi-square divergence principle and the information about moments of the random variable to determine the best probability distribution, then he estimated density function by using Gamma distribution.

The aim of this paper is the generalization of Kumar's method for the estimation of joint probability density function. Section 2 presents the definition of minimum bivariate chi-square divergence and the minimum bivariate chi-square divergence principle. Minimum chi-square divergence joint probability distributions by using Exponential distribution and product moment are determine in section 3. Section 4 provide the concluding remarks.

2 Minimum Chi-square Divergence Joint Density Function

Chi-square divergence between $f(x)$ and $g(x)$ defined as Pearson [10]:

$$\chi^2(f,g) = \int \frac{f^2(x)}{g(x)}\, dx - 1.$$

Now, we define bivariate Chi-square divergence between $f(x, y)$ and $g(x, y)$ as [11]:

$$\chi^2(f, g) = \int_X \int_Y \frac{f^2(x, y)}{g(x, y)}\, dxdy - 1. \tag{1}$$

We consider the minimum bivariate chi-square divergence principle as follows:
When a prior density function for (X, Y), i.e $g(x, y)$, which estimates the underlying probability density function $f(x, y)$ is given in addition to some constrains, when among all the density function $f(x, y)$ which satisfy the given constraints, we should select that probability density function which minimizes the bivariate χ^2-divergence between $f(x, y)$ and $g(x, y)$.

Definition 2.1. *If $h_t(x, y); t : 1, 2, \cdots, r$ is an arbitrary function of (X, Y), $f(x, y)$ is minimum chi-square divergence joint probability density function for (X, Y) if it minimize the chi-square divergence given*

1. $f(x, y) \geq 0$,
2. $\int_X \int_Y f(x, y)\, dxdy = 1$,
3. $\int_X \int_Y h_t(x, y) f(x, y)\, dxdy = m_{h_t,f}; \quad t : 1, 2, \cdots, r.$

By using following lemma and above definition, $f(x, y)$ is obtained.

Lemma 2.2. *Given a prior density function $g(x, y)$ for (X, Y) and the constraints expressed in Definition 1.1 the minimum bivariate chi-square divergence probability density function for (X, Y) has the probability density function*

$$f(x, y) = \frac{g(x, y)}{2}(\alpha_0 + \sum_{t=1}^{r} \alpha_t h_t(x, y)) \tag{2}$$

and the $(r + 1)$ constants $\alpha_0, \alpha_1, \cdots, \alpha_r$ are determined from

$$\int_X \int_Y \frac{g(x, y)}{2}(\alpha_0 + \sum_{t=1}^{r} \alpha_t h_t(x, y))\, dxdy = 1 \tag{3}$$

$$\int_X \int_Y \frac{h_t(x, y) g(x, y)}{2}(\alpha_0 + \sum_{t=1}^{r} \alpha_t h_t(x, y))\, dxdy = m_{h_t,f}; \quad t : 1, 2, \cdots, r \tag{4}$$

Proof. We apply the familiar method of finding extreme of afunction by introducing lagrangian multipliers, one for each constraint.Thus we minimize the lagrangian function

$$L = (\int_X \int_Y \frac{f^2(x, y)}{g(x, y)}\, dxdy - 1) - \alpha_0(\int_X \int_Y f(x, y)\, dxdy - 1)$$

$$-\sum_{t=1}^{r} \alpha_t(\int_X \int_Y h_t(x, y))\, dxdy - m_{h_t,f})$$

Minimizing L respect to $f(x, y)$, i.e differentiating the integrand respect to $f(x, y)$ and setting the derivative equal to zero, we get the following equation

$$\frac{2f(x, y)}{g(x, y)} - \alpha_0 - \sum_{t=1}^{r} \alpha_t h_t(x, y) = 0 \tag{5}$$

solving (1.3), we get the desired expression for $f(x, y)$. The $(r + 1)$ constants $\alpha_0, \alpha_1, \cdots, \alpha_r$ are determined from the $(r + 1)$ equation (1.3) and (1.3). □

The minimum χ^2-divergence measure is

$$\chi^2_{min}(f,g) = \int_X \int_Y \frac{g(x,y)}{4}(\alpha_0 + \sum_{t=1}^{r} \alpha_t h_t(x,y))^2 \, dxdy - 1. \tag{6}$$

Now, we consider only information about moments is $E(XY)$, then:

Theorem 2.3. *If $g(x,y)$ is prior density function for (X,Y) and $f(x,y)$ is satisfied in*

$$f(x,y) \geq 0\,, \int_X \int_Y f(x,y)\, dxdy = 1\,, \int_X \int_Y xyf(x,y)\, dxdy = m_{xy,f}$$

the minimum chi-square divergence probability density function is

$$f(x,y) = \frac{g(x,y)}{2}(\alpha_0 + \alpha_1 xy) \tag{7}$$

that α_0, α_1 is obtained from

$$\int_X \int_Y \frac{g(x,y)}{2}(\alpha_0 + \alpha_1 xy)\, dxdy = 1, \tag{8}$$

$$\int_X \int_Y \frac{xyg(x,y)}{2}(\alpha_0 + \alpha_1 xy)\, dxdy = m_{xy,f}. \tag{9}$$

3 The Joint Density Function Given Product of Two Exponential Distribution and $E(XY)$

We consider prior density function is product of two Exponential distribution, then

Theorem 3.1. *Given a prior density function for (X,Y) with the density function*

$$g(x,y) = \frac{1}{\lambda_1} e^{-\frac{x}{\lambda_1}} \frac{1}{\lambda_2} e^{-\frac{y}{\lambda_2}}; \quad x,y \geq 0\,,\, \lambda_1, \lambda_2 > 0$$

and the constraints

$$f(x,y) \geq 0\,, \int_0^\infty \int_0^\infty f(x,y)\, dxdy = 1\,, \int_0^\infty \int_0^\infty xyf(x,y)\, dxdy = m$$

then, the minimum chi-square divergence probability density function for (X,Y) has the probability density function

$$f(x,y) = \frac{1}{\lambda_1\lambda_2} e^{-\frac{x}{\lambda_1} - \frac{y}{\lambda_2}} (\frac{4\lambda_1^2\lambda_2^2 - m\lambda_1\lambda_2 + (m - \lambda_1\lambda_2)xy}{3\lambda_1^2\lambda_2^2}) \tag{10}$$

for $m \in [\lambda_1\lambda_2, 4\lambda_1\lambda_2]$. The minimum χ^2-divergence measure is

$$\chi^2_{min}(f,g) = \frac{(m - \lambda_1\lambda_2)^2}{3\lambda_1^2\lambda_2^2}$$

and

$$0 \leq \chi^2_{min}(f,g) \leq 3.$$

Proof. This theorem can be proved by using Theorem 3.1. □

Corollary 3.2. *If $m = \lambda_1\lambda_2$, then $f(x,y) = g(x,y)$ and the minimum χ^2-divergence measure is equal zero.*

Corollary 3.3. *If $m = 4\lambda_1\lambda_2$, then*

$$f(x,y) = \frac{x}{\lambda_1}e^{-\frac{x}{\lambda_1}}\frac{y}{\lambda_2}e^{-\frac{y}{\lambda_2}}.$$

The minimum χ^2-divergence measure is three.

Corollary 3.4. *For $m \in (\lambda_1\lambda_2, 4\lambda_1\lambda_2)$, the probability density function with minimum χ^2-divergence between $f(x,y)$ and $g(x,y)$ is as given in (3). For example, if $m = \frac{5}{2}\lambda_1\lambda_2$, then*

$$f(x,y) = \frac{1}{2}e^{-\frac{x}{\lambda_1}-\frac{y}{\lambda_2}}[\frac{\lambda_1\lambda_2 + xy}{\lambda_1^2\lambda_2^2}].$$

The minimum χ^2-divergence measure is equal 0.75.

4 Numerical Illustration

The minimum chi-square two Exponential divergence joint probability density function, having product of distribution with parameters $(\lambda_1, \lambda_2) = \{(1,1),(1,\frac{1}{2})\}$ and$E(XY)$ are presented in below table. The probability density function $f(x,y)$ with minimum chi-square divergence between $f(x,y)$ and $g(x,y)$ are given in column 3 and the corresponding value of chi-square divergence measure is presented in column 4. For instance, consider prior density function as $g(x,y) = e^{-x-y}$, the value of $E(XY)$ should be between 1 and 4. For $E(XY) = 1$ The minimum chi-square divergence density function is $f(x,y) = e^{-x-y}$ with $\chi^2_{min}(f,g) = 0$. If $E(XY) = \frac{5}{2}$, then $f(x,y) = \frac{1}{2}(1+xy)e^{-x-y}$ with $\chi^2_{min}(f,g) = \frac{3}{4}$.

λ_1, λ_2	$E(XY)$	$f(x,y)$	χ^2_{min}
	1	e^{-x-y}	0
$1, 1$	$\frac{5}{2}$	$\frac{1}{2}(1+xy)e^{-x-y}$	$\frac{3}{4}$
	4	xye^{-x-y}	3
	$\frac{1}{2}$	$2e^{-x-2y}$	0
$1, \frac{1}{2}$	$\frac{5}{4}$	$(1+2xy)e^{-x-2y}$	$\frac{3}{4}$
	2	$2xye^{-x-2y}$	3

Table 1: The joint density function given product of two Exponential distribution and $E(XY)$

References

[1] J. Abrahams, *On the selection of measures of distance between probability distributions*, Inform. Sci., 26 (1982), pp. 109-113.

[2] J. Berkson, *Minimum chi-square, not maximum likelihood (with discussion)*, Ann. Statist., 4 (1980), pp. 457-487.

[3] S. Guiasu, *The least weighted deviation*, Inform. Sc., 53 (1991), pp. 271-284.

[4] A. M. Kagan, V. J. Linnik and C. R. Rao, *Characterization Problems in Mathematical statistics*, John Wiley, New York, 1975.

[5] J. N. Kapur, *Twenty five years of maximum entropy principle*, Jour. Math. And Phys. Sc., 17 (2) (1982), pp. 103-456.

[6] T. Kawamura and K. Iwase, *Characterizations of the distributions of power inverse Gaussian and others based on the entropy maximization principle*, Jour. Japan Statist. Soe., 33 (1) (2003), pp. 95-104.

[7] H. K. Kesavan and J. N. Kapur, *The generalized maximum entropy principle*, IEEE Trans. Systems, Man and Cybernetics, 19 (5) (1989), pp. 1042-1052.

[8] S. Kullback, *Information Theory and Statistics*, New York, John Wiley, 1959.

[9] P. Kumar, *Probability Distributions Conditioned by the Available Information: Gamma Distribution and Moments*, Jour.Computes and Mathematics with applications, 52 (2006), pp. 289-304.

[10] K. Pearson, *On the criterion that a given system of deviations from the probable in the case of correlated system of variables is such that it can be reasonable supposed to have arisen from random sampling*, Phil. Mag, 50 (1900), pp. 157-172.

[11] S. Seifollahi Kosehlar, *Minimum chi-square distance probability density function given prior density function and moments*, MSC thesis, Iran, 2011.

Email:s.seifollahi.k@gmail.com

Email:bevrani@tabrizu.ac.ir

43rd Annual Iranian Mathematics Conference,
27-30 August 2012
University of Tabriz

Estimation of $P(Y < X)$ for Lindley distribution in the presence of one outlier

Parvin Fathipour
Univesity of Tabriz

Ali Abolhasani
Azarbaijan University of Tarbiat Moallem

Maryam Abolhasani
University of Mohaghegh Ardabili

Abstract
In this paper, we consider the problem of estimating $R = P(Y < X)$, where Y has lindley distribution with parameter a and X has lindley distribution with presence of one outlier with parameters b and c, such that X and Y are independent. the maximum likelihood estimator of R is derived and some results of simulation studies are presented.
Keywords: Lindley Distribution, Maximum Likelihood Estimator, Newton-Raphson Method, Outlier

Mathematics Subject Classification: 62F10

1 Introduction

In reliability context inferences about $R = P(Y < X)$, when X and Y are independently distributed, are a subject of interest. for example in mechanical reliability of a system if X is the strength of a component which is subject to stress Y, then R is a measure of system performance. the system fails, if at any time the applied stress is greater than its strength. Stress-strength reliability has been discussed in Kapur and Lamberson (1977). Sathe and Dixit (2001) have done estimate of R in the negative binomial distribution. Baklizi and Dayyeh (2003) have done shrinkage estimation of R in exponential case, and recently Deiri (2011) has done estimation of R with presence of two outliers in the exponential and gamma caces, respectively.
Dixit and Nasiri (2001) considered estimation of parameters of the exponential distribution in the presence of outliers generated from uniform distribution.Jafari (2011) obtain the moment, maximum likelihood and mixture estimators of R in Rayleigh distribution in the presence of one outlier. In this paper, we obtain the maximum likelihood estimator of R in Lindley distribution with presence of one outlier generated from the same distribution.
The probability density function of the lindley distribution with parameter of a is given by:

$$f(y; a) = \frac{a^2}{1+a}(1+y)e^{-ay}, \quad x > 0, a > 0$$

In this paper we assume that the random variables $(Y_1, Y_2, ..., Y_m)$ have lindley distribution with parameter a and the random variables $(X_1, X_2, ..., X_n)$ are such that one of them is distributed as lindley distribution with parameter c and the remaining $n-1$ random variables are distributed as lindley with parameter b.

2 Joint Distribution of $X_1, X_2, ..., X_n$ in the Presence of Outlier

Assume $X_1, X_2, ..., X_n$ are such that one of them is distributed with p.d.f $g(x, c)$ as lindley(c) and remaining (n-1) of them are distributed with p.d.f $f(x, b)$ as lindley(b). the joint distribution of $X_1, X_2, ..., X_n$ can be expressed as

$$\begin{aligned} f(x_1, x_2, ..., x_n; b, c) &= \frac{(n-1)!}{n!} \prod_{i=1}^{n} f(x_i; b) \sum_{A_1=1}^{n} \frac{g(x_{A_1}; c)}{f(x_{A_1}; b)} \\ &= \frac{(n-1)!}{n!} \frac{b^{2n}}{(1+b)^n} \prod_{i=1}^{n} (1+x_i) e^{-b\sum_{i=1}^{n} x_i} \sum_{A_1=1}^{n} \frac{\frac{c^2}{1+c}(1+x_{A_1})e^{-cx_{A_1}}}{\frac{b^2}{1+b}(1+x_{A_1})e^{-bx_{A_1}}} \\ &= \frac{(n-1)!}{n!} \frac{b^{2n-2}}{(1+b)^{n-1}} \frac{c^2}{1+c} \prod_{i=1}^{n} (1+x_i) e^{-b\sum_{i=1}^{n} x_i} \sum_{A_1=1}^{n} (1+x_{A_1}) e^{x_{A_1}(b-c)} \end{aligned} \tag{1}$$

see Dixit (1989), Dixit and Nasiri (2001), and Nasiri and Pazira (2009).from (1), the marginal distribution of X is

$$f(x; b, c) = \frac{1}{n} \frac{c^2}{1+c}(1+x)e^{-cx} + \frac{n-1}{n} \frac{b^2}{1+b}(1+x)e^{-bx}; \quad x > 0, b > 0, c > 0. \tag{2}$$

3 Maximum Likelihood Estimators of Parameters

Let $Y_1, Y_2, ..., Y_m$ be a random sample for Y with pdf,

$$f(y; a) = \frac{a^2}{1+a}(1+y)e^{-ay}, \quad x > 0, a > 0$$

the log likelihood function is given by

$$L(a) = 2mlna - mln(1+a) + \sum_{i=1}^{m} ln(1+y_i) - a\sum_{i=1}^{m} y_i$$

taking the derivative with respect to a and equating to 0, we obtain the MLE of a as

$$\hat{a} = \frac{m - \sum_{i=1}^{m} y_i \pm \sqrt{(\sum_{i=1}^{m} y_i - m)^2 + 8m\sum_{i=1}^{m} y_i}}{2\sum_{i=1}^{m} y_i} \tag{3}$$

now let $X_1, X_2, ..., X_n$ be a random sample for X with presence of one outlier with pdf,

$$f(x; b, c) = \frac{1}{n} \frac{c^2}{1+c}(1+x)e^{-cx} + \frac{n-1}{n} \frac{b^2}{1+b}(1+x)e^{-bx}; \quad x > 0, b > 0, c > 0$$

from (1), the log likelihood function is given by

$$\begin{aligned} L(b, c) = {} & ln(\frac{(n-1)!}{n!}) + (2n-2)lnb - (n-1)ln(1+b) + 2lnc - ln(1+c) \\ & + \sum_{i=1}^{n} ln(1+x_i) - b\sum_{i=1}^{n} x_i + ln(\sum_{A_1=1}^{n} e^{x_{A_1}(b-c)}) \end{aligned}$$

taking the derivatives with respect to b and c and equating the results to 0, we obtain the normal equations as

$$\frac{\partial L(b,c)}{\partial b} = \frac{2n-2}{b} - \frac{n-1}{1+b} - \sum_{i=1}^{n} x_i + \frac{\sum_{A_1=1}^{n} x_{A_1} e^{x_{A_1}(b-c)}}{\sum_{A_1=1}^{n} e^{x_{A_1}(b-c)}} \tag{4}$$

$$\frac{\partial L(b,c)}{\partial c} = \frac{2}{c} - \frac{1}{1+c} - \frac{\sum_{A_1=1}^{n} x_{A_1} e^{x_{A_1}(b-c)}}{\sum_{A_1=1}^{n} e^{x_{A_1}(b-c)}} \tag{5}$$

there is no closed-form solution to this system of equations, so we will solve for $\hat{b}$ and $\hat{c}$ iteratively, using the Newton-Raphson method, a tangent method for root finding. in our case we will estimate $\hat{\beta} = (\hat{b}, \hat{c})$ iteratively:

$$\hat{\beta}_{i+1} = \hat{\beta}_i - G^{-1}\mathbf{g} \tag{6}$$

where g is the vector of normal equations for which we want

$$\mathbf{g} = \begin{bmatrix} g_1 & g_2 \end{bmatrix}$$

with

$$g_1 = \frac{2n-2}{b} - \frac{n-1}{1+b} - \sum_{i=1}^{n} x_i + \frac{\sum_{A_1=1}^{n} x_{A_1} e^{x_{A_1}(b-c)}}{\sum_{A_1=1}^{n} e^{x_{A_1}(b-c)}}$$

$$g_2 = \frac{2}{c} - \frac{1}{1+c} - \frac{\sum_{A_1=1}^{n} x_{A_1} e^{x_{A_1}(b-c)}}{\sum_{A_1=1}^{n} e^{x_{A_1}(b-c)}}$$

and G is the matrix of second derivatives

$$G = \begin{bmatrix} \frac{dg_1}{db} & \frac{dg_1}{dc} \\ \frac{dg_2}{db} & \frac{dg_2}{dc} \end{bmatrix}$$

The Newton-Raphson algorithm converges, as our estimate of b and c change by less than a tolerated amount with each successive iteration, to $\hat{b}$ and $\hat{c}$.

4 The Maximum Likelihood Estimator of R

Let $Y \sim lindley(a)$ with pdf $h(y;a)$ and X be distributed with pdf $f(x;b,c)$ given in (2). the parameter R we want to estimate is

$$\begin{aligned} R &= P(Y < X) \\ &= \int_0^\infty \int_0^x h(y;a) f(x;b,c)\,dydx \\ &= \frac{1}{n}\int_0^\infty \int_0^x \frac{a^2}{1+a}(1+y)e^{-ay}\frac{c^2}{1+c}(1+x)e^{-cx}dydx \\ &+ \frac{n-1}{n}\int_0^\infty \int_0^x \frac{a^2}{1+a}(1+y)e^{-ay}\frac{b^2}{1+b}(1+x)e^{-bx}dydx \\ &= \frac{1}{n}[\frac{c^2(c(1+c)+(1+c)(3+c)a+(3+2c)a^2+a^3)}{(1+c)(1+a)(c+a)^3}] \\ &+ \frac{n-1}{n}[\frac{b^2(b(1+b)+(1+b)(3+b)a+(3+2b)a^2+a^3)}{(1+b)(1+a)(b+a)^3}] \end{aligned} \tag{7}$$

thus the MLE of R can be obtained from

$$\hat{R} = \frac{1}{n}[\frac{\hat{c}^2(\hat{c}(1+\hat{c}) + (1+\hat{c})(3+\hat{c})\hat{a} + (3+2\hat{c})\hat{a}^2 + \hat{a}^3)}{(1+\hat{c})(1+\hat{a})(\hat{c}+\hat{a})^3}] + \frac{n-1}{n}[\frac{\hat{b}^2(\hat{b}(1+\hat{b}) + (1+\hat{b})(3+\hat{b})\hat{a} + (3+2\hat{b})\hat{a}^2 + \hat{a}^3)}{(1+\hat{b})(1+\hat{a})(\hat{b}+\hat{a})^3}] \quad (8)$$

where $\hat{a}$, $\hat{b}$ and $\hat{c}$ can be obtained from (3) and (6).

5 Conclusion

According to the result of simulation, the values of biases and MSEs are often around 0 when the parameter value of outlier (c) is close to that of the rest $(n-1)$,X variables (b) and they increase when the difference between parameters b and c becomes greater than 1.

References

[1] Deiri, E. 2011a, Estimation of P(Y<X) for Exponential Distribution in the Presence of Two Outliers, *International Journal of Academic Research*, 3, 508-514.

[2] Deiri, E. 2011b, Estimation of P(Y<X) for Generalized Exponential Distribution in the Presence of Two Outliers when Scale Parameters is Known, *International Journal of Academic Research*, 3, 1179-1185.

[3] Deiri, E. 2011c. Estimation of Parameters of the Gamma Distribution in the Presence of Two Outliers, *Australian Journal of Basic and Applied Science*, 3, 846-852.

[4] Dixit, U.J. 1989, Estimation of Parameters of the Gamma Distribution in the Presence of Outliers, *Communications in Statistics - Theory and Methods*, 18, 3071-3085.

[5] Dixit, U.J., Moor, K.L. and Barnett,V. 1996, On the Estimation of the Power of the Scale Parameter of the Exponential Distribution in the Presence of Outlier generated from uniform distribution, *Metron*, 54, 201-211.

[6] Dixit, U.J. and Nasiri, P.F. 2001, Estimation of Parameters of the Exponential Distribution with Presence of Outliers Generated from Uniform Distribution, *Metron*, 49(3-4), 187-198.

[7] Jafari, A. 2011, Estimation of P(Y<X) in the Rayleigh Distribution with Presence of Outlier,Accepted in International Journal of Academic Research.

[8] Nasiri, P.F. and Pazira, H. 2010, Bayesian and Non-Bayesian Estimations on the Generalized Exponential Distribution in the Presence of Outliers, *Journal of Statistical Theory and Practice*, 4(3), 453-475.

Email:p.fathipoor@yahoo.com

43rd Annual Iranian Mathematics Conference,
27-30 August 2012
University of Tabriz

A random variable distributed between two random variables

M. Hamel Darbandi
University of Tabriz

A. Hamel Darbandi
Ankara University

R. Salimpoor
K.N. Toosi University

Abstract

Let X and Y be independent random variables and let Z be a random variable(which is uniform or not uniformly distributed) over [X,Y]. We study the distribution of the random variable Z and show that the arcsin distribution and cauchy distribution can be characterized in a particular way by means of this construction.

Keywords: Arcsin, Cauchy, Stieltjes Transform, Schwartz theory.

Mathematics Subject Classification: 60G50

1 Introduction

In a fundamental paper, Van Assche (1987) considered a random variable Z^* uniformly distributed between two independent random variables X and Y; in the sense that the conditional distribution of Z^* given $X = x$ and $Y = y$ is uniform over $[\min\{x, y\}, \max\{x, y\}]$. Mathematically this means that

$$P(Z^* \leq z \mid X = x, Y = y) = \begin{cases} 1, & z \geq \max(y, x) \\ \frac{z-x}{y-x}, & x < z < y \\ \frac{z-y}{x-y}, & y < z < x \\ 0, & z \leq \min(y, x), \end{cases}$$

By applying certain properties of the distributional derivatives, Van Assche (1987) derived the following interesting results. Assuming X and Y are identically distributed, then

Result (i):Under the assumption that X, Y are independent and continuous with distributions F_X, F_Y and F_Z^* is distribution of Z^*, relation between stieltjes transforms distributions of F_Z^*, F_X and F_Y is given that

$$-\mathcal{S}'(F_Z^*, z) = \mathcal{S}(F_X, z)\mathcal{S}(F_Y, z)$$

Result (ii): for X and Y on $[-1, 1]$, Z^* is uniform on $[-1, 1]$ if and only if X and Y have an arcsin distribution;

Result (iii): Z^* possesses the same distribution as X and Y if and only if X and Y are degenerated or have a Cauchy distribution.

Kotz and Johnson have cleverly done a further study of Van Assche's work.But they did not succeed to get the intended result(2). It also should be mentioned that the powerful method of Van Assche for calculation of distribution compared to Johnson and Kotz [1] method is more flexible. In recent years, several papers are presented on weighted average (see Johnson and kotz [1], Soltani-homei [5]). By using random division Soltani and Homei extended Van Assch -Johnson and kotz's work and extended weighted average. But it seemed their result hold only for the case of n=2 and even for getting the cauchy character one should solve an equation of order n that is very difficult. The special case of n=2, is equivalent to choosing a sample from uniform distribution and getting $(W,1-W)$. It seems that uniform distribution in Van Assche's paper is important. Main aim in paper is to show the result given Van Assche is not just for uniform distribution a lone.

In Section 2, we introduced the distribution similar construction of random variable Z^* uniformly distributed between two independent random variables X and Y. In Section 3, we establish the 2-th derivative of the Stieltjes transform of the distribution of Z is expressed in terms of the product of the Stieltjes transforms of the distributions one and derivative of other, Theorem 3.1.In Section 4, As, we observe that (Van Assche) two characterization for the Cauchy and arcsin distribution are true.

2 A Random Variable Distributed Between Two Independent Random Variables X And Y

Suppose that X and Y are two independent random variables (which are not necessarily identically distributed). We suppose that Z is a random variable between X and Y with conditional distribution, given $X=x,,Y=y$;

$$F_{Z|x,\,y}(z)=\begin{cases} 1, & z\geq \max(y,x)\\ (\frac{z-x}{y-x})^2, & x<z<y\\ -2(\frac{z-y}{y-x})-(\frac{z-y}{y-x})^2, & y<z<x\\ 0, & z\leq \min(y,x),\end{cases}$$

The conditional distribution of Z, given $X=x$ and $Y=y$, can be written in a more compact form using the Heaviside function:

$$U(x)=\begin{cases} 0, & x<0\\ 1, & x\geq 0.\end{cases}$$

namely

$$P(Z\leq z\mid X=x,Y=y)=(\frac{z-x}{y-x})^2U(z-x)+2(\frac{z-y}{x-y})U(z-y)-(\frac{z-y}{x-y})^2U(z-y). \tag{1}$$

The distribution function of Z can then be written down by integrating the X and Y out, giving

$$P(Z\leq z)=\int_{\mathbb{R}^2}(\frac{z-x}{y-x})^2U(z-x)+2(\frac{z-y}{x-y})U(z-y)-(\frac{z-y}{x-y})^2U(z-y)dF_XdF_Y. \tag{2}$$

Where, the distribution function of Z,X and Y, respectively are F_Z, F_X and F_Y.

3 Relation Between Stieltjes Transforms

Let us denote the Stieltjes transform of a distribution K by

$$\mathcal{S}(K,z)=\int_{\mathbb{R}}\frac{1}{z-x}K(dx), \tag{3}$$

for every z in the set of complex numbers $\mathbb{C}$ which does not belong to the support of K, $z\in\mathbb{C}\cap(\text{supp}K)^c$. For more on the Stieltjes transform see [7]. Let us also barrow the following tools from the Schwartz theory on distributional derivatives. Indeed every distribution Λ is viewed as a

Schwartz distribution on a certain infinitely differentiable functions φ through $\int_{-\infty}^{\infty}\varphi(x)\Lambda(dx)$. It is well know that Λ admits a derivative of order 2, denoted by $\Lambda^{[2]}$, possessing

$$\int_{-\infty}^{\infty}\varphi(x)\Lambda^{[2]}(dx)=\frac{1}{2}\int_{-\infty}^{\infty}\frac{d^2}{dx^2}\varphi(x)\Lambda(dx). \tag{4}$$

A distribution $\Lambda^{[2]}$ that possesses (4) is called the 2-th distributional derivative of Λ.

In this section we present relation between stieltjes transforms distributions of F_Z, F_X and F_Y.

Theorem 3.1. *Under the assumption that X, Y are independent and continuous, relation between stieltjes transforms distributions is given that*

$$-\frac{1}{2}\quad \mathcal{S}^{''}(F_Z,z)=\mathcal{S}(F_X,z)\mathcal{S}^{'}(F_Y,z) \tag{5}$$

Where $\mathcal{S}^{''}(F_Z,z)$ and $\mathcal{S}^{'}(F_Y,z)$ are second and first derivative of $\mathcal{S}(.,z)$ on z.

Proof. The conditional distribution $P(Z\le z\mid X=x,Y=y)$ given by (1) leads us to the following linear functional on complex-valued functions , defined on the set of real numbers $\mathbb{R}$;

$$P(Z\le z\mid X=x,Y=y)=(\frac{z-x}{y-x})^2U(z-x)+2(\frac{z-y}{x-y})U(z-y)-(\frac{z-y}{x-y})^2U(z-y) \tag{6}$$

It easily follows that

$$P(Z\le z\mid X=x,Y=y)=\frac{f_z(x)}{(y-x)^2}+\frac{\frac{df_z(y)}{dz}}{x-y}-\frac{f_z(y)}{(x-y)^2} \tag{7}$$

where $f_z(x)=(z-x)^2U(z-x)$. Also we note that

$$U(z-x)=\frac{1}{2}\quad\frac{d^2}{dx^2}f_z(x).$$

Thus

$$P(Z\le z)=\int_{\mathbb{R}}U(z-x)dF_Z(x)=\int_{\mathbb{R}^2}(\frac{f_z(x)}{(y-x)^2}+\frac{\frac{df_z(y)}{dz}}{x-y}-\frac{f_z(y)}{(x-y)^2})\,dF_XdF_Y$$

can be viewed as:

$$\frac{1}{2}\int_{\mathbb{R}}\frac{d^2}{dx^2}f_z(x)dF_Z(x)=\int_{\mathbb{R}^2}(\frac{f_z(x)}{(y-x)^2}+\frac{\frac{df_z(y)}{dz}}{x-y}-\frac{f_z(y)}{(x-y)^2})dF_XdF_Y \tag{8}$$

Now (8) together with (4) lead us to

$$\int_{\mathbb{R}}f_z(x)dF_Z^{[2]}(x)=\int_{\mathbb{R}^2}(\frac{f_z(x)}{(y-x)^2}+\frac{\frac{df_z(y)}{dz}}{x-y}-\frac{f_z(y)}{(x-y)^2})\,dF_XdF_Y \tag{9}$$

where $F_Z^{[2]}$ is the (2)-th distributional derivative of the distribution F_Z.

Therefore by using the linear property (6) and a standard argument in the integration theory, we obtain that

$$\int_{\mathbb{R}}f(x)dF_Z^{[2]}(x)=\int_{\mathbb{R}^2}(\frac{f(x)}{(y-x)^2}+\frac{\frac{df(y)}{dz}}{x-y}-\frac{f(y)}{(x-y)^2})dF_XdF_Y \tag{10}$$

for a suitable f.

We follows from (10) that, for $f(x) = \frac{1}{z-x}$

$$\begin{aligned}\int_{\mathbb{R}} f(x)dF_Z^{[2]}(x) &= \int_{\mathbb{R}^2} \frac{\frac{1}{z-x}}{(y-x)^2} + \frac{-\frac{1}{(z-y)^2}}{x-y} - \frac{\frac{1}{z-y}}{(x-y)^2}\, dF_X dF_Y \\ &= \int_{\mathbb{R}^2} \frac{\frac{1}{(y-x)^2}}{z-x} + \frac{-\frac{1}{x-y}}{(z-y)^2} - \frac{\frac{1}{(x-y)^2}}{z-y}\, dF_X dF_Y \\ &= \int_{\mathbb{R}^2} \frac{1}{z-x}\ \frac{1}{(z-y)^2}\, dF_X dF_Y,\end{aligned}$$

Thus

$$-\mathcal{S}(F_Z^{[2]}, z) = \mathcal{S}(F_X, z)\mathcal{S}'(F_Y, z) \qquad z \in \mathbb{C} \bigcap (\mathrm{supp} F_X \cap \mathrm{supp} F_Y)^c \tag{11}$$

Therefore

$$\begin{aligned}\frac{d^2}{dz^2}\,\mathcal{S}(F_Z, z) &= \int_{\mathbb{R}} \frac{2}{(z-x)^3}\, F_Z(dx) \\ &= \int_{\mathbb{R}} \frac{d^2}{dx^2} \frac{1}{z-x}\, F_Z(dx) \\ &= 2 \int_{\mathbb{R}} \frac{1}{z-x}\, F_Z^{[2]}(dx) \\ &= 2\,\mathcal{S}(F_Z^{[2]}, z) \\ &= -2\,\mathcal{S}(F_X, z)\,\mathcal{S}'(F_Y, z).\end{aligned}$$

giving the result. The proof of the theorem is complete.

□

4 Characterization

Now we are to review Van Assche's findings.

Theorem 4.1. *Let X and Y be i.i.d random variables on [-1,1], then Z is uniformly distributed on [-1,1] if and only if X and Y have an arcsin distribution on [-1,1], i.e,*

$$P(Z \leq t) = P(Y \leq t) = \frac{1}{\pi} \int_{-1}^{t} \frac{1}{\sqrt{1-x^2}} dx$$

Proof. The random variable Z has a uniform distribution on [-1,1] and F is the distribution function of X and Y, then it follows from the theorem 3.1

$$\begin{aligned}\mathcal{S}(F_X, z)\mathcal{S}'(F_Y, z) &= -\frac{1}{2}\mathcal{S}''(F_Z, z) \\ &= -\frac{z}{(z^2-1)^2}\end{aligned}$$

The solution $\mathcal{S}(F_X, z)$ is

$$\mathcal{S}(F_X, z) = \frac{1}{\sqrt{z^2-1}}$$

Which is the stieltjes transform of the arcsin distribution . □

Theorem 4.2. *Let X and Y be i.i.d. random variables, then Z has the same distribution as X and Y if and only if X and Y are almost surely constant or have a Cauchy distribution.*

Proof. The random variable Z has the same distribution as X and Y, so that we have to find the Idempotent probability measures for F, here F is the probability distribution function of X and Y. Then it follows from the theorem 3.1,

$$-\frac{1}{2} \quad \mathcal{S}^{''}(F,z) = \mathcal{S}(F,z)\mathcal{S}^{'}(F,z)$$

By using a Reduction Order in ODE method the solution is obtain as

$$\mathcal{S}(F,z) = \frac{1}{z+c}$$

Where c is a constant.In order that this is the stieltjes transform of a probability distribution, we need to have that $\text{Im}\mathcal{S}(F,z) < 0$ whenever $\text{Im}(z) > 0$, $b \neq 0$, Therefore

$$\mathcal{S}(F,z) = \frac{1}{z-a+ib} \quad \text{Im}(z) > 0, \quad b \neq 0,$$

Where a is real and $b \geq 0$. the case $b = 0$ corresponds to $F(x) = U(x-a)$ whence X and Y are almost surely constant. When $b > 0$ we have

$$F(x) = \frac{1}{\pi}\int_{-\infty}^{x} \frac{b}{b^2+(t-a)^2}dt$$

Which is the Cauchy distribution with center parameter a and spread parameter b. □

References

[1] N. L. Johnson and S. Kotz, *Randomly weighted averages: Some aspects and extensions*, Amer. Statist., 44 (1990), pp. 245–249.

[2] W. Rudin, *Real and Complex Analysis*, McGraw and Hill, 1987.

[3] W. Rudin, *Functional Analysis*, McGraw and Hill, 1991.

[4] A. R. Soltani and H. Homei, *A generalization for two-sided power distributions and adjusted method of moments*, Statistics, 43 (2009), pp. 611–620.

[5] A. R. Soltani and H. Homei, *Weighted averages with random proportions that are jointly uniformly distributed over the unit simplex*, Statistics and Probability Letters, 9 (2009), pp. 1215–1218.

[6] W. Van Assche, *A random variable uniformly distributed between two independent random variables*, Sankhya Ser. A, 49 (1987), pp. 207–211.

[7] A. I'Zayed, *Handbook of Function and Generalized Function Transformations*, CRC Press, London, 1996.

Email:darbandi89@ms.tabrizu.ac.ir

Email:azizdarbandi@gmail.com

Stieltjes transform

H. Homei

University of Tabriz

Abstract

The concept of a randomly weighted averages introduced recently with random proportions that are jointly uniformly distributed over the unit simplex. In this article we study some generalization this concept and results. Our study on conditional distribution of randomly weighted averages also leads us to a new rich class of extension triangular distribution. We also find characterization for Cauchy, arcsin, semicircle and power semicircle distribution.

Keywords: Arcsin distribution, Stieltjes transform, Schwartz theory on distributional derivatives, uniform Distribution, Cauchy distribution.

Mathematics Subject Classification: 42A38

1 Introduction

Many articles have been distributed about random mixed variables sofar. Mix model has been exerted in many diciplines such as sociology and biology and other sciences. This model is being used in total probability, Bayes's theorem and posterior distributions. We focus on special family of distribution.

In a fundamental paper, Van Assche(1987) considered a random variable S uniformly distributed between two independent random variables X and Y, in the sense that the conditional distribution of S given X=x and Y=y is uniform over [min{x,y},max{x,y}]. He employed the Stieltjes transform and derived that: (i) for X and Y on [-1,1], S is uniform on [-1,1] if and only if X and Y have an arcsin distribution;(ii)S possesses the same distribution as X and Y if and only if X and Y are degenerated or have a Cauchy distribution. The work of Van Assche was commented on by Johnson and Kotz(1990), whereby a simple method derived the arcsin result(i), but acknowledged that their approach was not useful to obtain(ii). Johnson and Kotz noticed that the random variable S is indeed a weighted average of X and Y with random proportions U and $1-U$; $S = UX + (1-U)Y$, U uniform [0,1] independent of X and Y. Also Johnson and Kotz extended the randomly weighted average of more than two random variables ($\sum W_i X_i$, where W is Dirichlet distribution). Jonson and Kotz initialized a discussion but did not finish it. In fact they neither proved nor refused Van Assche's findings. Soltani and Homei (2009) followed Johnson and kotz and investigated Van Assche's findings. They let $X_1, \cdots, X_n$ to be independent, and considered

$$S_n = R_1X_1 + R_2X_2 + \cdots + R_{n-1}X_{n-1} + R_nX_n, \quad n \geq 2,$$

where random proportions are $R_i = U_{(i)} - U_{(i-1)}, i = 1, \cdots, n-1$ and $R_n = 1 - \sum_{i=1}^{n-1} R_i$, $U_{(1)}, \cdots, U_{(n-1)}$ order statistics from a uniform distribution on [0,1], $U_0 = 0$. These random proportions are said to be uniformly distributed over the unit simplex, and were introduced and studied by Dempster and kleyle(1968).

They employed Stieltjes transform and derived that (i) $\frac{(-1)^{n-1}}{(n-1)!}\frac{d^{n-1}}{dz^{n-1}}\mathcal{S}(F_{S_n}, z) = \prod_{i=1}^{n} \mathcal{S}(F_{X_i}, z)$;

(ii) S_n possesses the same distribution as X_1,...,X_n if and only if X_1,...,X_n are degenerated or have a Cauchy distribution;(iii) The arcsin's result of Van Assche (1987) is only true for S_2.
In this work, we extended the work of Soltani and Homei concept and result.
We let X_1,...,X_n to be independent, and consider

$$S_{n^*} = R_1X_1 + R_2X_2 + \cdots + R_{n-1}X_{n-1} + R_nX_n, \quad n \geq 2,$$

where random proportions are $R_i = U_{k_i} - U_{k_{i-1}}, i = 1, \cdots, n-1$.
In section 2, we derive the distribution of S_{n^*}, for given distinct values $X_1 = x_1, \cdots, X_n = x_n$. In section 3, we establish the main result of this article where the $(n^*-1)-th$ derivative of the Stieltjes transform of the distribution of S_{n^*} is expressed in terms of the product of the Stieltjes transforms of the distributions of $X_1, \cdots, X_n$, Theorem 3.1. In section 4, we review Van Assche-Soltani and Homei's findings.

2 Conditional distribution

In this section we provide the conditional distribution of $S_{n^*} = \sum_{j=1}^n R_{k_j}X_j$ for given $(X_1, ..., X_n) = (x_1, ..., x_n)$ at z, denoted by $k(z|x_1, ..., x_n)$. At first we assume $x_1 > x_2 > ... > x_n > 0$, but later we will remove this restriction on $x_1, ..., x_n$. We recall that $(u_{(1)}, ..., u_{(n)})$ is the order statistics of a random sample $u_1, ..., u_n$ for the uniform $[0, 1]$. For the sequence of indices $\{k_1, ..., k_{n-1}\}$ is an ordered sequence in $\{1, ..., n^* - 1\}$. Thus $\{u_{k_1}, ..., u_{k_{n-1}}\} \subset \{u_{(1)}, ..., u_{(n^*-1)}\}$ the increments R_{k_j} are defined by $R_{k_j} = u_{k_j} - u_{k_{j-1}}, j = 1, ..., n-1$, where $u_{k_0} = 0$ and $u_{k_n} = 1$. Since $\sum_{j=1}^n R_{k_j} = 1$, $k(z|x_1, ..., x_n)$ can be expressed as

$$P(\sum_{j=1}^{n-1} c_jR_{k_j} \leq z - x_n); \quad c_j = x_j - x_n, \, j = 1, ..., n-1 \, , \tag{1}$$

the distribution $\sum_{j=1}^{n-1} c_jR_{k_j}$ was derived by Weisberg (1971):

$$P(\sum_{j=1}^{n-1} c_jR_{k_j} \leq z - x_n) = 1 - \sum_{j=1}^{r} \frac{h_j^{m_j-1}(x_j; z)}{(m_j - 1)!}, \tag{2}$$

where $m_j = k_j - k_{j-1}, \; j = 1, ..., n, \; k_n = n^*, \; \sum_{j=1}^n m_j = n^*, \; h_j^{(m_j-1)}(c_j)$ is the $(m_j - 1)-$th derivative of

$$h_j(x; z) = \frac{(x - z)^{n^*-1}}{c_j \prod_{i \neq j}^n (x - x_i)^{m_i-1}} \tag{3}$$

at x evaluated at $x_1, x_2, ..., x_n$, where r is the largest positive integer such that $z < x_r$.
The distribution of $\sum_{j=1}^{n-1} c_jR_{k_j}$ in (2), alternatively can be expressed as

$$P(\sum_{j=1}^{n-1} c_jR_{k_j} \leq z - x_n) = \sum_{j=r^*+1}^{n} \frac{f_j^{(m_j-1)}(x_j; z)}{(m_j - 1)!} \, , \tag{4}$$

where r^* is the largest positive integer such that $x_{r^*} \geq z$, and $f_j^{(m_j-1)}(x_j; z)$ is the $(m_j - 1)-$th derivatives of

$$f_j(x; z) = \frac{(x - z)^{n^*-1}}{\prod_{i \neq j}^n (x - x_i)^{m_i}} \, . \tag{5}$$

at $x = x_j$.
By using the heaviside function

$$U(x) = \begin{cases} 0 & x < 0 \\ 1 & x > 0, \end{cases} \tag{6}$$

the distribution in (4) can be expressed as

$$k(z \mid x_1, ..., x_n) = \sum_{j=1}^{n} \frac{f_j^{(m_j-1)}(x_j; z)U(z - x_j)}{(m_j - 1)!} \ , \tag{7}$$

for any set of distinct values $x_1, ..., x_n$, and any $z \in [min(x_1, ..., x_n), max(x_1, ..., x_n)]$. The conditional

$$k(z \mid x_1, ..., x_n) = \begin{cases} 0 & z < min(x_1, ..., x_n) \\ 1 & z \geq min(x_1, ..., x_n), \end{cases} \tag{8}$$

$k(z; x_1, ..., x_n)$ is a big family of distribution that includes two-sided power (TSP) distributions and general two-sided power (GTSP) distributions.
If $k_1 = 1, k_2 = n, x_1 = 1, x_2 = \theta, x_3 = \cdots = x_n = 0$, then (1) readily reduces to the STSP distribution (see more [11]) and let $x_1 = b, x_2 = (b - a)\theta + a, x_3 = a, x_4 = \cdots = x_n = 0$ then (1) readily reduces to the GTSP distribution ([8]). In the following $k_j = j$, new family distribution include triangular distribution is introduced.

$$F(x|\ \theta_1, \cdots, \theta_n) = \begin{cases} \sum_{j=0}^{i} c_j (x - \theta_{n-j})^{n-1}, \\ \theta_{n-i} < x \leq \theta_{n-i-1},\ i = 0, \cdots, n-2, \end{cases} \tag{9}$$

where

$$c_j = \left[\prod_{i=1, i \neq n-j}^{n} (\theta_i - \theta_{n-j}) \right]^{-1},\ j = 0, \cdots, n-1.$$

Xwill be said to follow a N-sided power distribution, $NSP(\theta_1, \cdots, \theta_n), \theta_1 > \theta_2 > \cdots > \theta_n, n > 0$, where n is an integer. The density of (9) is unimodal with the mode at $\bar{\theta}$.

For $n = 3$, $F(. \mid \theta_1, \theta_2, \theta_3)$ simplifies to a triangular distribution. The density function of a $NSP(\theta_1, \theta_2, \theta_3)$ distribution follows from expression (9) as

$$f(x|\ \theta_1, \theta_2, \theta_3) = \begin{cases} \frac{2(x-\theta_3)}{\theta_{13}\ \theta_{23}}, & \theta_3 < x \leq \theta_2 \\ \\ \frac{2(\theta_1 - x)}{\theta_{13}\ \theta_{12}}, & \theta_2 < x \leq \theta_1\ . \end{cases} \tag{10}$$

$f(z|\theta_1, \theta_2, \theta_3)$ is triangular distribution. Figure.1 provides examples of symmetric $NSP(\theta_1, \cdots, \theta_n)$ distributions, i.e $\bar{\theta} = 0.5$, including, triangular distribution. Figure.2 presents examples of positively and negatively skewed $NSP(\theta_1, \cdots, \theta_n)$ distributions, including examples of triangular distributions.

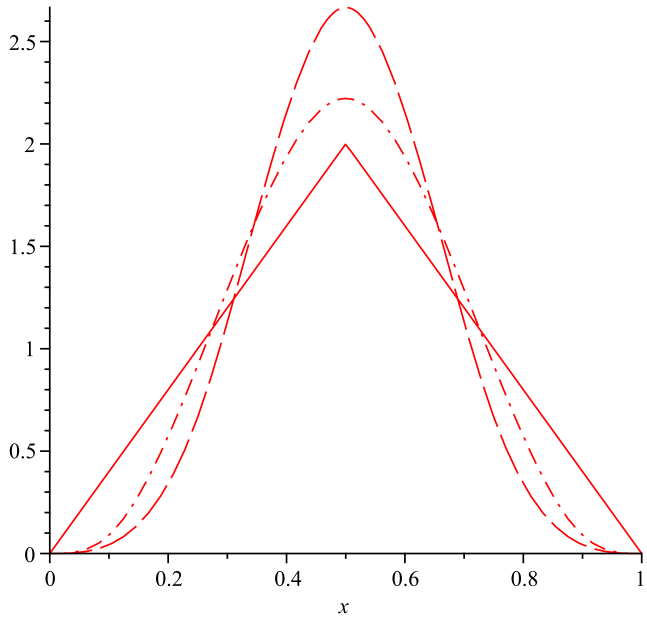

Figure1:Symmetric NSP(θ1,θ2,θ3,θ4,θ5) distribution:—·—·— , θ1=1,θ2=0.9 θ3=0.5,θ4=0.1,θ5=0 ; ——— , θ1=1,θ2=0.75,θ3=0.5,θ4=0.25,θ5=0

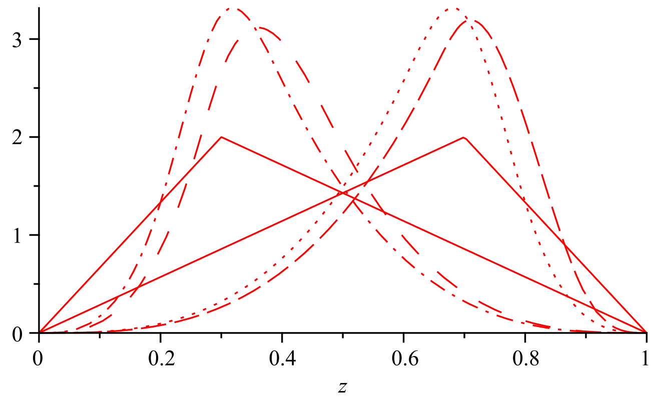

Figure 2:Positively skewed NSP(θ1,θ2,θ3,θ4,θ5) distributions: ···· , θ1=1,θ2=0.8,θ3=0.7,θ4=0.6,θ5=0; — — — , θ1=1,θ2=0.9 ,θ3=0.7 ,θ 4=0.65,θ5=0 and negatively skewed NSP(θ1,θ2,θ3, θ4,θ5) distrbutions:-·-·-· ,θ1=1,θ2=0.4,θ3=0.3,θ4=0.2,θ5=0;- - - , θ1=1, θ2=0.5,θ3=0.3,θ4=0.25,θ5=0

The expressions for the mean and the variance can be obtained from expression (9) and simplify to

$$E(X) = \bar{\theta} \,, \tag{11}$$

$$Var(X) = \frac{S_\theta^2}{n+1} \tag{12}$$

The meaning of the parameters is as follows:θ_1, θ_n are the end points of the parameter , n is the shape parameter and $\theta_2, \cdots, \theta_{n-1}$ are the threshold parameters for a change in the form of

the probability density function. However, N-sided power distribution is a novel extension of the triangular distribution. We can used triangular distributions as a proxy to the Beta distribution, specifically in problems of assessment of risk and uncertainty, such as the project evaluation and review technique (PERT).
The expression for conditional distribution $k(z \mid x_1, ..., x_n)$ will play an essential role in deriving the $(n^* - 1) - th$ derivative of the Stieltjes transform of the distribution S_{n^*} that will be established in the following section.

3 The Stieltjes transform for the distribution of S_{n^*}

Let us first present the following lemma.

Lemma 3.1. *For given distinct real $x_1, ..., x_n$, z and integers $m_i \geq 1$, $i = 1, ..., n$, the following*

$$\sum_{j=1}^{n} \frac{(-1)^{m_j}}{(m_j - 1)!} [\frac{d^{m_j - 1}}{dx_j^{m_j - 1}} \frac{1}{(z - x_j)} \frac{1}{\prod_{i \neq j}^{n} (x_i - x_j)^{m_i}}] = \prod_{j=1}^{n} \frac{1}{(x_j - z)^{m_j}} \ . \tag{13}$$

The conditional kernel presented in (7) leads us to a fairly large class of conditional kernel. Indeed we define

$$k(g|x_1, ..., x_n) = \sum_{j=1}^{n} \frac{(-1)^{m_j - 1}}{(m_j - 1)!} \frac{d^{m_j - 1}}{dx_j^{m_j - 1}} \frac{g(x_j)}{\prod_{i \neq j} (x_i - x_j)^{m_i}} \ , \tag{15}$$

for any function $g : \mathbb{R} \to \mathbb{C}$, the kernel given in (15) wil be reduced to (7) if $g_z(x) = (z - x)^{n^* - 1} U(z - x)$.
Following, Van Assche (1987), we will also apply the concept of the distribution function, then it possess a unique distributional derivative $\Lambda^{(n)}$ for a given integer n, such that

$$\int_{-\infty}^{+\infty} \varphi(x) \Lambda^{(n)}(dx) = \frac{(-1)^n}{n!} \int_{-\infty}^{+\infty} \left[\frac{d^n}{dx^n} \varphi(x) \right] \Lambda(dx). \tag{16}$$

For certain infinitely differentiable function φ, the following lemma provide a useful integral relation between the conditional kernel (15) and the distribution derivative of F the distribution of the random mixture S_{n^*}.

Lemma 3.2. *The $(n^* - 1)$-th distributional derive of the distribution F and the conditional kernel (15) are subject to*

$$\int_{\mathbb{R}} g(x) dF^{(n^* - 1)}(x) = \int_{\mathbb{R}^n} k(g|x_1, ..., x_n) \prod_{i=1}^{n} F_{X_i}(dx_i), \tag{17}$$

for any infinitely differential function g for which the integrals are finite.

Theorem 3.3. *Under the assumptions that $X_1, ..., X_n$ are independent and continuous,*

$$\frac{(-1)^{n^* - 1}}{(n^* - 1)!} \frac{d^{n^* - 1}}{dz^{n^* - 1}} S(F, z) = \prod_{i=1}^{n} \frac{(-1)^{m_i - 1}}{(m_i - 1)!} \frac{d^{m_i - 1}}{dz^{m_i - 1}} S(F_{X_i}, z), \qquad z \in \mathbb{C} \bigcap_{i=1}^{n} (suppF_{X_i})^c.$$

4 Characterization

4.1 Van Assche's findings and it's generalization with Beta random proportion

Let Z be randomly weighted average with random proportions U and $1-U$; $Z = UX_1+(1-U)X_2$, U $Beta(m_1, m_2)$ independent of X_1 and X_2, then

$$-Beta(m_1, m_2)\ \mathcal{S}^{(m_1+m_2-1)}(F_Z, z) = \mathcal{S}^{(m_1-1)}(F_{X_1}, z)\mathcal{S}^{(m_2-1)}(F_{X2}, z)\ , \quad z \in \mathbb{C}\bigcap_{i=1}^{2}(\text{supp}F_{X_i})^c. \tag{21}$$

Now we let $m_1 = 1, m_2 = 1$, we have

$$-\mathcal{S}^{'}(F_Z, z) = \mathcal{S}(F_{X_1}, z)\mathcal{S}(F_{X_2}, z).$$

Now we are to review Van Assche's findings for some beta distributions, so the results given by Van Assche is not only for uniform distribution.

Theorem 4.1. *Assume Z is randomly weighted average with Beta(2,1) random proportion and X_1, X_2 are random variables with a common distribution function F, then Z has distribution F if and only if X_1 and X_2 are almost surely constant or have a Cauchy distribution.*

Theorem 4.2. *Let X_1 and X_2 be i.i.d random variables on [-1,1] and Z be randomly weighted average with Beta(2,1) random proportion, then Z is uniformly distributed on [-1,1] if and only if X_1 and X_2 have an arcsin distribution on [-1,1].*

Theorem 4.3. *Let X_1 and X_2 be independent random variables and Z be randomly weighted average with Beta(2,1) random proportion. Then if X_1 has uniform distribution on [-1,1], then X_2 has arcsin distribution on [-1,1] if and only if Z has semicircle distribution on [-1,1].*

Theorem 4.4. *Let X_1 and X_2 be i.i.d. random variables on [-1,1] and Z be randomly weighted average with Beta(2,2) random proportion. Then X_1 and X_2 have a uniform distribution on [-1,1], if and only if Z has power semicircle distribution, i.e,*

$$f(z) = \frac{3(1-z^2)}{4}\ , \quad -1 \le z \le 1.$$

4.2 Soltani and Homei's findings

By using theorem 3.3 and $m_1 = m_2 = \cdots = m_n = 1$, we have

$$\frac{(-1)^{n-1}}{(n-1)!}\frac{d^{n-1}}{dz^{n-1}}\mathcal{S}(F, z) = \prod_{i=1}^{n}\mathcal{S}(F_{X_i}, z), \quad z \in \mathbb{C}\bigcap_{i=1}^{n}(\text{supp}F_{X_i})^c. \tag{22}$$

This relation is given by Soltani and Homei. We will give examples of derivation based on (22).

Theorem 4.5. *Let X_1,X_2 and X_3 be i.i.d random variables on [-1,1], then S_3 has semicircle distribution on [-1,1] if and only if X_1,X_2 and X_3 have arcsin distribution.*

Theorem 4.6. *Let X_1,X_2 and X_3 be i.i.d random variables on [-1,1], then X_1,X_2 and X_3 have semicircle distribution on [-1,1] if and only if S_3 has power semicircle distribution on [-1,1],i.e,*

$$f(z) = \frac{16}{5\pi}(1-z^2)^{\frac{5}{2}}\ , \quad -1 \le z \le 1.$$

Theorem 4.7. *Let X_1,X_2 and X_3 be independent random variables on [-1,1] and X_1,X_2 have semicircle distribution on [-1,1], then X_3 has arcsin distribution if and only if S_3 has power semicircle distribution on [-1,1],i.e,*

$$f(z) = \frac{8}{3\pi}(1-z^2)^{\frac{3}{2}}\ , \quad -1 \leq z \leq 1.$$

Theorem 4.8. *Let X_1,X_2,X_3 and X_4 be i.i.d random variables on [-1,1], then X_1,X_2,X_3 and X_4 have arcsin distribution on [-1,1] if and only if S_4 has power semicircle distribution on [-1,1],i.e ,*

$$f(z) = \frac{3(1-z^2)}{4}\ , \quad -1 \leq z \leq 1.$$

Theorem 4.9. *Let X_1,X_2,X_3 and X_4 be i.i.d random variables on [-1,1], then X_1,X_2,X_3 and X_4 have semicircle distribution on [-1,1] if and only if S_4 has power semicircle distribution on [-1,1], i.e ,*

$$f(z) = \frac{128}{35\pi}\,(1-z^2)^{\frac{7}{2}}\ , \quad -1 \leq z \leq 1.$$

Theorem 4.10. *Let X_1,X_2,X_3,X_4 be independent random variables on [-1,1], X_1,X_2,X_3 have semicircle distribution on [-1,1], then X_4 has arcsin distribution if and only if S_4 has power semicircle distribution on [-1,1], i.e ,*

$$f(z) = \frac{16}{5\pi}(1-z^2)^{\frac{5}{2}}\ , \quad -1 \leq z \leq 1.$$

4.3 Some characterization

In this section, we also observe application of theorem 3.3, as:

Theorem 4.11. *Let $m_1 = 3, m_2 = 1, m_3 = 1$ and X_1,X_2 and X_3 be independent random variables on [-1,1]. Then if X_2, X_3 have arcsin distribution, then X_3 has semicircle distribution on $[-1,1]$ if and only if S_3 has power semicircle distribution on $[-1,1]$,i.e ,*

$$f(z) = \frac{8}{3\pi}(1-z^2)^{\frac{3}{2}}\ , \quad -1 \leq z \leq 1.$$

Theorem 4.12. *Let $m_1 = 1, m_2 = 1, m_3 = 2$ and X_1,X_2 and X_3 be independent random variables on [-1,1], then X_1,X_2 and X_3 have arcsin distribution on [-1,1] if and only if S_3 has a semicircle distribution on [-1,1].*

References

[1] A. P. Dempster and R. M. Kleyle, *Distribution by cutting a simplex with hyperplanes*, Ann. Math. Statist., 39 (1968), pp. 1473–1478.

[2] M. C. Diniz, E. de Souza e Silva and H. R. Gail, Calculating the distribution of a linear combination of uniform order statistics, Informs J. Computing, 14 (2002), pp. 0124–0131.

[3] N. L. Johnson and S. Kotz, *Randomly weighted averages: Some aspects and extensions*, Amer. Statist., 44 (1990), pp. 245–249.

[4] O. Arizmendi and V. Perez-Abreu, *On the non-classical infinite divisibility of power semicircle distributions*, Communications on Stochastic Analysis, Vol. 4, No. 2 (2010), pp. 161–178.

[5] A. R. Soltani and H. Homei, *Weighted averages with random proportions that are jointly uniformly distributed over the unit simplex*, Statistics and Probability Letters, 9 (2009), pp. 1215–1218.

[6] W. Van Assche, *A random variable uniformly distributed between two independent random variables*, Sankhya Ser. A, 49 (1987), pp. 207–211.

[7] J. R. Van Dorp and S. Kotz, *The standard Two Sided Power distribution and its Properties: with Applications in Financial Engineering*, The American Statistician, 56, No. 2 (2002), pp. 90–99.

[8] H. Weisberg, *The distribution of linear combinations of order statistics from the uniform distribution*, Ann. Math. Statist., 42 (1971), pp. 704–709.

Email:homei@tabrizu.ac.ir

43rd Annual Iranian Mathematics Conference,
27-30 August 2012
University of Tabriz

Transformation method for estimating $P(X < Y)$ in the case of three parameter generalized Rayleigh distribution

Hossein Jabbari Khamnei
University of Tabriz

Parvin Fathipour
University of Tabriz

Abstract

In this paper, we will use transformation method to find MLE of $P(X < Y)$, where X and Y are independent random variables from three parameter generalized rayleigh distribution. The results of simulation shows that the mean of errors, by using transformation method, is often around zero.

Keywords: Stress-strength model, Transformation method, Three parameter generalized rayleigh distribution, Incomplete Beta function

Mathematics Subject Classification: 62N02, 62N05, 62F10, 90B25

1 Introduction

In the last four decades, estimation of $R = P(X < Y)$ is very common between statisticians. Tong (1974) discusses on estimation of R in exponential case, also had a look at estimation of R for exponential familie (1977). Basu (1981), has obtained the MLE for R when the distribution of variables are gamma and exponential. The MLE for R, in the case of bivariate exponential distribution have been considered by Awad et al (1981). Johnson (1988) had a good review on estimation of R for exponential families. Dinh et al (1991) discusses on the UMVUE of R in bivariate normal case. Cramer and Kamps (1997), considered the UMVUE of R when the distribution of X and Y are exponential. Ahmad et al (1977) have studied about empirical Bayes estimation of R for Burr-type X distribution. Surles and Padgett (1998) have made inference for R when the distribution of variables are scaled Burr-type X. Comparison of different estimations of R for a scaled Burr-type X, have been made by Raqab and Kundu (2005).

2 Main Result

As Kotz et al (2003) states in chapter 2, one can use transformation method to find point and interval estimation for R. Transformation method have been overlooked by statisticians, but rarely applied to stress-strength models before. Suppose that random vector (X, Y) has the pdf $f(x, y|\theta)$ where θ is a scalar or vector-value parameter with considering that there exist random variable ξ and η and a monotone function $u(.)$ with the inverse $\upsilon = u^{-1}$ such that

$$X = u(\xi) \Leftrightarrow \xi = \upsilon(X), Y = u(\eta) \Leftrightarrow \eta = \upsilon(Y)$$

Assume that the function u and υ are strictly increasing, then (ξ, η) is a random vector with the pdf

$$g^*(\xi, \eta|\theta) = f(u(\xi), u(\eta)|\theta)u'(\xi)u'(\eta) \quad (1)$$

Parametrization of (1) can be carried out in other way. Let (ξ, η) be a random vector with pdf $g(\xi, \eta|\tau)$ where the scalar or vector-value parameter τ is connected to θ by one-to-one transformation ϱ with the inverse ν

$$\theta = \nu(\tau) \Leftrightarrow \tau = \varrho(\theta)$$

Thus, there exists the following correspondence between $f(x, y|\theta)$ and $g(\xi, \eta|\tau)$

$$\begin{array}{c} g(\xi, \eta|\tau) = f(u(\xi), u(\eta)|\nu(\tau))u'(\xi)u'(\eta) \\ f(x, y|\theta) = g(\upsilon(x), \upsilon(y)|\varrho(\theta))\upsilon'(x)\upsilon'(y) \end{array}$$

Since function u is monotonically increasing, then, $P(\xi < \eta) = P(u(\xi) < u(\eta)) = P(X < Y)$. So, to find the MLE for $P(X < Y)$, it is just enough to find MLE for $P(\xi < \eta)$.

To find MLE for R in the case of three parameter generalized rayleigh distribution, firs we transform it into gamma distribution then using the relationship between gamma and incomplete beta distribution, we estimate R. The advantage of using this method is shown by simulation.

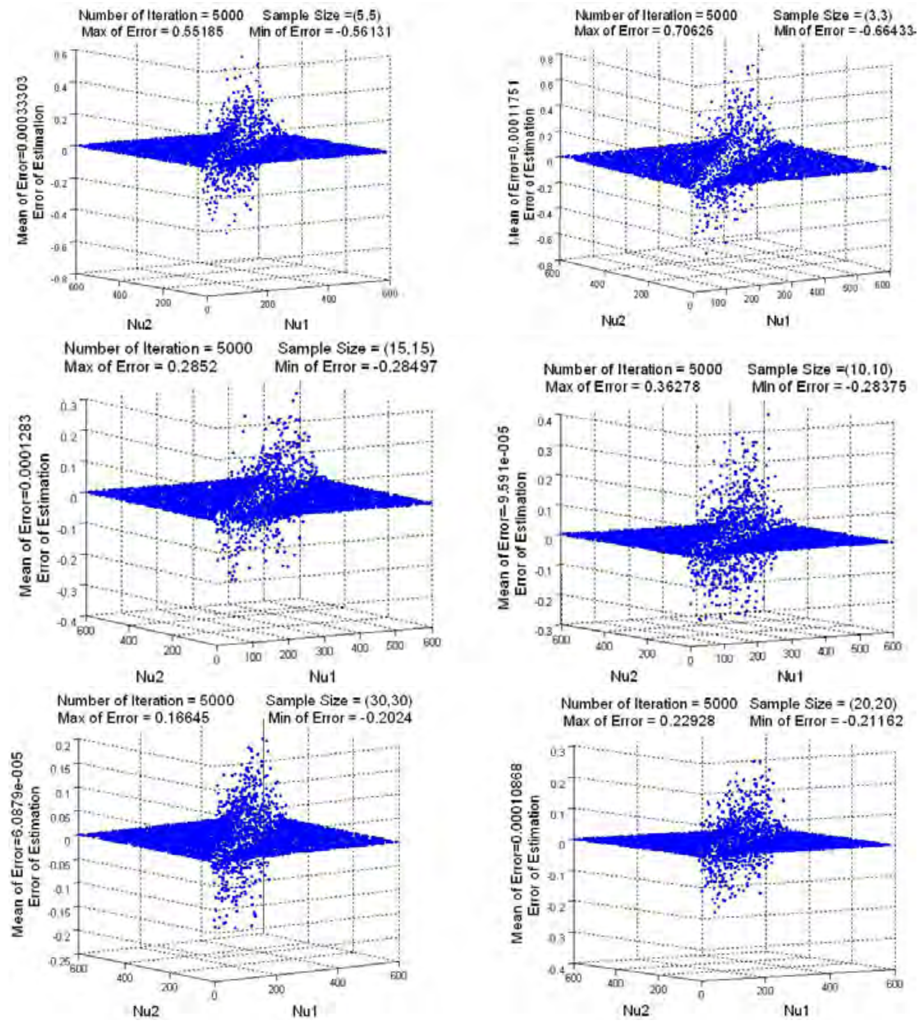

Figure 1: Sample sizes is considered (3,3), (5,5), (15,15), (20,20), (30,30) and iteration number is 5000. In all graphs error of estimation is around zero and by increasing sample size, the range of error of estimation decreases.

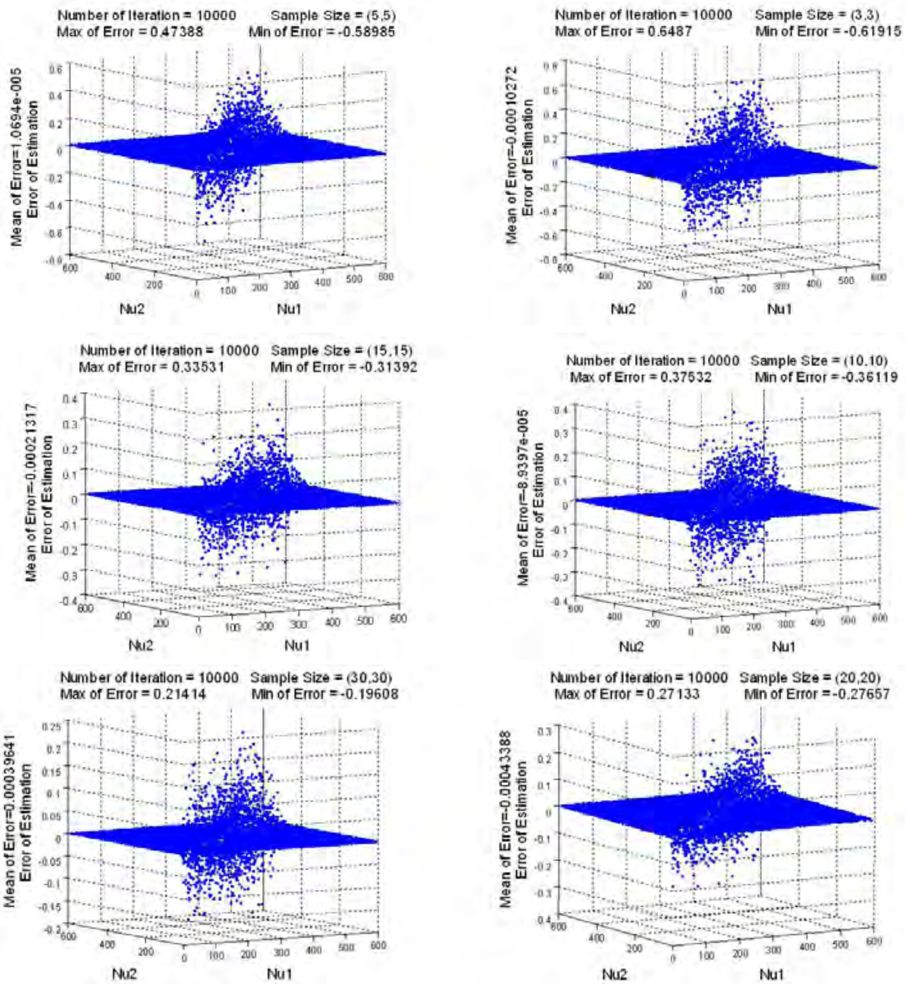

Figure 2: Sample sizes is considered (3,3), (5,5), (15,15), (20,20), (30,30) and iteration number is 10000. In all graphs error of estimation is around zero and by increasing sample size, the range of error of estimation decreases.

References

[1] Ahmad, K. E., Fakhry, M. E. and Jaheen, Z. F. (1997), Empirical Bayes Estimation of $P(X < Y)$ and Characterization of the Burr-type X Model, *Journal of Statistical Planning and Inference*, **64**, 297-308.

[2] Awad, A. M., Azzam, M. M. and Hamadan, M. A. (1981), Some Inference Results in $P(Y < X)$ in the Bivariate Exponential Model, *Communication in Statistics Theory and Methods*, **10**, 2515-2524.

[3] Basu, A. (1981), The Estimation $P(X > Y)$ for Distributions Useful in Life Testing, *Naval Research Logistics Quarterly*, **3**, 383-392.

[4] Dinh, K. T., Singh, J. and Gupta, R. C. (1991), Estimation of Reliability in Bivariate Distributions, *Statistics*, **22(3)**, 409-417.

[5] Johnson, R. A. (1988), *Stress-Strength Models for Reliability*, Handbook of Statistics, 7 (Kirshnaiah, P.R. and Rao, C.R.), Amsterdam, North Holland.

[6] Kotz, S. and Lumelskii, Y. and Pensky, M. (2003), *The Stress-Strength Model and generalizations, Theory and Applications*, World Scientific Publishing Co. Pte. Ltd. Singapore.

[7] Surlse, J. G. and Padgett, W. J. (2001), Inference for Reliability and Stress-Strength for a Scaled Burr-type X Distribution, *Lifetime Data Analysis*, **7**, 187-200.

[8] Tong, H. (1977), On the Estimation of $P(X < Y)$ for Exponential Families, *IEEE Trans. on. Reliability*, **R-26**, 54-56.

[9] Tong, H. (1977), A Note on the Estimation for $P(X < Y)$ in the Exponential Case, *Techno Metrics*, **16**, 625.

Email:h_jabbari@tabrizu.ac.ir

43^{rd} Annual Iranian Mathematics Conference,
27-30 August 2012
University of Tabriz

A computational approach test in two power law distributions

A. A. Jafari
Yazd University

K. Abdollahnezhad
Golestan University

F. Yeghmai
Golestan University

S. Mahmoudi
Golestan University

Abstract
In this article, we apply a computational approach test (CAT) based on maximum likelihood estimator, introduced by Pal et al. (2007),in power law distribution. The CAT does not require the knowledge of any sampling distribution, depends heavily on numerical computations and Monte-Carlo simulation. We apply the CAT for testing the equality of scale parameters in two power law populations. Simulation studies show that the actual size and power of this method is satisfactory. At end, an example with simulation data is given.
Keywords: Power distribution, Hypothesis test, Actual size, computational approach test

Mathematics Subject Classification: 62F04, 62F99, 62F03, 62E17

1 Introduction

Advances in the technology of computational tools have affected the significantly statistical inference and estimation. Complex theoretical results can now be better realized through numerical computations and/or Monte-Carlo simulations well before that can be verified analytically. Recently, Pal et al. (2007) developed a simple computational technique, called CAT, for hypothesis testing problems. This method uses the MLE for the purpose of statistical inference on unknown parameters. The CAT is a simple procedure based on a simple set of computational steps which can be implemented easily by applied researchers. The computational mechanism is such that the CAT finds the critical region automatically. Chang and Pal (2008a) showed that for the one-way ANOVA problem with the usual assumptions, the CAT provides power which is very close to that of the classical F-test. Also Chang and Pal (2008b) applied the proposed CAT for testing the common mean of several normal distributions.The CAT does not require the knowledge of any sampling distribution, depends heavily on numerical computations and Monte-Carlo simulation. Hence, the conventional approach had been to use either the asymptotic theory, or use some sort of approximations to the null distributions of the test statistics.

Let us consider the power law density function with parameters θ and β (pow(θ,β))

$$f(x;\theta,\beta)=\frac{\beta x^{\beta-1}}{\theta^{\beta}},\quad 0<x<\theta,\quad \beta>0, \tag{1}$$

where θ is the scale parameter and β is the shape parameter. The theoretical mean and variance are derived as:

$$E(X)=\frac{\theta\beta}{\beta+1},\quad Var(X)=\frac{\theta^2\beta}{(\beta+2)(\beta+1)^2}. \tag{2}$$

The MLE of the parameters θ and β are

$$\hat{\theta} = \max\{X_i\},\ i = 1, 2, .., n, \qquad \hat{\beta} = \frac{1}{\log(\hat{\theta}) - \frac{1}{n}\sum_{i=1}^{n} \log(X_i)}, \tag{3}$$

and

$$\frac{2n\beta}{\hat{\beta}} \sim \chi^2_{(2n-2)}, \qquad -2n\beta\left[\log(\hat{\theta}) - \log(\theta)\right] \sim \chi^2_{(2)}. \tag{4}$$

In this article, we consider testing the equality of two scale parameters in two power law populations. There is no exact test for this problem. Therefore, we propose the CAT for this hypothesis test.

2 The CAT for Testing Two Power Law Distribution

Let $X_{i1}, ..., X_{in_i}$, $i = 1, 2$, be two independent random samples from power law distribution, i.e. $X_{ij} \sim$pow(θ_i, β_i), $i = 1, 2$, $j = 1, ..., n_i$, where the parameters $(\theta_1, \theta_2, \beta_1, \beta_2)$ are unknown. We consider problems of testing

$$H_0 : \theta_1 = \theta_2 \ \ vs. \ \ H_1 : \theta_1 \neq (< or >)\theta_2, \tag{5}$$

which this test are equivalent to the following test

$$H_0^* : \delta = 0 \ \ vs. \ \ H_1^* : \delta \neq (< or >)0, \tag{6}$$

where $\delta = \theta_1 - \theta_2$.

The general format of CAT is given in Pal et al. (2007) and in the following we present the specific version of it as applicable for the given power law problem. Note that $\delta = \theta_1 - \theta_2$ is the parameter of interest.

Step 1. Obtain the MLE of the parameters as $\hat{\theta}_i = \max\{X_{i1}, ..., X_{in_i}\}$, $i = 1, 2$, and then $\hat{\delta} = \hat{\theta}_1 - \hat{\theta}_2$.

Step 2. Assume that H_0^* is true, i.e. $\theta_1 = \theta_2 = \theta$. Under this restricted model, we have $X_{ij} \sim$pow(θ, β_i), $i = 1, 2$, $j = 1, ..., n_i$. Get the MLE's of three parameters $(\theta, \beta_1, \beta_2)$ which are called the restricted MLEs as given

$$\hat{\theta}_R = \frac{\max\{X_{11}, ..., X_{1n_1}\} + \max\{X_{21}, ..., X_{2n_2}\}}{2}, \tag{7}$$

and

$$\hat{\beta}_{i(R)} = \frac{1}{\log(\hat{\theta}_R) - \frac{1}{n_i}\sum_{j=1}^{n_i} \log(X_{ij})}, \quad i = 1, 2. \tag{8}$$

Step 3. Generate artificial data X_{ij}, $i = 1, 2$, $j = 1, ..., n_i$ from pow$(\hat{\theta}_R, \hat{\beta}_{i(R)})$ a large number of times (say, M times). For each of these replicated samples, recalculate the MLE of δ. Thus we will have $\hat{\delta}_{01}, \hat{\delta}_{02}, ..., \hat{\delta}_{0M}$.

Step 4. Let $\hat{\delta}_{0(1)} \leq \hat{\delta}_{0(2)} \leq ... \leq \hat{\delta}_{0(M)}$ be the ordered values of $\hat{\delta}_{0l}$, $l = 1, ..., M$.

Step 5. (i) For testing $H_0^* : \delta = 0$ against $H_1^* : \delta < 0$, define the critical value $\hat{\delta}_L$ as $\hat{\delta}_L = \hat{\delta}_{0(\alpha M)}$. If $\hat{\delta} < \hat{\delta}_L$, then H_0^* is rejected. Alternatively the p-value can be defined as

$$p = \frac{1}{M}\sum_{l=1}^{M} I(\hat{\delta}_{0(l)} < \hat{\delta}). \tag{9}$$

(ii) For testing $H_0^* : \delta = 0$ against $H_1^* : \delta > 0$, define $\hat{\delta}_U = \hat{\delta}_{0((1-\alpha)M)}$. If $\hat{\delta} > \hat{\delta}_U$, then H_0^* is rejected. Alternatively the p-value can be defined as

$$p = \frac{1}{M} \sum_{l=1}^{M} I(\hat{\delta}_{0(l)} > \hat{\delta}). \tag{10}$$

(iii) For testing $H_0^* : \delta = 0$ against $H_1^* : \delta \neq 0$, define $\hat{\delta}_L = \hat{\delta}_{0((\alpha/2)M)}$ and $\hat{\delta}_U = \hat{\delta}_{0((1-\alpha/2)M)}$. If $\hat{\delta}$, is either greater than $\hat{\delta}_U$ or less than $\hat{\delta}_L$, then H_0^* is rejected. Alternatively the p-value can be defined as

$$p = 2\min(p_1, p_2), \tag{11}$$

where $p_1 = \frac{1}{M}\sum_{l=1}^{M} I(\hat{\delta}_{0(l)} < \hat{\delta})$ and $p_2 = \frac{1}{M}\sum_{l=1}^{M} I(\hat{\delta}_{0(l)} > \hat{\delta})$.

3 Numerical Studies

A simulation study is performed for evaluating the size and power of CAT in comparing the scale parameters of two Power distributions. Two set sample are generated; first data set with size n_1, is generated from power distribution with parameters $\theta_1 = 3$ and $\beta_1 = 1$ and second data set with size n_2, is generated from power distribution with different parameters θ_2 and β_2. The p-value for CAT with $M = 10000$ simulations is computed for testing $H_0 : \theta_1 = \theta_2$ vs. $H_1 : \theta_1 \neq \theta_2$. For $N = 10000$ replication, the p-values are computed. The size (power) of the CAT is the number of cases that the p-values are smaller than nominal level $\alpha = 0.05$. The results for different value are given in Table 1. We can find that the actual sizes of the CAT are close to nominal level and powers of the CAT are satisfactory.

Table 1: The actual sizes and powers of test at 5% significant level when $\theta_1 = 3$ and $\beta_1 = 1$.

		θ_2					
β_2	(n_1, n_2)	3.0	3.5	4.0	4.2	4.5	5.0
0.5	(10, 10)	0.0717	0.2385	0.5006	0.5943	0.7131	0.8303
	(20, 10)	0.0706	0.4154	0.6986	0.7535	0.8306	0.8961
	(10, 20)	0.0695	0.2413	0.5886	0.7149	0.8416	0.9370
	(20, 20)	0.0657	0.5039	0.8636	0.9161	0.9572	0.9845
	(30, 35)	0.0639	0.7982	0.9803	0.9913	0.9977	0.9993
	(50, 50)	0.0598	0.9487	0.9980	0.9991	1.0000	1.0000
	(100, 50)	0.0599	0.9743	0.9994	0.9999	0.9999	1.0000
	(50, 100)	0.0573	0.9948	1.0000	1.0000	1.0000	1.0000
1.0	(10, 10)	0.0698	0.2323	0.5672	0.6953	0.8265	0.9364
	(20, 10)	0.0682	0.4882	0.8665	0.9178	0.9589	0.9855
	(10, 20)	0.0679	0.2413	0.6311	0.7750	0.9007	0.9797
	(20, 20)	0.0669	0.5805	0.9623	0.9843	0.9971	0.9998
	(30, 35)	0.0615	0.9155	0.9993	0.9999	1.0000	1.0000
	(50, 50)	0.0591	0.9952	1.0000	1.0000	1.0000	1.0000
	(100, 50)	0.0565	0.9988	1.0000	1.0000	1.0000	1.0000
	(50, 100)	0.0553	1.0000	1.0000	1.0000	1.0000	1.0000
1.5	(10, 10)	0.0667	0.2293	0.5901	0.7330	0.8677	0.9662
	(20, 10)	0.0655	0.5326	0.9305	0.9695	0.9874	0.9968
	(10, 20)	0.0631	0.2400	0.6422	0.7853	0.9145	0.9863
	(20, 20)	0.0613	0.6079	0.9872	0.9961	0.9999	1.0000
	(30, 35)	0.0607	0.9518	0.9999	1.0000	1.0000	1.0000
	(50, 50)	0.0570	0.9998	1.0000	1.0000	1.0000	1.0000
	(100, 50)	0.0552	0.9999	1.0000	1.0000	1.0000	1.0000
	(50, 100)	0.0547	1.0000	1.0000	1.0000	1.0000	1.0000
2.0	(10, 10)	0.0656	0.2303	0.6023	0.7515	0.8880	0.9776
	(20, 10)	0.0669	0.5553	0.9603	0.9870	0.9960	0.9997
	(10, 20)	0.0621	0.2410	0.6449	0.7890	0.9195	0.9888
	(20, 20)	0.0609	0.6181	0.9945	0.9991	1.0000	1.0000
	(30, 35)	0.0578	0.9665	1.0000	1.0000	1.0000	1.0000
	(50, 50)	0.0550	0.9999	1.0000	1.0000	1.0000	1.0000
	(100, 50)	0.0555	1.0000	1.0000	1.0000	1.0000	1.0000
	(50, 100)	0.0543	1.0000	1.0000	1.0000	1.0000	1.0000

Example. To illustrate the CAT we simulated data on $X \sim$pow$(2,1)$ and on $Y \sim$pow$(3,5)$ with $n_1 = n_2 = 10$. Therefore, two data set are generated from power law distribution. The two set data are

Data 1	0.6436	1.0204	1.7635	1.7734	1.1493	1.3233	0.0103	1.7982	1.5776	0.1547
Data 2	2.3280	2.8040	1.6894	2.4329	2.6147	2.9358	2.6694	2.9960	2.6147	2.8855

The MLEs are $\hat{\theta}_1 = 1.7982$, $\hat{\beta}_1 = 0.9874$, $\hat{\theta}_2 = 2.9960$, $\hat{\beta}_2 = 6.4718$. Using given steps in this Section, with $M = 100000$, we computed the p-value for testing $H_0 : \theta_1 = \theta_2 \;\; vs. \;\; H_2 : \theta_1 \neq \theta_2$, as 0.0106, which suggests that the data provide evidence against H_0.

References

[1] Casella, G., Berger, R. L. (1990). *Statistical Inference*. Wadsworth, Pacific Grove, CA.

[2] Chang, C. H., Pal, N. (2008a). A Revisit to the Behrens-Fisher Problem: Comparison of Five Test Methods, *Communications in statistics-Simulation and Computation*, **37**, 1064-1085.

[3] Chang, C. H., Pal, N. (2008b). Testing on the common mean of several normal distributions. *Computational Statistics & Data Analysis* **53**, 321-333.

[4] Chang, C. H., Pal, N., Lim, W., and Lin, J. J. (2010). Comparing several population means: a parametric bootstrap method, and its comparison with usual ANOVA F test as well as ANOM, Computational Statistics,**25**, issue 1, 71-95.

[5] Johnson, N. L., Kotz, S. and Balakrishnan, N. (1995). *Continuous Univariate Distributions* (2nd ed.), Wiley, New York.

[6] Meniconi, M. and Barry, D. M. (1996). The Power Function Distribution: A useful and simple distribution to assess Electrical component reliability. *Microelectron. Reliab.*, **36**, No. 9, pp. 1207-1212.

[7] Pal, N., Lim, W. K. and Ling, C. H. (2007). A computational approach to statistical inferences, *J. Appl. Probab. Statist.*, *2*, 13-35.

Email:k.abdollahnezhad@gu.ac.ir

Email:aajafari@yazduni.ac.ir

43rd Annual Iranian Mathematics Conference,
27-30 August 2012
University of Tabriz

On Log-concavity of skew-symmetric distributions and their applications in penalized linear models

Vahid Nassiri
Vrije Universiteit Brussel, Belgium

Ignace Loris
Université Libre de Bruxelles, Belgium

Abstract
Log-concavity of the skew-symmetric class of distributions is studied. Also the possibility of using them as error distribution in a sparse linear model is investigated. A procedure to estimate the penalized model is discussed.
Keywords: Log-concavity, skew-symmetric distributions, linear models, sparsity

Mathematics Subject Classification: 62J07, 65K05

1 Introduction

Linear models, thanks to their good results and simplicity, lie at the heart of mathematical and statistical modeling. A linear model has the following general form:

$$Ax = b + \epsilon \tag{1}$$

where ϵ is the noise (error), $A \in \mathbb{R}^{n\times k}$, $b \in \mathbb{R}^{n\times 1}$, and $x \in \mathbb{R}^{k\times 1}$ is the unknown parameter vector.

Least squares (LS) models, which are connected to the Gaussian error distribution, are the most famous linear models. As observing asymmetry or outliers is very common in practice, using LS models does not seem reasonable in many applications. Therefore, the extension of Gaussian models to some models which allow asymmetry and heavy tails is very interesting. Quantile regression of [3] is an alternative for LS models, but its non-differentiable loss function makes it hard to deal with.

The class of skew normal distributions was introduced by [1]. His skewing strategy can be extended to any other symmetric distribution. Consider $f_0(x)$ a continuous symmetric probability density function (PDF) around $x = 0$. Its skewed version, $f(x)$, is defined as:

$$f(x) = 2f_0(x)G_0(\theta x), \tag{2}$$

where $G_0(.)$, the skewing function, is a continuous cumulative distribution function (CDF) on the real line, where $g_0 = G_0'$ is an even density function. $\theta \in \mathbb{R}$ is the skewness parameter. Considering the asymmetry of the class of distributions in (2), choosing a kernel with heavier tails comparing with Gaussian, would provide a suitable alternative for the LS models. The two main difficulties are estimating the model and the curse of dimensionality.

The use of sparse models is a promising approach for dealing with problem where many variables are involved, especially when the number of variables is higher than the observations. Penalizing the parameter vector with a suitable penalty (such as e.g. the ℓ_1-norm of the coefficients) produces a sparse model. A penalized estimation of x in model 1, is the solution of the following problem:

$$\min_x \{T(Ax - b) + \lambda \|x\|_1\} \tag{3}$$

where T is a loss function, and $\|.\|_1$ is the ℓ_1-norm: $\|x\|_1 = \sum_i^k |x_i|$. If the loss function T is convex and has a Lipschitz continuous derivative, [2] suggested the so-called fast iterative soft thresholding algorithm (FISTA) to solve (3). Their algorithm will be adapted to estimate our penalized model for a suitable skew-symmetric error.

2 Main Result

2.1 Log-concavity of the skew-symmetric densities

Considering the density in expression (2), its logarithm is as follows: $\log f(x) = \log f_0(x) + \log G_0(\theta x)$.

Theorem 2.1. *If f_0 and G_0 in relation (2) are twice differentiable, then the density in relation (2) is log-concave if and only if $\forall x \in \mathbb{R}$,*

$$G_0^2(\theta x)\left(f'^2_0(x) - f_0(x)f_0''(x)\right) + \theta^2 f_0^2(x)\left(g_0^2(\theta x) - G_0(\theta x)g_0'(\theta x)\right) \geq 0 \quad (4)$$

Corollary 2.2. *If in expression (4) we set $\Phi = f'^2_0(x) - f_0(x)f_0''(x)$ and $\Psi = g_0^2(\theta x) - G_0(\theta x)g_0'(\theta x)$ and also Φ^+, Φ^-, Ψ^+ and Ψ^- as positive and negative parts of Φ and Ψ, then for different types of f_0 and g_0 the Condition 4 is simplified in Table 1.*

Table 1: Log-concavity of f

f_0/G_0	*log-concave*	*log-convex*	*none*
log-concave	$\forall x \in \mathbb{R}$ *(4) is true*	$\Phi \geq -\Psi$	$\Phi + \Psi^+ \geq \Psi^-$
log-convex	$\Phi \geq -\Psi$	$\nexists x \in \mathbb{R}$ *that (4) is true*	$\Phi + \Psi^+ \geq \Psi^-$
none	$\Psi + \Phi^+ \geq -\Phi^-$	$\Phi^+ \geq -(\Phi^- + \Psi)$	$(\Phi^+ + \Psi^+) \geq -(\Phi^- + \Psi^-)$

[4] has shown the skewing distribution doesn't have a huge effect on the resulting skewness. Therefore, we may omit Student's-t as the skewing distribution for its very complicated form. Table 2 shows the results for some famous distributions.

Table 2: Log-concavity of different kernels (✓: log-concave, ⊖: not log-concave)

f_0/G_0	Gaussian	Logistic	Student's-$t(\nu)$
Gaussian	✓	✓	not studied
Logistic	✓	✓	not studied
Students'$-t(\nu)$	⊖	⊖	⊖

The case of Student's-t kernel with Logistic skewing function needs to be studied more.

Corollary 2.3. *The density of the form $\frac{2}{\sqrt{\nu}\beta(\nu/2,1/2)}\left(1+\frac{x^2}{\nu}\right)^{-(1+\nu)/2}\frac{1}{1+e^{-\theta x}}$ is not log-concave unless $\nu \to \infty$.*

2.2 Lipschitz continuity of $-\frac{\partial}{\partial x}\log f(x)$

As it has been mentioned, the derivative of the loss function needs to be Lipschitz continuous, one would also need to find the Lipschitz constant (LC). Assume the $f(x)$ is twice differentiable. Then:

$$-\frac{\partial^2}{\partial x^2}\log f(x) = -\left(\frac{f_0''(x)}{f_0(x)} - \left(\frac{f_0'(x)}{f_0(x)}\right)^2\right) - \theta^2\left(\frac{g_0'(\theta x)}{G_0(\theta x)} - \left(\frac{g_0(\theta x)}{G_0(\theta x)}\right)^2\right) \quad (5)$$

When $x \to +\infty$, the second part in (5) is zero. Otherwise, it would be $\frac{0}{0}$, one may use l'Hospital's rule to find the derivative. So it would be possible that the second derivative be unbounded. Consider the following theorem for a Gaussian skewing function.

Theorem 2.4. *If $G_0(\theta x) = \Phi(\theta x)$ and $\Phi(.)$ is the Gaussian CDF, then $\frac{\partial}{\partial x}(-\log(G_0(\theta x))$ is not Lipschitz continuous.*

Consequently, one may just consider two cases:

$$\begin{array}{lll} \text{logistic skewed logistic}: & -\frac{\partial^2}{\partial x^2}\log f(x) = 2\frac{e^{-x}}{(1+e^{-x})^2} + \theta^2\frac{e^{-\theta x}}{(1+e^{-\theta x})^2}, & LC = \frac{\theta^2}{4} + \frac{1}{2} \\ \text{logistic skewed Gaussian}: & -\frac{\partial^2}{\partial x^2}\log f(x) = 1 + \theta^2\frac{e^{-\theta x}}{(1+e^{-\theta x})}, & LC = \frac{\theta^2}{4} + 1. \end{array} \qquad (6)$$

References

[1] A. Azzalini, *A class of distributions which includes the normal ones*, Scand. J. Statist., 12 (1985), pp. 171–178.

[2] A. Beck and M. Teboulle, *A fast iterative shrinkage-threshold algorithm for linear inverse problems*, SIAM Journal on Imaging Sciences, 2 (2009), pp. 183–202.

[3] R. Koenker and G. Bassett, *Regression quantiles*, Econometrics, 46 (1978), pp. 33–50.

[4] D. Umbach, *The effect of the skewing distribution on skew-symmetric families*, Soochow J. Math., 33 (2007), pp. 657–668.

Email:vnassiri@vub.ac.be

Email:igloris@ulb.ac.be

43^{rd} Annual Iranian Mathematics Conference,
27-30 August 2012
University of Tabriz

On the quartic loss estimation

M. Arashi
Shahrood University of Technology

M. Norouzirad
Islamic Azad University, Mashhad Branch

Abstract
It is sometimes interesting to find an estimator of the underlying loss function as a measure of closeness. In this presentation, a general form of an unbiased estimator of the quartic loss function which is polynomial in coordinates, is derived. It is worthwhile to point that the unbiased estimator of the loss is constructed according and by the use of Stein-type shrinkage estimators which are almost biased.
Keywords: Almost differentiable, Quartic loss function, Shrinkage, Unbiased.

Mathematics Subject Classification: 62F10, 62H12

1 Introduction

Let $\boldsymbol{X} \sim N_p(\boldsymbol{\theta}, \boldsymbol{I}_p)$, where $p \geq 3$. Formally we want to estimate $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_p)$ using an estimator $\boldsymbol{\delta}(\boldsymbol{X}) = (\delta_1(\boldsymbol{X}), \delta_2(\boldsymbol{X}), \ldots, \delta_p(\boldsymbol{X}))$. $\boldsymbol{X} = (X_1, X_2, \ldots, X_p)$, is an inadmissible estimator of $\boldsymbol{\theta}$ under the quadratic loss (QDL) function for $p \geq 3$ (see Stein, 1956). James and Stein (1961) showed that the estimator $\boldsymbol{\delta}^a_{JS}(\boldsymbol{X}) = \left(1 - \frac{a}{\|\boldsymbol{X}\|^2}\right)\boldsymbol{X}$, dominates $\boldsymbol{\delta}_{ML}(\boldsymbol{X})$ in the sense of having smaller risk function under the QDL. When $0 < a \leq (p-2)$, the quadratic risk of $\boldsymbol{\delta}^a_{JS}(\boldsymbol{X})$ is minimized at $a = p-2$. Further, they showed that for sufficiently small a and large b (independent of $\boldsymbol{\theta}$) the estimator, $\boldsymbol{\delta}^{a,b}_{JS}(\boldsymbol{X}) = \left(1 - \frac{a}{b+\|\boldsymbol{X}\|^2}\right)\boldsymbol{X}$, has smaller quadratic risk than the usual estimator $\boldsymbol{X}$.

In the estimation of a loss function, two view points arise naturally; one is to measure the financial loss, which is not proposed so far. The second is to formulate a function, which has minimum risk. Since in reality the loss is not known in the above two situations, studying each of them would be beneficial in one sense. However the latter has been studied at some length in the work of Sandved (1968), Kiefer (1976), Rukhin (1988), Lu and Berger (1989), Lele (1992), Wan and Zou (2004) and Norouzirad and Arashi (2012) among others. They considerably proposed some sorts of estimators for the QDL function.

Berger (1978) considered loss functions L which are polynomials in the coordinates of $\boldsymbol{\delta} - \boldsymbol{\theta}$. He gave a developed example when $L(\boldsymbol{\delta}, \boldsymbol{\theta}) = \|\boldsymbol{\delta} - \boldsymbol{\theta}\|^4 = \left(\sum_{i=1}^{p}(\delta_i - \theta_i)^2\right)^2$. He stated that James-Stein type estimator $\boldsymbol{\delta}^{a,b}_{JS}(\boldsymbol{X})$ dominates $\boldsymbol{\delta}_{ML}(\boldsymbol{X})$ for all $\theta \in \mathcal{R}^p$ provided that $p \geq 3$, $0 \leq a \leq 2(p-2)$ and $b \geq 2 + \frac{p+1}{p+2}a$.

2 Preliminaries

In the following we propose some necessary tools.

Definition 2.1. *$\hat{L}$ is a unbiased estimator of L if it is an unbiased estimator of $R = E[L]$, i.e. $E[\hat{L}] = R = E[L]$.*

Using Stein's identity (1981) we have the following result.

Lemma 2.2. *Let $\boldsymbol{X} \sim N_p(\boldsymbol{\theta}, \boldsymbol{I}_p)$. Assume $g(\boldsymbol{x})$ is a real-valued p times continuously differentiable function. Then*

$$E\left[g(X_i)(X_i-\theta_i)^2\right] = E\left[\frac{\partial^2}{\partial X_i^2}g(X_i)\right] + E\left[g(X_i)\right]$$

$$E\left[g(X_i)\left\|\boldsymbol{X}-\boldsymbol{\theta}\right\|^2\right] = pE\left[g(X_i)\right] + E\left[\frac{\partial^2}{\partial X_i^2}g(X_i)\right]$$

$$E\left[g(X_i,X_j)(X_i-\theta_i)(X_j-\theta_j)\right] = E\left[\frac{\partial^2}{\partial X_i \partial X_j}g(X_i,X_j)\right]$$

$$E\left[g(X_i)\left\|\boldsymbol{X}-\boldsymbol{\theta}\right\|^2(X_i-\theta_i)\right] = (p+2)E\left[\frac{\partial}{\partial X_i}g(X_i)\right] + 2E\left[\frac{\partial^3}{\partial X_i^3}g(X_i)\right] + \sum_{\substack{j=1\\ i\neq j}}^{p} E\left[\frac{\partial^3}{\partial X_i \partial X_j^2}g(X_i)\right]$$

3 Main Result

In this section we propose an unbiased shrinkage type estimator of the quartic loss function.

Theorem 3.1. *The unbiased estimator of $L(\boldsymbol{\delta}^a_{JS}(\boldsymbol{X}), \boldsymbol{\theta})$ under the quartic loss function is given by*

$$\begin{aligned}\hat{L}(\boldsymbol{\delta}^a_{JS}(\boldsymbol{X}), \boldsymbol{\theta}) \quad = \quad & p(p+2) + \frac{\Psi_1(a)\|\boldsymbol{X}\|^2}{(b+\|\boldsymbol{X}\|^2)^2} + \frac{\psi_2(a)}{(b+\|\boldsymbol{X}\|^2)^2} + \frac{\psi_3(a)\|\boldsymbol{X}\|^2}{(b+\|\boldsymbol{X}\|^2)^3} \\ & \frac{\psi_4(a)\sum_{i=1}^p X_i^4}{(b+\|\boldsymbol{X}\|^2)^4} + \frac{\psi_5(a)\|\boldsymbol{X}\|^4}{(b+\|\boldsymbol{X}\|^2)^4} + 192a^2\frac{\sum_{i=1}^p \sum_{\substack{j=1\\ j\neq i}}^p X_i^2 X_j^2}{(b+\|\boldsymbol{X}\|^2)^4} \\ & +192a\frac{\sum_{i=1}^p \sum_{j=i+1}^p X_i^2 X_j^2}{(b+\|\boldsymbol{X}\|^2)^4}\end{aligned}$$

where $\psi_1(a) = a^2(2p+4) - 4(p-4)a$,$\psi_2(a) = 4ap(a(p+2)-3b)$,$\psi_3(a) = -4(p+2)a^2(a+10) + 32a(p-4)$, $\psi_4(a) = 96a(a+2)$ and $\psi_5(a) = a^2(a^2+24a+48)$

Proof. By definition 1.1, we have,

$$\begin{aligned} E\left[\left\|\left(1-\frac{a}{b+\|\boldsymbol{X}\|^2}\right)\boldsymbol{X}-\boldsymbol{\theta}\right\|^4\right] \quad = \quad & E\left[\|\boldsymbol{X}-\theta\|^4\right] + 4a^2E\left[\sum_{i=1}^p \frac{(X_i-\theta_i)X_i}{(b+\|\boldsymbol{X}\|^2)}\right]^2 \\ & +a^4E\left[\frac{\|\boldsymbol{X}\|^4}{(b+\|\boldsymbol{X}\|^2)^4}\right] - 4aE\left[\frac{\|\boldsymbol{X}-\theta\|^2\sum_{i=1}^p(X_i-\theta_i)X_i}{(b+\|\boldsymbol{X}\|^2)}\right] \\ & +2a^2E\left[\frac{\|\boldsymbol{X}-\theta\|^2\|\boldsymbol{X}\|^2}{(b+\|\boldsymbol{X}\|^2)^2}\right] - 4a^3E\left[\frac{\|\boldsymbol{X}\|^2\sum_{i=1}^p(X_i-\theta_i)X_i}{(b+\|\boldsymbol{X}\|^2)^3}\right] \\ = \quad & E\left[\|\boldsymbol{X}-\theta\|^4\right] + 4a^2\sum_{i=1}^p E\left[\frac{(X_i-\theta_i)^2X_i^2}{(b+\|\boldsymbol{X}\|^2)}\right] \\ & +8a^2\sum_{i=1}^p\sum_{j=i+1}^p E\left[\frac{X_iX_j(X_i-\theta_i)(X_j-\theta_j)}{(b+\|\boldsymbol{X}\|^2)}\right] \\ & +a^4E\left[\frac{\|\boldsymbol{X}\|^4}{(b+\|\boldsymbol{X}\|^2)^4}\right] - 4a\sum_{i=1}^p E\left[\frac{\|\boldsymbol{X}-\theta\|^2(X_i-\theta_i)X_i}{(b+\|\boldsymbol{X}\|^2)}\right] \end{aligned}$$

$$+2a^2 E\left[\frac{\|\boldsymbol{X}-\theta\|^2\|\boldsymbol{X}\|^2}{(b+\|\boldsymbol{X}\|^2)^2}\right]-4a^3\sum_{i=1}^{p}E\left[\frac{\|\boldsymbol{X}\|^2(X_i-\theta_i)X_i}{(b+\|\boldsymbol{X}\|^2)^3}\right]$$

Using Lemma 2.2, after some algebraic computations, we get

$$\begin{aligned}
E\left[L(\boldsymbol{\delta}^a_{JS}(\boldsymbol{X}),\boldsymbol{\theta})\right] &= E\left[\|\boldsymbol{X}-\boldsymbol{\theta}\|^4\right]+4a^2\left\{E\left[\frac{\|\boldsymbol{X}\|^2}{(b+\|\boldsymbol{X}\|^2)^2}\right]+2pE\left[\frac{1}{(b+\|\boldsymbol{X}\|^2)^2}\right]\right.\\
&\left.-20E\left[\frac{\|\boldsymbol{X}\|^2}{(b+\|\boldsymbol{X}\|^2)^3}\right]+24E\left[\frac{\sum_{i=1}^p X_i^4}{(b+\|\boldsymbol{X}\|^2)^4}\right]\right\}\\
&+8a^2\left\{\frac{p(p-1)}{2}E\left[\frac{1}{(b+\|\boldsymbol{X}\|^2)^2}\right]-4(p-1)E\left[\frac{\|\boldsymbol{X}\|^2}{(b+\|\boldsymbol{X}\|^2)^3}\right]\right.\\
&\left.+24E\left[\frac{\sum_{i=1}^p\sum_{j=i+1}^p X_i^2X_j^2}{(b+\|\boldsymbol{X}\|^2)^4}\right]\right\}+a^4E\left[\frac{\|\boldsymbol{X}\|^4}{(b+\|\boldsymbol{X}\|^2)^4}\right]\\
&-4a\left\{3pbE\left[\frac{1}{(b+\|\boldsymbol{X}\|^2)^2}\right]+(p-4)E\left[\frac{\|\boldsymbol{X}\|^2}{(b+\|\boldsymbol{X}\|^2)^2}\right]\right.\\
&-8(p-4)E\left[\frac{\|\boldsymbol{X}\|^2}{(b+\|\boldsymbol{X}\|^2)^3}\right]-48E\left[\frac{\sum_{i=1}^p X_i^4}{(b+\|\boldsymbol{X}\|^2)^4}\right]\\
&\left.-48E\left[\frac{\sum_{i=1}^p\sum_{\substack{j=1\\ j\neq i}}^p X_i^2X_j^2}{(b+\|\boldsymbol{X}\|^2)^4}\right]\right\}+2a^2\left\{pE\left[\frac{\|\boldsymbol{X}\|^2}{(b+\|\boldsymbol{X}\|^2)^2}\right]\right.\\
&+2pE\left[\frac{1}{(b+\|\boldsymbol{X}\|^2)^2}\right]-4(p+4)E\left[\frac{\|\boldsymbol{X}\|^2}{(b+\|\boldsymbol{X}\|^2)^3}\right]\\
&\left.+24E\left[\frac{\|\boldsymbol{X}\|^4}{(b+\|\boldsymbol{X}\|^2)^4}\right]\right\}-4a^3\left\{(p+2)E\left[\frac{\|\boldsymbol{X}\|^2}{(b+\|\boldsymbol{X}\|^2)^3}\right]\right.\\
&\left.-6E\left[\frac{\|\boldsymbol{X}\|^4}{(b+\|\boldsymbol{X}\|^2)^4}\right]\right\}\\
&= p(p+2)+\left\{a^2(2p+4)-4(p-4)a\right\}E\left[\frac{\|\boldsymbol{X}\|^2}{(b+\|\boldsymbol{X}\|^2)^2}\right]\\
&+4ap(a(p+2)-3b)E\left[\frac{1}{(b+\|\boldsymbol{X}\|^2)^2}\right]\\
&+\left\{-4(p+2)a^2(a+10)+32a(p-4)\right\}E\left[\frac{\|\boldsymbol{X}\|^2}{(b+\|\boldsymbol{X}\|^2)^3}\right]\\
&+96a(a+2)E\left[\frac{\sum_{i=1}^p X_i^4}{(b+\|\boldsymbol{X}\|^2)^4}\right]+a^2(a^2+24a+48)E\left[\frac{\|\boldsymbol{X}\|^4}{(b+\|\boldsymbol{X}\|^2)^4}\right]\\
&+192a^2E\left[\frac{\sum_{i=1}^p\sum_{\substack{j=1\\ j\neq i}}^p X_i^2X_j^2}{(b+\|\boldsymbol{x}\|^2)^4}\right]+192aE\left[\frac{\sum_{i=1}^p\sum_{j=i+1}^p X_i^2X_j^2}{(b+\|\boldsymbol{X}\|^2)^4}\right]\\
&= E\left[p(p+2)+\frac{\Psi_1(a)\|\boldsymbol{X}\|^2}{(b+\|\boldsymbol{X}\|^2)^2}+\frac{\psi_2(a)}{(b+\|\boldsymbol{X}\|^2)^2}+\frac{\psi_3(a)\|\boldsymbol{X}\|^2}{(b+\|\boldsymbol{X}\|^2)^3}\right.\\
&\frac{\psi_4(a)\sum_{i=1}^p X_i^4}{(b+\|\boldsymbol{X}\|^2)^4}+\frac{\psi_5(a)\|\boldsymbol{X}\|^4}{(b+\|\boldsymbol{X}\|^2)^4}+192a^2\frac{\sum_{i=1}^p\sum_{\substack{j=1\\ j\neq i}}^p X_i^2X_j^2}{(b+\|\boldsymbol{X}\|^2)^4}\\
&\left.+192a\frac{\sum_{i=1}^p\sum_{j=i+1}^p X_i^2X_j^2}{(b+\|\boldsymbol{X}\|^2)^4}\right]
\end{aligned}$$

which completes the proof. □

References

[1] J. O. Berger,*Minimax estimation of a multivariate normal mean under polynomial loss*,J. Multivariate Anal., 8, (1978), pp. 173–180.

[2] J. Kiefer,*Admissibility conditional confidence procedures*,Anal. Statist., 4, (1976), pp. 836–865.

[3] W. James and C. Stein, *Estimation with quadratic loss, in Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press: Berkeley, (1961), pp. 361–380.

[4] M. Norouzirad and M. Arashi, *A note on loss estimation*, Appl. Math. Letters, 25, (2012), pp. 779–783.

[5] C. Lele, *Inadmissibility of loss estimators*, Statist. Decisions., 10, (1992), pp. 309–322.

[6] K. L. Lu and J. O. Berger,*Estimation of normal means: frequentist estimation of loss*, Anal. Statist., 17, (1989), pp. 890–906.

[7] A. L. Rukhin, *Estimated loss and admissible loss estimators* , in Statistical Decision Theory and Related Topics IV (S.S. Gupta and J.O. Berger, eds.), Springer Verlag: New York, (1988), pp. 409–418.

[8] E. Sandved, *Ancillary statistics and estimation of the loss in estimation problems*, Ann. Math. Statist., 39, (1968), pp. 1755–1758.

[9] C. Stein, *Inadmissibility of the usual estimator for the mean of a multivariate normal distribution*, in Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, University of California Press: Berkeley, (1956), pp. 197–206.

[10] A. T. K. Wan and G. Zou, *On Unbiased and improved loss estimation for the mean of a multivariate normal distribution with unknown variance*, J. Statist. Plann. Inf., 119, (2004), pp. 17–22.

Email:m_arashi_stat@yahoo.com

Email:mina.norouzirad@gmail.com

Mathematics in Science and Technology

43^{rd} Annual Iranian Mathematics Conference,
27-30 August 2012
University of Tabriz

Multiparty semiquantum secret sharing using entangled states

Massoud Hadian Dehkordi
Iran University of Science and Technology

Elham Fattahi
Iran University of Science and Technology

Abstract

Quantum secret sharing (QSS) is a procedure of sharing secret information by using quantum states. Eavesdropping, the most important problem of classical secret sharing is solved by presenting the idea of QSS. but what happens if not all the participants are quantum? In this paper we present a semiquantum secret sharing scheme for quantum Alice and the arbitrary number classical agents. The presented scheme is a combination of all kinds of (k, n) threshold schemes for $k = 1, ..., n$. Our protocol is shown to be secure against eavesdropping.
Keywords: Secret sharing, quantum cryptography, Multiparty

1 Introduction

The combination of quantum mechanics with information security has produced many interesting and important developments, most notably Quantum Cryptography. Quantum Secret Sharing (QSS) generalizes classical secret sharing scheme into the quantum scenario. QSS is a method that takes care of the problem of eavesdropping. If an eavesdropper or one dishonest agent eavesdrop on the communication channel to obtain all shares in the secret, then no cooperation is needed to recover Alice s secret. This is where QSS is useful, since it enables secret sharing and eavesdropping detection at the same time. The first QSS scheme proposed by Hillery *et al* [1] used three-particle entangled Greenberger-Horne-Zeilinger (GHZ) states. A number of examples of QSS schemes have been reported [2] with different intriguing properties. Other entangled states such as Einstein-Podolsky-Rosen (EPR) pairs have been used in QSS protocols , and even multi-particle product states have been proposed for this purpose [3]. Thus, entanglement is not necessary for QSS.

Recently, Qin Li *et al.*have proposed an interesting semiquantum secret sharing scheme in which Alice has full quantum capabilities and shares a secret with two parties, Bob and Charlie, who have restricted quantum capabilities (called “classical”) [4]. This scheme uses GHZ states for carrying information.

In this paper, we have proposed a multiparty semiquantum secret sharing scheme. In the four-party case, the “quantum” Alice shares a secret with a “classical” Bob, Charlie and Dave so that these three agents need to collaborate to reconstruct the secret. The interesting property of this protocol is that Alice may, depending on choices made by the three other agents, share a *Private key* with one of the agents, a *Pair key* with two agents, or a *Triple key* with all three agents. These alternatives will occur with equal probability. In other words, this protocol is a probabilistic combination of all three of $(1, 3)$, $(2, 3)$ and $(3, 3)$ threshold secret sharing, while still allowing for detection of eavesdropping, and also allowing restricted quantum abilities of the three agents Bob, Charlie, and Dave.

In this scheme Alice needs to have certain quantum abilities: she needs to be able to prepare a specific entangled state (see below) and also certain measurements such as Bell state discrimination, GHZ state discrimination, ..., up to a number of qubits equal to the number of sharing agents. The classical parties, Bob, Charlie, Dave, ..., only need to be able to perform these three operations:

1. measure the qubits in the standard basis,
2. reorder the qubits,
3. return the qubits without disturbance.

2 Semiquantum scheme with classical Bob

In this section we describe the scheme with quantum Alice and classical Bob. Bob is able to do three above operations. while Alice has all of the quantum capabilities. The algorithm of the scheme is as following:

(i) Alice prepares a long string (N is the length of the string) of two-qubit entangled state in the form of EPR pair

$$|\Phi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$$

(ii) she keeps the first particle of each EPR pair for herself and sends the second one to Bob.

(iii) Bob randomly decides to measure each particle in standard basis (SHARE) or resend it to Alice without any changes (CHECK). But before returning the particles, Bob has to rearrange them in the new order in order to make it impossible for eavesdropper to realize which particles are resent.

(iv) Alice announces that she has received the particles, Now Bob reveals the original order of particles and which qubits are resent. According to this report, Alice can reorder the qubits in original order.

(v) In this case according to the action which Bob did, Alice performs various actions for each qubit. If Bob choose to SHARE then Alice measures her corresponding particle in standard basis as well. Alice can guess Bob's measurement result by observing her measurement outcome. On the other hand if Bob decides to CHECK, then Alice combines her particle with reflected one and measures them by Bell basis. If she gets any states except the original Bell state as an outcome, she can realize the existence of eavesdropping on the channel.

(vi) Alice evaluates the error rate, if this amount be higher than the predefined threshold value, the protocol aborts.

Suppose an eavesdropper Eve (with quantum abilities)is going to find Alice's secret message without being detected. She received all of the particles on the channel before Bob can access them. Eve measures all of them by standard basis. The problem occurs when Bob decides to CHECK the particle which is measured by Eve. when this particle is combined by Alice's corresponding qubit, they produce a product state instead of a Bell state. Then with probability $1-(\frac{1}{2})^N$, eavesdropping is detected.

3 Three-party semiquantum secret sharing scheme

Now we describe our protocol in some steps:

(i) Alice prepares a long string of four-qubit entangled state in the form of

$$\begin{aligned}|\Phi\rangle = \frac{1}{2\sqrt{2}}(&|0000\rangle + |0011\rangle + |0101\rangle + |0110\rangle \\ &+|1001\rangle + |1010\rangle + |1100\rangle + |1111\rangle) \qquad (1)\end{aligned}$$

(the length of the string is N and the states are indexed from 1 to N).Then Alice sends the second, the third and the fourth qubit of each entangled state to Bob, Charlie and Dave respectively. She keeps the first qubit for herself.

(ii) After Bob receives the qubits, for each qubit he randomly decides to measure that using the standard basis (we call this action as SHARE)or to resend the qubit to Alice without any changes (we refer to this action as CHECK), but before returning the qubits, Bob has to rearrange them in the new order. Because by reordering the qubits nobody could realize which qubits are returned. The outcome of each measurement is a binary bit 0 or 1. Charlie and Dave also does the same thing.

(iii) Alice announces that she has received agents's reflected qubits, Now Bob, Charlie and Dave reveal which qubits were reflected and the original order of them. According to their reports, Alice can reorder the qubits in the original order.

(iv) In this step, for each qubit Alice performs one of the eight actions which are defined as below, on her own qubit according to Bob's, Charlie's and Dave's actions, as illustrated in Table 1. The eight cases in Table 1 appears with the same probability.

Case	Bob	Charlie	Dave	Alice
Case 1	SHARE	SHARE	SHARE	Action 1
Case 2	SHARE	SHARE	CHECK	Action 2
Case 3	SHARE	CHECK	SHARE	Action 3
Case 4	CHECK	SHARE	SHARE	Action 4
Case 5	SHARE	CHECK	CHECK	Action 5
Case 6	CHECK	SHARE	CHECK	Action 6
Case 7	CHECK	CHECK	SHARE	Action 7
Case 8	CHECK	CHECK	CHECK	Action 8

Table 1: Participants' actions on the qubits in each position

(v) Alice checks the error rate in cases (2),(3),....,(7), and (8). If the error rate in any case is higher than some predefined threshold value, the protocol aborts.

(vi) In case(1), when Bob, Charlie and Dave choose to SHARE, Alice requires all of the agents to reveal a random subset of the bits which are used to generate Alice's secret bit. This process is used to check the error rate in case(1). All of the actions are described completely in [5]

4 Efficiency and Security

We show the preceding three-party semiquantum secret sharing protocol is secure against eavesdropping in two situations. The first situation is that one of the Alice's agents, including Bob, Charlie or Dave, is dishonest and attempts to recover Alice's secret message without cooperating with the other two parties.

The second situation is that an eavesdropper Eve (with quantum abilities) is going to find Alice's secret message without being detected. We first suppose the dishonest classical Bob can access all of Alice's transmissions. In one case, when Bob receives three transmissions from Alice, he measures each particle using standard basis. He keeps one of them for himself and resend the two remaining particles in the state he found to Charlie and Dave. According to Eq. (1), Bob can recover Alice's secret bit by operating an exclusive-or addition on measurement's outcomes. If both of Charlie and Dave choose to SHARE, then Bob is not detected and can be succeed.

Otherwise if both Charlie and Dave or one of them choose to CHECK, then with probability $(3/4)^2$, Bob escapes from being detected. In other case, Bob just measures his own particle and one of the Charlie's or Dave's particle and resend the other qubit without any changes. In this situation, the probability that Bob escapes from being detected is $(3/4)$.

Assume that Bob has to measure Charlie's and Dave's qubits in k ($k \leq N$) positions to obtain the significant information of Alice's secret without cooperating with Charlie and Dave, then the probability that Bob goes undetected is $(3/4)^{2k}$ which is arbitrarily small if k and N are picked large enough. We have discussed the second situation in the extended version of paper in [5].

References

[1] M. Hillery, V. Buzěk, and A. Berthiaume, *Quantum secret sharing*, Phys. Rev. A, 59 (1999), pp. 1829–1834.

[2] L. Xiao ,G. Long, F. G. Deng, and J. .W. Pan, *Efficient multiparty quantum-secret-sharing schemes.*, Phys. Rev. A, 69 (2004), pp. 052307.

[3] G. P. Guo, and G. C. Guo, *Quantum secret sharing without entanglement*, Phys. Lett. A, 310 (2003), pp. 247.

[4] Q. Li, W. H. Chan, and D. Y. Long, *Semi-quantum secret sharing using entangled states* , Phys. Rev. A, 82 (2010), pp. 022303.

[5] M. Hadian Dehkordi, E. Fattahi, *Semiquantum secret sharing* , preprint.

Email:mhadian@iust.ac.ir

Email:elham_fattahi@iust.ac.ir

43rd Annual Iranian Mathematics Conference,
27-30 August 2012
University of Tabriz

Traveling wave solutions of a biological reaction-convection-diffusion equation model by using (G'/G) expansion method

Sh. Javadi
Kharazmi University

M. Jani
Kharazmi University

J. Tayyebi
University of Birjand

Abstract

In this paper, the (G'/G)-expansion method is applied to solve generalization of Fisher and Burgers equations arising in mathematical biology. Exact traveling wave solutions are obtained by this method. This scheme can be easily applied to a wide class of nonlinear partial differential equations.

Keywords: (G'/G)-expansion method, Reaction-convection-diffusion equation, Biological model

Mathematics Subject Classification: 35A20, 35C07, 35C09

1 Introduction

Mathematical modeling of physical and biological systems often leads to nonlinear evolution equations. Various ansatze have been proposed for seeking traveling wave solutions of these equations. Recently, some new methods have been proposed to find some particular solutions of these problems, for instance, the homogenous balance method [1], the Jacobi elliptic function method [2], etc.

The objective of this paper is to use the (G'/G)-expansion method to investigate exact traveling wave solutions of a biological model that is a reaction-convection-diffusion equation namely Murray equation. The (G'/G)-expansion method is based on the explicit linearization of nonlinear evolution equations for traveling waves with a certain substitution which leads to a second-order differential equation with constant coefficients [3, 4]. Through the use of the method we can obtain more general solutions with some free parameters. Moreover, the (G'/G)-expansion method transforms a nonlinear equation to a simple algebraic system of equations which can be solved easily by means of a symbolic computational software like *Maple*. In this paper, we use *Maple*15.

2 Description of the (G'/G)-expansion method

The application of the (G'/G)-expansion method is as follows[4]: At first the traveling wave variable $\xi = kx + wt$ is used to reduce the following partial differential equation

$$P(u, u_t, u_x, u_{tt}, u_{xt}, u_{xx}, \ldots) = 0, \tag{1}$$

into an ordinary differential equation for $U = U(\xi)$ of the form

$$Q(U, -wU', kU', w^2U'', k^2U'', kwU'', \ldots) = 0. \tag{2}$$

After solving the ODE (2), we have some traveling wave solutions $u(x,t) = U(\xi)$ of the PDE (1). To solve the ODE suppose it is assumed that the solution of the ODE (2) can be expressed as a polynomial in (G'/G) as follows:

$$U(\xi) = \sum_{i=0}^{n} \alpha_i (\frac{G'}{G})^i, \tag{3}$$

where $G = G(\xi)$ satisfies the second-order linear ODE in the form

$$G'' + \lambda G' + \mu G = 0. \tag{4}$$

where $\alpha_i, i = 0 \ldots n$, λ and μ are constants to be determined and the leading coefficient $\alpha_n \neq 0$. The positive integer n can be determined by considering the homogeneous balance between the highest order derivatives and nonlinear terms appearing in (2). Substituting (3) into (2) and using (4), and then collecting all terms with the same power of (G'/G) together, the left-hand side of Eq. (2) is converted into a polynomial in (G'/G). By equating each coefficient of the resulted polynomial to zero, it is obtained a system of algebraic equations for $\alpha_i, k, w, \lambda, \mu$ that can be solved by a computational algebraic system (CAS) like Maple. Since the general solutions of the ODE (4) are well known for us, then substituting α_i, k, w and the general solutions of Eq. (4) into Eq. (3), we have more traveling wave solutions of the nonlinear evolution equation (1).

Remark 2.1. *The function $G(\xi)$ which is the solution of Eq. (4) has the property that different order derivatives of the function (G'/G) can be expressed as a second order polynomial with respect to (G'/G). In fact, using (4) we have*

$$\frac{d(\frac{G'}{G})}{d\xi} = -(\mu + \lambda(\frac{G'}{G}) + (\frac{G'}{G})^2), \tag{5}$$

and so we have the following derivatives of $U(\xi)$

$$U' = \frac{dU}{d\xi} = -(\mu + \lambda(\frac{G'}{G}) + (\frac{G'}{G})^2)\frac{dU}{d(\frac{G'}{G})}, \tag{6}$$

$$U'' = \frac{d^2U}{d\xi^2} = (\mu + \lambda(\frac{G'}{G}) + (\frac{G'}{G})^2)^2\frac{d^2U}{d(\frac{G'}{G})^2} + (\lambda + 2(\frac{G'}{G}))(\mu + \lambda, (\frac{G'}{G}) + (\frac{G'}{G})^2)\frac{dU}{d(\frac{G'}{G})}. \tag{7}$$

etc. So by considering $U(\xi)$ as Eq. (3), it follows that its derivatives are polynomials of (G'/G).

Remark 2.2. *The general solution $G(\xi)$ of Eq. (4) required in the last step has an explicit form and so the expression G'/G required in step 4. is as follows*

$$\frac{G'(\xi)}{G(\xi)} = \begin{cases} \frac{\sqrt{\lambda^2-4\mu}}{2}(\frac{C_1 \sinh\frac{1}{2}\sqrt{\lambda^2-4\mu}\xi + C_2\cosh\frac{1}{2}\sqrt{\lambda^2-4\mu}\xi}{C_1\cosh\frac{1}{2}\sqrt{\lambda^2-4\mu}\xi + C_2\sinh\frac{1}{2}\sqrt{\lambda^2-4\mu}\xi}) - \frac{\lambda}{2}, & \lambda^2 - 4\mu > 0, \\ \frac{\sqrt{4\mu-\lambda^2}}{2}(\frac{-C_1\sin\frac{1}{2}\sqrt{4\mu-\lambda^2}\xi + C_2\cos\frac{1}{2}\sqrt{4\mu-\lambda^2}\xi}{C_1\cos\frac{1}{2}\sqrt{4\mu-\lambda^2}\xi + C_2\sin\frac{1}{2}\sqrt{4\mu-\lambda^2}\xi}) - \frac{\lambda}{2}, & \lambda^2 - 4\mu < 0. \end{cases} \tag{8}$$

3 Application of the method

Consider the following reaction-convection-diffusion equation of the form

$$u_t = (\lambda + \lambda_0 u)u_{xx} + \lambda_1 uu_x + \lambda_2 u - \lambda_3 u^2, \tag{9}$$

where $\lambda, \lambda_0, \lambda_1, \lambda_2, \lambda_3$ are real constants [5]. In the particular case $\lambda = 1$, $\lambda_0 = 0$, this equation coincides with the Murray equation

$$u_t = u_{xx} + \lambda_1 u u_x + \lambda_2 u - \lambda_3 u^2, \tag{10}$$

Which itself is a generalization of the well-known Fisher and Burgers equations. We introduce the traveling wave variable $u(x,t) = U(\xi)$, $\xi = kx + wt$ into Eq. (10) to find

$$wU' - k^2 U'' - \lambda_1 k U U' - \lambda_2 U + \lambda_3 U^2 = 0, \tag{11}$$

where prime denotes the derivatives with respect to ξ. Considering the homogeneous balance between the highest linear term U'' and the nonlinear term UU' in Eq. (11), the parameter n is determined. By doing so, we have $n = 1$ and so

$$U(\xi) = \alpha_0 + \alpha_1 (G'/G), \tag{12}$$

where α_i, $i = 0, 1$ are constants to be determined later. Substituting Eq. (12) along with (6) and (7) into Eq. (11) and collecting all terms with the same power of (G'/G) together, the left hand side of Eq. (11) is converted into a polynomial in (G'/G). Equating each coefficient to be zero yields a set of simultaneous algebraic equations where the nontrivial traveling wave solutions are as follow:

$$\begin{aligned} &k = \pm\frac{\lambda_1\lambda_2}{2\lambda_3\sqrt{\lambda^2-4\mu}}, \quad w = \pm\frac{\lambda_2(\lambda_1^2\lambda_2 + 4\lambda_3^2)}{4\lambda_3^2\sqrt{\lambda^2-4\mu}}, \\ &\alpha_0 = \frac{\lambda_2}{2\lambda_3}(1 \pm \frac{\lambda}{\sqrt{\lambda^2-4\mu}}), \quad \alpha_1 = \pm\frac{\lambda_2}{\lambda_3\sqrt{\lambda^2-4\mu}}, \end{aligned} \tag{13}$$

Case 1 when $\lambda^2 - 4\mu > 0$ by substituting (13) into Eq. (12) and using (8) the hyperbolic solutions are obtained as follows

$$\begin{aligned} U^{\pm}_{hyper}(\xi) &= \frac{\lambda_2}{2\lambda_3}(1 \pm \frac{\lambda}{\sqrt{\lambda^2-4\mu}}) \\ &\pm \frac{\lambda_2}{\lambda_3\sqrt{\lambda^2-4\mu}}(\frac{\sqrt{\lambda^2-4\mu}}{2}(\frac{C_1 \sinh\frac{1}{2}\sqrt{\lambda^2-4\mu}\xi + C_2\cosh\frac{1}{2}\sqrt{\lambda^2-4\mu}\xi}{C_1\cosh\frac{1}{2}\sqrt{\lambda^2-4\mu}\xi + C_2\sinh\frac{1}{2}\sqrt{\lambda^2-4\mu}\xi}) - \frac{\lambda}{2}), \end{aligned} \tag{14}$$

where $\xi = (\pm\frac{\lambda_1\lambda_2}{2\lambda_3\sqrt{\lambda^2-4\mu}})x + (\pm\frac{\lambda_2(\lambda_1^2\lambda_2+4\lambda_3^2)}{4\lambda_3^2\sqrt{\lambda^2-4\mu}})t$ and C_1, C_2 are arbitrary constants. So we get different solutions $u(x,t) = U(\xi)$.

Case 2 when $\lambda^2 - 4\mu < 0$ by substituting (13) along with (8) into (12), we have the following trigonometric solution

$$\begin{aligned} U^{\pm}_{trig}(\xi) &= \frac{\lambda_2}{2\lambda_3}(1 \pm \frac{\lambda}{\sqrt{\lambda^2-4\mu}}) \\ &\pm \frac{\lambda_2}{\lambda_3\sqrt{\lambda^2-4\mu}}(\frac{\sqrt{4\mu-\lambda^2}}{2}(\frac{-C_1\sin\frac{1}{2}\sqrt{4\mu-\lambda^2}\xi + C_2\cos\frac{1}{2}\sqrt{4\mu-\lambda^2}\xi}{C_1\cos\frac{1}{2}\sqrt{4\mu-\lambda^2}\xi + C_2\sin\frac{1}{2}\sqrt{4\mu-\lambda^2}\xi}) - \frac{\lambda}{2}) \end{aligned} \tag{15}$$

Where the traveling wave is as

$$\xi = \pm\frac{\lambda_1\lambda_2}{2\lambda_3\sqrt{\lambda^2-4\mu}}x \pm \frac{\lambda_2(\lambda_1^2\lambda_2+4\lambda_3^2)}{4\lambda_3^2\sqrt{\lambda^2-4\mu}}t$$

and C_1, C_2 are arbitrary constants. So we get different traveling wave solutions $u(x,t) = U(\xi)$ also in this case.

4 Conclusions

In this paper, the (G'/G) expansion method is successfully applied for obtaining exact traveling wave solutions of a reaction-convection-diffusion equation i.e. Murray biological model. It is shown that the (G'/G)-expansion method is quiet efficient and well suited for finding exact solutions. The reliability of the method and reduction in the size of computational domain give this method a wider applicability. With the aid of Maple and by putting them back into the original equation, we have assured the correctness of the obtained solutions.

References

[1] M.L. Wang, Y.B. Zhou, *Application of a homogeneous balance method to exact solutions of nonlinear equations in mathematical physics*, Phys. Lett. A 216 (1996) pp. 67–75.

[2] Y.B. Zhou, M.L. Wang, *The periodic wave solutions and solitary wave solutions for a class of nonlinear partial differential equations*, Phys. Lett. A 323 (2004) pp. 77–88.

[3] M. Wang, X. Li, (G'/G)*-expansion method and traveling wave solutions of nonlinear evolution equations in mathematical physics*, Phys. Lett. A 372 (2008) pp. 417–423.

[4] M. Wang, J. Zhang, *Application of the* (G'/G)*-expansion to travelling wave solutions of the Broer-Kaup and the approximate long water wave equations*, Appl. Math. Comput. 206 (2008) pp. 321–326.

[5] A. Yildirim, Z. Pinar, *Application of the exp-function method for solving nonlinear reaction-diffusin equations arising in mathematical biology*, Comp. Math. Appl. 60 (2010) pp. 1873-1880.

Email:javadi@tmu.ac.ir

Email:mostafa.jani@tmu.ac.ir

Email:javadtayyebi@alum.sharif.edu

43rd Annual Iranian Mathematics Conference,
27-30 August 2012
University of Tabriz

The inverse 1-median problem on a plane

Mohammadreza Galavii

Zabol University

Abstract
Given n points in the plane with nonnegative weights, the inverse Euclidean 1-median problem consists in changing the weights at minimum cost such that a prespecified point in the plane becomes the Euclidean 1-median. We derive a purely combinatorial algorithm which solves the inverse Euclidean 1-median problem with unit cost in $O(n \log n)$ time.
Keywords: Location problem, Inverse location problem, Fermat-Weber problem.

1 Introduction

In recent years inverse optimization problems found an increased interest. The *inverse optimization problem* consists in changing parameters of the problem at minimum cost such that a prespecified solution becomes optimal. Recently, inverse p-median problem has been investigated by Burkard, Pleschiutschnig and Zhang [2]. They showed that the discrete inverse p-median problem with real weights can be solved in polynomial time provided p is fixed and not an input parameter. They presented a greedy-like $O(n \log n)$ time algorithm for the inverse 1-median problem in the plane provided the distances between the points are measured in the Manhattan or maximum metric. Galavii [3] showed that the inverse 1-median problem on a tree with positive weights can be solved in linear time. In this paper we investigate the inverse Euclidean 1-median problem in the plane.

2 The Euclidean 1-median problem on a plane and its inverse

Given n points $P_1, P_2, ..., P_n$ in a metric space (X, d) and positive weights $w_1, w_2, ..., w_n$ the 1-*median problem* asks for a point $P \in X$ which minimizes

$$\sum_{i=1}^{n} w_i d(P_i, P).$$

In the *inverse* 1-*median problem* a point P_0 is given in addition to the points $P_1, P_2, ..., P_n$. The weight of these points have to be modified within given bounds $[\underline{w}_i, \overline{w}_i]$ such that P_0 becomes a 1-median and the sum of weight changes [or: the cost for the weight changes] is as small as possible.

Definition 2.1. *If $P_0 \neq P_i$ for all $i = 1, 2, ..., n$ the resultant force $R(P_0)$ at P_0 given by:*

$$R(P_0) := \sum_{i=1}^{n} \frac{w_i}{d(P_i, P_0)}(P_i - P_0);$$

and for $P_i = P_j$ for some $j = 1, 2, ..., n$,

$$R(P_0) := \max(\|R_j\| - w_j, 0)\frac{R_j}{\|R_j\|\|}.$$

Differential calculus tells that P_0 is a 1-median iff the resultant $R(P_0) = 0$. In case that $P_i \neq P_j$ we can assume that P_0 lies in the interior of the convex hull of the points P_i, $i = 1, 2, ..., n$ and that P_0 is the origin. Next we project the given points on the unit circle. The point P_0=(0,0) is a Euclidean 1-median if and only if

$$R_x(w) := \sum_{i=1}^{n} w_i x_i = 0, \tag{1}$$

$$R_y(w) := \sum_{i=1}^{n} w_i y_i = 0. \tag{2}$$

Since the Euclidean distance is invariant with respect to rotation and reflection, we can always assume that

$$R_x(w) = 0,$$
$$R_y(w) \leq 0.$$

If $R_y(w) = 0$, then the weights w_i, $i = 1, 2, ..., n$, provide an optimal solution. Therefore we assume in the following $R_y(w) < 0$. We call $\mid R_y(w) \mid$ the *optimality gap* $G(w)$. The solution method is based on a sequence of weight changes. If by chance one of the given points coincides with $A := (0, 1)$ or $B := (0, -1)$, then we can decrease $G(w)$ by changing the weight of this point without violating $R_x(w) = 0$. Otherwise we reduce the optimality gap by simultaneously changing the weights of two points, say point P_s and point P_t. If we want to decrease the optimality gap by δ, the weight change δ_s of point P_s and δ_t of point P_t have to fulfill according to (1) and (12):

$$x_s\delta_s + x_t\delta_t = 0,$$
$$y_s\delta_s + y_t\delta_t = \delta.$$

Then Cramer's rule yields

$$\delta_s = -\frac{x_t}{x_s y_t - x_t y_s}\delta,$$
$$\delta_t = \frac{x_s}{x_s y_t - x_t y_s}\delta.$$

δ is to be chosen as large as possible such that $\delta \leq G(w)$ and the weight bounds for w_s and w_t are fulfilled. The maximal possible value of δ is called the augmentation value δ_{st}. If $\delta_{st} > 0$, we call (P_s, P_t) an *augmenting pair*. The cost of an augmentation δ by the pair (P_s, P_t) is given by $\mid \delta_s \mid + \mid \delta_t \mid$. We can evaluate the efficiency of the weight change incurred by the augmenting pair (P_s, P_t) by defining the *efficiency* e_{st} as fraction of the gain in closing the optimality gap divided by costs:

$$e_{st} := \frac{\delta}{\mid \delta_s \mid + \mid \delta_t \mid}.$$

An augmenting pair with maximum efficiency is called *maximal augmenting pair*. For each point $P_i = (x_i, y_i)$ let $\frac{y_i}{x_i}$ denote the *slop* of P_i. Then (P_s, P_t) is a maximal augmenting pair, if P_s has maximum slop and P_t has a minimum slop. If an augmenting pair with maximal efficiency is used for a pair exchange then there exists an optimal solution which can be obtained without revoking the weight increase or decrease of this weight transformation. This leads to a greedy-like algorithm. The algorithm terminates after at most n weight exchanges. It yields a solution where at most two of the change weights lie strictly between their lower and upper bound. This method solves the unit cost model in $O(n \log n)$ time. Successively choosing maximal augmenting pairs in case of arbitrary linear cost for the weight changes, however, does yield an optimal solution.

Now we consider the case that the prespecified point coincides with one of the given points. According to the optimality condition point P_0 is 1-median if and only if

$$R_x^2(w) + R_y^2(w) \leq w_0^2$$

holds. This condition does not lead to a convex problem. However, it is possible to fix the optimal weight w_0^* in advance to

$$w_0^* = \min\{\overline{w}_0, \sqrt{R_x^2(w) + R_y^2(w)}\}.$$

The remaining problem is convex and can be solved by any algorithm for convex programming.

References

[1] R. E. Burkard, M. Galavii and E. Gassner, *The inverse Fermat-Weber problem*, European Journal of Operational Research, 206 (2010), pp. 11–17.

[2] R. E. Burkard, C. Pleschiutschnig and J. Z. Zhang, *Inverse p-median problems*, J. Discrete Optimization, 1 (2004), pp. 23–39.

[3] M. Galavii, *Inverse* 1-*Median Problems*, Ph.D. Thesis, Institute of Optimization and Discrete Mathematics, Graz University of Technology, Graz, Austria, 2008.

Email:galavi@opt.math.tu-graz.ac.at

43rd Annual Iranian Mathematics Conference,
27-30 August 2012
University of Tabriz

Some properties of Laplacian matrix of saturated hydrocarbons

A. M. Nazari
Arak University

M. Ahmadi
Arak University

Abstract

In this paper we study the adjacency and Laplacian matrices of graph of saturated hydrocarbons C_nH_{2n+2} for $n = 1, 2, \ldots$. We find a bound for spectrum of Laplacian matrix of C_nH_{2n+2} better than bound which is obtained from Greshgoorin Theorem.

Keywords: Saturated hydrocarbon, eigenvalues, Laplacian matrix

Mathematics Subject Classification: 15A18; 05C50; 05C05.

1 Introduction

Let G be the graph of saturated hydrocarbon C_nH_{2n+2} for a given integer number n. We denote the Laplacian matrix associated to G by $L(G) = D(G) - A(G)$, where $D(G)$ is diagonal matrix with entries $(d_1, d_2, ..., d_n)$, which d_i is degree of v_i of G, and $A(G)$ is adjacency matrix of G. We have $L(G) = MM^T$, where M is event matrix of G.

In this paper at first we recall the algorithm of Jacobs and Trevisan and by this algorithm we find a new bound for spectrum of $L(G)$.

1.1 Algorithm of Jacobs and Trevisan

This algorithm for a given tree T, and given interval J compute the number of eigenvalues of $L(T)$ that lies in interval J, and based on diagonal entries of the matrix $L(T) + \alpha I$, where $L(T)$ is Laplacian matrix and $\alpha \in \mathbb{R}$ and I is identity matrix. We can use straightforward this algorithm on a tree without the matrices $L(T)$ and $A(T)$. Now recall the Jacobs and Trevisan algorithm from [2].

Input: Tree T, scalar α

Output: diagonal matrix D congruent to $L(T) + \alpha I$

Algorithm Diagonal of (T, α)

  Initialize $a(v) := d(v) + \alpha$, for all vertices v
  order vertices bottom up
  for $k = 1$ to n
    if v_k is a leaf then continue
    else if $a(c) \neq 0$ for all children c of v_k then
      $a(v_k) := a(v_k) - \sum \frac{1}{a(c)}$, summing over all children of v_k
    else
      select one children v_j of v_k for which $a(v_j) = 0$
      $d(v_k) := -\frac{1}{2}$

$d(v_j) := 2$
if v_k has a parent v_l, remove the edge v_kv_l.
end loop

Lemma 1.1. [1] *Let T be a tree and let D be the diagonal matrix that generated by diagonalizing algorithm $(T, -\alpha)$, then the following proposition is satisfied.*
a) *The number of positive entries of D, is the number of eigenvalues of $L(T)$ which is greater than of α.*
b) *The number of negative entries of D, is the number of eigenvalues of $L(T)$ which is less than of α.*
c) *If there is j zero entries in main diagonal of D, then α is the eigenvalue of T with multiplicity j.*

2 The eigenvalues of Laplacian matrix of saturated hydrocarbons

The largest element of Laplacian matrices of saturated hydrocarbons in each row is 4, and this element lies in main diagonal. The Laplacian matrix of saturated hydrocarbons is diagonally dominant. Since the sum of each rows and columns is zero, then by Greshgoorin Theorem, for all eigenvalue λ, we have $|\lambda| \leq 8$.
Consider saturated hydrocarbons C_nH_{2n+2}, We apply algorithm Jacobs - Trevisan, then for $\alpha = -1$ we have Figure 1, therefore there is $n+2$ zero in the vertices of this graph, and this show that the eigenvalue 1 has multiplicity $n+2$. The n another positive number of Figure 1, shows n eigenvalues of Laplacian matrix greater than 1 and n negative numbers shows the number of eigenvalues which are less than 1 equal n.
Now for finding an upper bound for eigenvalues of $L(T)$ associated to saturated hydrocarbons, by checking and correction we set $\alpha = -6.5$, then we have the graph of Figure 3 by Jacobs - Trevisan algorithm.
Now from left to right we calculate the number of each vertices by implementing algorithm:
-The first carbon atom $x_1 = -\frac{5}{2} - (-\frac{2}{11} - \frac{2}{11} - \frac{2}{11})$

-The second carbon atom $x_2 = -\frac{5}{2} - (-\frac{2}{11} - \frac{2}{11} + \frac{1}{x_1})$
-The 3th carbon atom $x_3 = -\frac{5}{2} - (-\frac{2}{11} - \frac{2}{11} + \frac{1}{x_2})$
⋮ ⋮
-The $(n-1)$th carbon atom $x_{n-1} = -\frac{5}{2} - (-\frac{2}{11} - \frac{2}{11} + \frac{1}{x_{n-2}})$
-The (n)th carbon atom $x_n = -\frac{5}{2} - (-\frac{2}{11} - \frac{2}{11} + \frac{1}{x_{n-1}})$
Now by considering induction we show that all x_i are nonnegative.

$$x_1 = -\frac{5}{2} + \frac{6}{11} < -1$$

let $x_k < -1$, we show that x_{k+1} also less than 1.
Since $x_k < -1$ then $\frac{1}{x_k} > -1$, so $-\frac{1}{x_k} < +1$,

$$x_{k+1} = -\frac{5}{2} + \frac{4}{11} - \frac{1}{x_k} < -\frac{5}{2} + \frac{4}{11} + 1 = -\frac{25}{22} < -1$$

and this show that all $(3n+2)$ eigenvalues of $L(G)$ less than 6.5.
If we apply algorithm Jacobs-Trevisan for $\alpha = -6$ we find some positive number of vertices. Then we think 6.5 is a good choice for the bound of eigenvalues of $L(G)$.

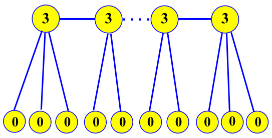

Figure 1:

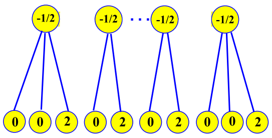

Figure 2:

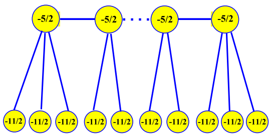

Figure 3:

References

[1] D. P. JACOBS AND V. TREVISAN, *Locating the eigenvalues of a tree*, Linear Algebra Appl. **434** (2011) 81-88.

[2] E. FRITSCHER, C. HOPPEN, I. ROCHA, V. TREVISAN, *On the sum of the laplacian eigenvalues of a tree*, Linear Algebra Appl. **435** (2011) 371-399.

[3] D. P. JACOBS,V. TREVISAN , *The determinant of a trees neighborhood matrix*, Linear Algebra Appl. **256** (1997) 235-249.

[4] D. P. JACOBS, C.M.S.MACHADO,V. TREVISAN, *An $O(n^2)$ algorithm for the characteristic polynomial of a tree*, J. Combin. Math. Combin. Comput. **54** (2005) 213-221.

Email:a-nazari@araku.ac.ir

Email:m-ahmadi@arashad.araku.ac.ir