Journal of Mathematical Extension
Vol. 3, No. 1 (2008), 27-42

On a New Generalization of the Exponential Distribution

S. M. Mirhossaini and A. Dolati
Yazd University

Abstract. This paper introduces a two-parameter family of distributions which includes the ordinary exponential distribution as a special case. This distribution exhibits monotone hazard rate and may be a competitor to the families of two parameter gamma and Weibull distributions. Various statistical and reliability aspects of this model is explored. Several numerical examples based on real data show the flexibility of the new distribution for modeling proposes.

AMS Subject Classification: Primary 60E05; secondary 62H20.
Keywords and Phrases: Exponential distribution, gamma distribution, Hazard rate, lifetime data, Weibull model.

1. Introduction

There are many lifetime distributions that have been proposed in the literature. Exponential distributions because of their useful properties and convenient theory, play a central role in analyses of lifetime data. In situations where the ordinary exponential distribution is not sufficiently broad, a number of wider families such as the gamma and Weibull models are in common use. Both distributions have the increasing as well as the decreasing failure rate. This gives an extra edge over the exponential distribution which has only constant failure rate. A complete treatment of these distributions is given by Johnson et al. [12, Chapters 17, 19 and 21]. In recent years, there have been some attempts to provide exponential-based alternatives to the gamma and Weibull distributions to increase the flexibility for modeling purposes; see, e.g., Mudholkar and Srivastava [15], Marshall and Olkin [14], Adamidis and Loukas [1], Gupta and Kundu [10], Adamidis et al. [2] and Nadarajaha & Kotz [16]. In this paper we introduce a family of two-parameter univariate distributions having both decreasing and increasing failure rate which includes the exponential distribution as a special case. This model is obtained using a certain mixture of the order statistics of a sample of size 2 from ordinary exponential distribution. The study examines various properties of this distribution. The paper is organized as follows: Section 2, discusses a general system of univariate distributions. Section 3, introduces the new generalization of the exponential distribution and presents its basic properties including the behavior of the density and hazard rate functions, expressions for the moments and related measures, the distribution of the sums and a characterization based on its mean residual life function. Section 4, devoted to the estimation of the parameters. We also provided several numerical examples in this section, where the new two-parameter exponential distribution fits better than the ordinary exponential, gamma and Weibull models.

2. A General System of Univariate Distributions

Let X_1 and X_2 be two independent and identically distributed random variables having the survival function $\bar{F} = 1 - F$. For $-1 \leqslant \alpha \leqslant 1$, consider the random variable U defined by

$$U = \begin{cases} X_{1:1}, & \text{with probability } \frac{1+\alpha}{2} \\ X_{2:2}, & \text{with probability } \frac{1-\alpha}{2}, \end{cases} \tag{1}$$

where $X_{1:1} = \min(X_1, X_2)$ and $X_{2:2} = \max(X_1, X_2)$ are the corresponding order statistics of X_1 and X_2. Since the distribution functions of $X_{2:2}$ and $X_{1:1}$ are given by $F_{2:2}(x) = F^2(x)$ and $F_{1:1}(x) = 2F(x) - F^2(x)$, it is then a simple exercise to show that the random variable U has the distribution function

$$G(x;\alpha) = F(x)\{1 + \alpha\bar{F}(x)\}, \quad -\infty < x < \infty, \ -1 \leqslant \alpha \leqslant 1. \tag{2}$$

Clearly, for $\alpha = 0$, we get $G = F$. When $\alpha = -1$, $G = F_{2:2}$, and when $\alpha = 1$, $G = F_{1:1}$. Since $G(.;\alpha)$ is increasing in α, we have the inequality

$$F_{2:2}(x) \leqslant G(x;\alpha) \leqslant F_{1:1}(x),$$

for all x and $-1 \leqslant \alpha \leqslant 1$.

Remark 1. *Note that if H is a distribution function, by solving the equation $F(x)\{1+\alpha\bar{F}(x)\} = H(x)$ for F, we obtain*

$$F(x;\alpha) = \frac{1+\alpha-\sqrt{(1+\alpha)^2-4\alpha H(x)}}{2\alpha}, \tag{3}$$

which is also a distribution function for any $-1 \leqslant \alpha \leqslant 1$.

Remark 2. *The equation constitute the system of distributions defined by (2), could be seen as a univariate version of the well-known Farlie-Gumbel-Morgenstern (FGM, for short) family of bivariate distributions, which is usually written as*

$$H(x,y) = F_1(x)F_2(y)\{1+\alpha\bar{F}_1(x)\bar{F}_2(y)\},$$

with $\alpha \in [-1,1]$; see, Drouet-Mari and Kotz ([8, Chapter 5]) for a good review.

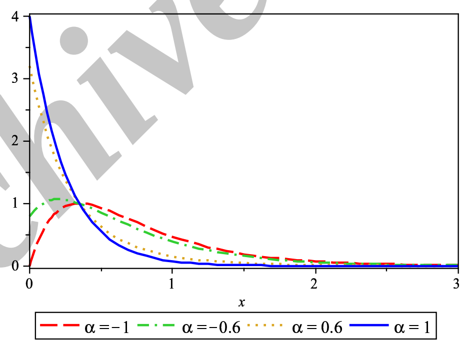

Fig.1. Density function of ME distribution for $\theta = 2$ and $\alpha = -1, -0.6, -0.3, 0, 0.3, 0.6, 1$.

3. A Family of Two-Parameter Exponential Distributions

The particular case that F is an exponential distribution with the parameter θ, (2) yields a new two-parameter family of distributions given by

$$G(x;\alpha,\theta) = (1-e^{-\theta x})(1+\alpha e^{-\theta x}), \quad x,\theta > 0, \ -1 \leqslant \alpha \leqslant 1. \tag{4}$$

The corresponding density function is

$$g(x;\alpha,\theta) = \theta e^{-\theta x}\{1+\alpha(2e^{-\theta x}-1)\}. \tag{5}$$

We say that the random variable X has a '*mixture exponential*' (ME) with the shape parameters α and the scale parameter θ, denoted by $X \sim$ME(α,θ), if X has the density function defined by (5).

3.1. Shape of the Density Function

Proposition 1. *The function* $\log g(.;\alpha,\theta)$, *is concave for* $-1 \leqslant \alpha \leqslant 0$ *and convex for* $0 \leqslant \alpha \leqslant 1$.

Proof. The result follows by observing that the second derivative of the logarithm of the density function with respect to x is

$$\frac{d^2}{dx^2}\log g(x;\alpha,\theta) = \frac{2\alpha(1+\alpha)\theta^2 e^{-\theta x}}{(1+\alpha(2e^{-\theta x}-1))^2}\ .$$

As a result of Proposition 1, for $0 \leqslant \alpha \leqslant 1$, $g(x.;\alpha,\theta)$ is decreasing, and for $-1 \leqslant \alpha < 0$, $g(x.;\alpha,\theta)$ is unimodal. By solving $d\log g(x;\alpha,\theta)/dx = 0$, it is readily verified that a random variable X with the density $g(.;\alpha,\theta)$ has the mode

$$\text{mode}(X) = \begin{cases} 0, & \alpha \geqslant \frac{-1}{3} \\ \frac{-1}{\theta}\ln\left(\frac{\alpha-1}{4\alpha}\right), & \alpha \leqslant \frac{-1}{3} \end{cases}. \tag{6}$$

ME probability density functions are displayed in Fig. 1, for selected values of α and θ. □

3.2. Moments and other Measures

If $X \sim$ME(α,θ), then the generating function (m.g.f.) of X defined by $M(t) = E(e^{tX})$, is given by

$$M(t) = \frac{\theta\{2\theta - (1+\alpha)t\}}{(\theta - t)(2\theta - t)}. \tag{7}$$

By straightforward integration the raw moments of X about the origin are found to be

$$E(X^r) = \frac{(1+\alpha(2^{-r}-1))r!}{\theta^r}, \tag{8}$$

for $r \in N$. By using the identity $E(X-\mu)^k = \sum_{r=0}^{k} \binom{n}{k} E(X^r)(-\mu)^{k-r}$, the central moments can be obtained. In particular:

$$\begin{aligned} \text{Var}(X) &= \frac{4-2\alpha-\alpha^2}{4\theta^2}, \\ E(X-\mu)^3 &= \frac{8-3\alpha-3\alpha^2-\alpha^3}{4\theta^3}, \\ E(X-\mu)^4 &= \frac{144-72\alpha-48\alpha^2-12\alpha^3-3\alpha^4}{16\theta^4}. \end{aligned}$$

The coefficient of variation (γ), skewness ($\sqrt{\beta_1}$) and kurtosis (β_2) of X are given by

$$\gamma = \frac{\sqrt{4-2\alpha-\alpha^2}}{2-\alpha}, \tag{9}$$

$$\sqrt{\beta_1} = \frac{2(8-3\alpha-3\alpha^2-\alpha^3)}{(4-2\alpha-\alpha^2)^{\frac{3}{2}}}, \tag{10}$$

$$\beta_2 = \frac{144-72\alpha-48\alpha^2-12\alpha^3-3\alpha^4}{(4-2\alpha-\alpha^3)^2}\ . \tag{11}$$

The coefficients γ, $\sqrt{\beta_1}$ and β_2, are independent of the scale parameter θ, and for the exponential distribution they are given by 1, 2 and 9, respectively. The coefficient of variation is less than 1 for $\alpha < 0$ and is greater than 1 for $\alpha > 0$.

By using the relation (3), we obtain the qth quantile x_q of ME distribution as

$$x_q = \frac{-1}{\theta}\ln\left(\frac{\alpha-1+\sqrt{(1+\alpha)^2-4\alpha q}}{2\alpha}\right). \tag{12}$$

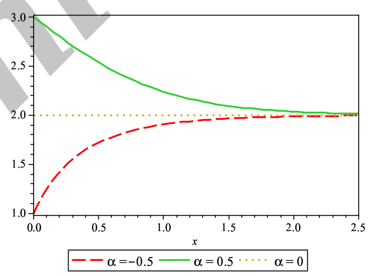

Fig2. Hazard rate function of ME distribution for $\theta = 2$ and $\alpha = -1, -0.6, -0.3, 0, 0.3, 0.6$.

In particular, the median of X is given by

$$\text{med}(X) = \frac{-1}{\theta} \ln\left(\frac{\alpha - 1 + \sqrt{1+\alpha^2}}{2\alpha}\right). \quad (13)$$

Note that if $\alpha \to 0$, then $x_q \to \frac{-1}{\theta}\ln(1-q)$, which is the qth quantile of the exponential distribution.

It is easy to see that med(X), mode(X) and $E(X)$ are all decreasing in α and θ. It also follows that mode$(X) \leqslant$ med$(X) \leqslant E(X)$.

3.3. Reliability Properties

The reliability (or survival) function corresponding to the ME distribution is given by

$$\bar{G}(x;\alpha,\theta) = e^{-\theta x}\{1 + \alpha(e^{-\theta x} - 1)\}, \quad (14)$$

and thus the ME failure rate (also known as hazard rate) function is

$$\begin{aligned} h(x;\alpha,\theta) &= g(x;\alpha,\theta)\{\bar{G}(x;\alpha,\theta)\}^{-1} \quad (15)\\ &= \frac{\theta\{1+\alpha(2e^{-\theta x} - 1)\}}{1+\alpha(e^{-\theta x} - 1)}. \end{aligned}$$

Proposition 2. *The function $h(.;\alpha,\theta)$ is increasing in x for $-1 \leqslant \alpha < 0$, constant for $\alpha = 0$ and decreasing for $0 < \alpha \leqslant 1$.*

Proof. The proof follows using the log-convexity and the log-concavity of the density function of the ME distribution [4]. □

It can be verified that $h(., \alpha, \theta)$ is increasing in α and decreasing in θ. Thus, the ME distribution is positively, (resp, negatively) ordered with respect to α (resp, θ) according to the hazard rate ordering. Note that, $\lim_{x\to 0} h(x; \alpha, \theta) = (1+\alpha)\theta$. Therefore, at the origin this hazard rate varies continuously with the parameters. This is in contrast with the families of Weibull or gamma distributions; for both of those families, $h(0) = 0$, or $h(0) = \infty$, so that $h(0)$ is discontinuous in the parameters. For ME distribution, $\lim_{x\to\infty} h(x; \alpha, \theta) = \theta$, is bounded and continuous in the parameters, like gamma distribution but unlike the Weibull distribution. From Proposition 2., it follows that

$$(1+\alpha)\theta \leqslant h(x; \alpha, \theta) \leqslant \theta \quad (-1 \leqslant \alpha \leqslant 0),$$

and

$$\theta \leqslant h(x; \alpha, \theta) \leqslant (1+\alpha)\theta \quad (0 \leqslant \alpha \leqslant 1).$$

Fig 2. shows the hazard rate function of the ME(α, θ), for selected values of α and θ.

Given that there was no failure prior to time t, the residual life distribution of the random variable X, distributed as ME distribution with the parameters α and θ, has the survival function

$$\begin{aligned} \bar{G}_t(x; \alpha, \theta) &= P\{X > x + t \mid X > t\} \\ &= \frac{e^{-\theta(x+t)}\{1 + \alpha(e^{-\theta(x+t)} - 1)\}}{e^{-\theta t}\{1 + \alpha(e^{-\theta t} - 1)\}} \\ &= e^{-\theta x}\{1 + \beta(e^{-\theta x} - 1)\} \\ &= \bar{G}(x; \beta, \theta), \end{aligned} \tag{16}$$

where $\beta = \beta(t) = \alpha e^{-\theta t}\{1 + \alpha(e^{-\theta t} - 1)\}^{-1}$. Thus, the residual life distribution of a random variable X distributed as ME(α, θ) at time t, is another ME distribution with the shape parameter depending upon time t. The limit distribution as $t \to \infty$ is an ordinary exponential distribution because the limit of $\beta(t)$ is 0.

Since distributions with an increasing (decreasing) hazard rate are 'new better (worse) than used' ([4, page 159]), it follows that, when X distributed as ME(α, θ), we have

$$\bar{G}_t(x; \alpha, \theta) \begin{cases} \leqslant \bar{G}(x; \alpha, \theta), & (-1 \leqslant \alpha \leqslant 0) \\ \geqslant \bar{G}(x; \alpha, \theta), & (0 \leqslant \alpha \leqslant 1) \end{cases} \cdot$$

By using the equality (16) and that $E(X; \alpha, \theta) = \frac{2-\alpha}{2\theta}$, when X distributed as ME(α, θ), the mean residual life function, i.e., the mean of the residual life distribution could be obtained as

$$\begin{aligned} m(t; \alpha, \theta) &= E(X - t | X > t) \\ &= \frac{2 - \beta(t)}{2\theta} \\ &= \frac{1}{\theta} \frac{1 + \alpha(\frac{1}{2}e^{-\theta t} - 1)}{1 + \alpha(e^{-\theta t} - 1)} \cdot \end{aligned} \tag{17}$$

It is easy to see that $m(t; \alpha, \theta)$, is increasing in t for $0 \leqslant \alpha \leqslant 1$ and decreasing for $-1 \leqslant \alpha \leqslant 0$, with $\lim_{t\to\infty} m(t; \alpha, \theta) = 1/\theta = E(X; 0, \theta)$ and $\lim_{t\to 0} m(t; \alpha, \theta) = (2 - \alpha)/2\theta = E(X; \alpha, \theta)$. It also follows that

$$\frac{1}{\theta} \leqslant m(t; \alpha, \theta) \leqslant \frac{2 - \alpha}{2\theta} \quad (-1 \leqslant \alpha \leqslant 0),$$

and

$$\frac{2-\alpha}{2\theta} \leqslant m(t;\alpha,\theta) \leqslant \frac{1}{\theta} \quad (0 \leqslant \alpha \leqslant 1).$$

3.4. A Characterization

The following result provides a characterization for ME distribution.

Proposition 3. *The mean residual life function, given by (17), characterizes the ME distribution.*

Proof. The derivation of (17), established the necessity. To prove the sufficiency, it is seen that (17) may be written as

$$\begin{aligned} -\frac{d}{dt}\ln\int_t^\infty \bar{G}(x)dx &= \frac{d}{dt}\ln\frac{1}{\theta}e^{-\theta t}\{1+\alpha(2^{-1}e^{-\theta t}-1)\} \\ &= \theta\{1+\alpha(e^{-\theta t}-1)\}\{1+\alpha(2^{-1}e^{-\theta t}-1)\}^{-1} \\ &= \frac{1}{m(t;\alpha,\theta)}, \end{aligned}$$

and therefore by solving for $\int_t^\infty \bar{G}(x)dx$, differentiating the result with respect to x and using $\bar{G}(0)=1$, to determine the constant, the same survival function given by (14) is obtained, which completes the proof. □

3.5. Distribution of Sums

The following result provides a mixture representation for the distribution of the sum of two independent random variables distributed according to ME distribution.

Proposition 4. *Let X_1 and X_2 be two independent random variables from ME(α_i, θ), for $i = 1, 2$. Let $T = X_1 + X_2$. Then the density function of T is given by*

$$f_T(t) = \frac{(1+\alpha_1)(1+\alpha_2)}{4} f_1(t) + \frac{(1-\alpha_1\alpha_2)}{2} f_2(t) + \frac{(1-\alpha_1)(1-\alpha_2)}{4} f_3(t),$$

where, $f_1(x) = 4\theta^2 x e^{-2\theta x}$, $f_2(x) = 4\theta e^{-\theta x}\{1 - (\theta x + 1)e^{-\theta x}\}$ and

$f_3(x) = 4\theta e^{-\theta x}\{(\theta x + 2)e^{-\theta x} + \theta x - 2\}$.

Proof. The proof is straightforward using the fact that ME random variables have representation (1). □

4. Estimation

4.1. Maximum Likelihood Estimators

Assuming a random sample of n observations, $x_1, ..., x_n$, from ME(α, θ), the log-likelihood function, $l(\alpha, \theta)$, is given by

$$l(\alpha, \theta) = n \log(\theta) - \theta \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} \log(1 + \alpha(2e^{-\theta x_i} - 1)).$$

Differentiating with respect to α and θ and equating to zero we obtain

$$\frac{\partial l}{\partial \theta} = n\theta^{-1} - \sum_{i=1}^{n} x_i - 2\alpha \sum_{i=1}^{n} x_i e^{-\theta x_i}\{1 + \alpha(2e^{-\theta x_i} - 1)\}^{-1} = 0,$$

$$\frac{\partial l}{\partial \alpha} = \sum_{i=1}^{n} (2e^{-\theta x_i} - 1)\{1 + \alpha(2e^{-\theta x_i} - 1)\}^{-1} = 0.$$

The solution of the two non-linear equations must be found using a numerical method.

The density function of ME distribution satisfies all the regularity conditions [3, pp. 86-87] and therefore applying the usual large sample approximation, the estimators $(\hat{\alpha}, \hat{\theta})$ treated as being approximately bivariate normal with the mean vector (α, θ) and variance-covariance matrix I^{-1}, where

$$I = -\frac{1}{n}\begin{bmatrix} E(\frac{\partial^2 l}{\partial \alpha^2}) & E(\frac{\partial^2 l}{\partial \alpha \partial \theta}) \\ E(\frac{\partial^2 l}{\partial \theta \partial \alpha}) & E(\frac{\partial^2 l}{\partial \theta^2}) \end{bmatrix},$$

is the inverse of the Fisher information matrix. After some algebra operations, the elements of I are given by

$$\begin{aligned} E(\frac{\partial^2 l}{\partial \theta^2}) &= -\frac{1}{\theta^2}\big(1 + 2(1-\alpha) + \frac{(1-\alpha)^2}{\alpha}\mathrm{p}(3, \frac{\alpha-1}{2\alpha})\big), \\ E(\frac{\partial^2 l}{\partial \alpha \partial \theta}) &= -\frac{1}{2\theta\alpha^2}\big((1-\alpha)\mathrm{d}(\frac{1+\alpha}{1-\alpha}) + 2\alpha\big), \\ E(\frac{\partial^2 l}{\partial \alpha^2}) &= \frac{1}{\alpha} + \frac{1}{2\alpha^3}\ln\left(\frac{1-\alpha}{1+\alpha}\right), \end{aligned}$$

where $\mathrm{p}(r,t) = \sum_{n=1}^{\infty} t^n n^{-r}$, $|t| \leqslant 1$, and $\mathrm{d}(t) = \int_0^t \frac{\ln(x)}{1-x}dx$, are the polylogarithm and dilogarithm functions, respectively ([9, page 27]), which are available in standard software such as *Mathematica* or *Maple.*

4.2. Data Analysis

Two sets of real data are considered. The first set of data involves the intervals in days between successive coal-mining disasters in Great Britan for the period 1875–1951. A disaster is defined as involving the death of 10 or more men. They were originally discussed by Maguire et al. [13] and analyzed by Cox & Lewis [6, page 4] and Adamidia & Loukas [2]. The second set of data, are survival times of 43 patients suffering from chronic granulocytic leukaemia. This data set reported by Bryson & Siddiqui [5] and reanalyzed in Hollander & Proschan [11]. In addition to the exponential and ME distributions, the gamma and Weibull distributions with respective densities $f_1(x) = \theta^\alpha x^{\alpha-1} e^{-\theta x} \{\Gamma(\alpha)\}^{-1}$, and $f_2(x) = \alpha\theta^\alpha x^{\alpha-1} e^{-(\theta x)^\alpha}$, were fitted to the data sets. The maximum likelihood estimates, the log-likelihood and the Kolmogrov-Smirnov (K-S) statistic presented in Table 1. It is observed that, the ME distribution fits marginally better than usual exponential distribution and two popular alternatives the gamma and Weibull models in both cases.

TABLE 1.

Estimates, log-likelihood and Kolmogrov-Smirnov statistic

Data Set	Distribution	α	θ	LL	K–S
$1(n=109)$	Exponential	–	0.0043	-703.3133	0.0776
	ME	0.7126	0.0028	-700.9823	0.0667
	gamma	0.8560	0.0037	-702.4007	0.0796
	Weibull	0.8847	0.0046	-701.7724	0.2965
$2(n=43)$	Exponential	–	0.0010	-336.6865	0.1162
	ME	-0.6103	0.0014	-335.2714	0.0678
	gamma	1.3027	0.0014	-335.8229	0.0852
	Weibull	1.2400	0.0010	-335.3089	0.6082

References

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

[2] K. Adamidisa, T. Dimitrakopouloub, and S. Loukas, On an extension of the exponential-geometric distribution, *Statistics and Probability Letters*, 73 (2005), 259-269.

[3] L. J. Bain, *Statistical analysis of reliability and life testing models*, New York: Marcel Dekker, Inc., 1978.

[4] R. E. Barlow and F. Proschan, *Statistical theory of reliability and life testing*, New York: Holt, Rinehart and Winston, 1975.

[5] M. C. Bryson and M. M. Siddiqui, Some criteria for aging, *Journal of the American Statistical Association*, 64 (1969), 1472-1483.

[6] D. R. Cox and P. A. W. Lewis, *The statistical analysis of series of events*, Chapman and Hall, London, 1978.

[7] H. David and H. N. Nagaraja, *Order statistics*, Wiley, Hoboken, NJ, 3rd Edition, 2003.

[8] D. Drouet-Mari and S. Kotz, *Correlation and dependence*, Imperial College Press, London, 2001.

[9] A. Erdelyi, W. Magnus, F. Oberhettinger, and F. G. Tricomi, *Higher transcendental functions*, New York: McGraw-Hill, Book Company Inc. 1953.

[10] R. D. Gupta and D. Kundu, Generalized exponential distributions, *Austral. New J. Zealand Statist.* 41 (1999), 173-188.

[11] M. Hollander and F. Proschan , Tests for the mean residual life, *Biometrika*, 62 (1975), 585-593.

[12] N. L. Johnson, S. Kotz, and N. Balakrishnan, *Continuous univariate distributions*, Vol. 1 (2nd Edition), Wiley: New York, 1994.

[13] B. A. Maguire, E. S. Pearson, and A. H. A. Wynn, The time intervals between industrial accidents, *Biometrika*, 39 (1952), 168-180.

[14] A. W. Marshall and I. Olkin, Adding a parameter to a family of distributions with application to the exponential and Weibull families, *Biometrika*, 84 (1997), 641-652.

[15] G. S. Mudholkar and D. K. Srivastava, Exponentiated Weibull family for analyzing bathtub failure-rate data, *IEEE Transactions on Reliability*, 42 (1993), 299-302.

[16] S. Nadarajaha and S. Kotz, The beta exponential distribution, *Reliability Engineering and System Safety*, 91 (2006), 689-697.

[17] M. Shaked and G. Shanthikumar, *At stochastic orders*, Springer: New York, 2007.

Seyed Mohsen Mirhossaini
Department of Statistics
Faculty of Science
Yazd University
Yazd, 89195-741, Iran

Ali Dolati
Department of Statistics
Faculty of Science
Yazd University
Yazd, 89195-741, Iran
E-mail: adolati@yazduni.ac.ir