On Record Values and Reliability Properties of Marshall–Olkin Extended Exponential Distribution

K.K. Jose[1], E. Krishna[2], and Miroslav M. Ristic[3]

[1] *Department of Statistics, Central University of Rajasthan, India-305802*

[2] *Department of Statistics, St. Joseph's College for Women, Alappuzha, Kerala, India-688001*

[3] *Faculty of Sciences and Mathematics, University of Nis, Serbia, 1800*

kkjose@curaj.ac.in, ekrishna48@gmail.com, miristic@ptt.rs

Received 12 June 2014

Accepted 22 July 2014

The Marshall–Olkin Extended Exponential distribution is introduced and reliability properties are studied. The p.d.f.'s of n^{th} record value, joint p.d.f.'s, of m^{th} and n^{th} record values are derived to obtain the expression for mean, variance and covariance of reord values. The entropy of j^{th} record value is derived.The stress strength analysis for the new model is carried out. We develop autoregressive processes and sample path properties are explored. The results are verified using simulations as well as graphical studies.The model is extended to higher orders also.

Keywords: Auto regressive processes; Entropy; Exponential distribution; Marshall–Olkin distribution; Record values; Reliability; Sample path; Simulation; Stress-strength analysis.

1. Introduction

Exponential distributions play a central role in analysis of lifetime or survival data, in part because of their convenient statistical theory, their important 'lack of memory' property and their constant hazard rates. In circumstances where the one-parameter family of exponential distributions is not sufficiently broad, a number of wider families such as the gamma, Weibull and Gompertz-Makeham distributions are in common use; these families and their usefulness are described by various authors (see Johnson, Kotz and Balakrishnan, 2004).

By various methods, new parameters can be introduced to expand families of distributions for added flexibility or to construct covariate models. Introduction of a scale parameter leads to the accelerated life model, and taking powers of the survival function introduces a parameter that leads to the proportional hazards model. For instance, the family of Weibull distributions contains the exponential distributions and is constructed by taking powers of exponentially distributed random variables. The family of gamma distributions also contains the exponential distributions, and is constructed by taking powers of the Laplace transform.

Marshall and Olkin (1997) introduced a new family of distributions in an attempt to add a parameter to a family of distributions. Let $\overline{F}(x) = P(X > x)$ be the survival function of a random

Copyright: the authors

variable X, and $\alpha > 0$ be a parameter. Then

$$\overline{G}(x,\alpha) = \frac{\alpha\overline{F}(x)}{1-(1-\alpha)\overline{F}(x)}; \quad -\infty < x < \infty, \ \ \alpha > 0 \tag{1.1}$$

is a proper survival function. The new family $\{\overline{G}(x,\ \alpha)\}$ is called Marshall–Olkin family of distributions. The p.d.f. corresponding to (1.1) is given by

$$g(x,\alpha) = \frac{\alpha f(x)}{[1-(1-\alpha)\overline{F}(x)]^2} \tag{1.2}$$

where f(x) is the p.d.f. corresponding to $F(x)$. The new hazard (failure) rate function is given by

$$h(x,\alpha) = \frac{r(x)}{1-(1-\alpha)\overline{F}(x)} \ \text{ where } \ r(x) = \frac{f(x)}{\overline{F}(x)} \tag{1.3}$$

Alice and Jose (2003, 2004 a,b) studied these in detail in the case of Pareto models.

In this paper, we introduce the Marshall–Olkin Extended Exponential distribution MOEE(α,λ) in section 2 and its properties are studied. In section 3, we discuss MOEE(α) distributions with special emphasis on record value theory. In section 4, we derive the entropy of record value distribution and entropy is calculated for various record values.In section 5, we obtain an estimate of reliability in the context of stress strength analysis and average bias,average mean square error, average confidence interval and coverage probability for the estimate is tabulated numerically for the simulated data. In section 6 we introduce first order stationary autoregressive processes with exponential marginals and the sample path properties are explored. The probability p is estimated and the standard error of the estimated value is calculated numerically by simulation.

2. Marshall–Olkin Extended Exponential Distribution

When $\overline{F}(x) = \mathrm{e}^{-\lambda x}, x \geqslant 0$, is the survival function of exponential distribution, we have the Marshall–Olkin Extended Exponential MOEE (α,λ) distribution with survival function,

$$\overline{G}(x) = \frac{\alpha}{\mathrm{e}^{\lambda x}-\overline{\alpha}}, \quad x \geqslant 0, \ \ \lambda > 0, \ \ \alpha > 0, \ \ \overline{\alpha} = 1-\alpha \tag{2.1}$$

Then the p.d.f. is

$$g(x) = \frac{\alpha\lambda \mathrm{e}^{\lambda x}}{[\mathrm{e}^{\lambda x}-\overline{\alpha}]^2}, \quad x \geqslant 0, \ \ \lambda > 0, \ \ \alpha > 0, \ \ \overline{\alpha} = 1-\alpha. \tag{2.2}$$

Direct evaluation shows that,

$$E(X) = -\frac{\alpha\log\alpha}{\lambda\overline{\alpha}}$$

The hazard rate is

$$h(x) = \frac{\lambda \mathrm{e}^{\lambda x}}{\mathrm{e}^{\lambda x}-\overline{\alpha}}, \quad x \geqslant 0, \ \ \alpha > 0. \tag{2.3}$$

The graph of $h(x)$ is drawn. It can be seen that the hazard rate is DFR for $\alpha < 1$, and IFR for $\alpha > 1$. Note that for $\alpha = 1$, $h(x) = 1$, showing constant failure rate. This establishes the wide applicability of the MOEE distribution in reliability modeling.

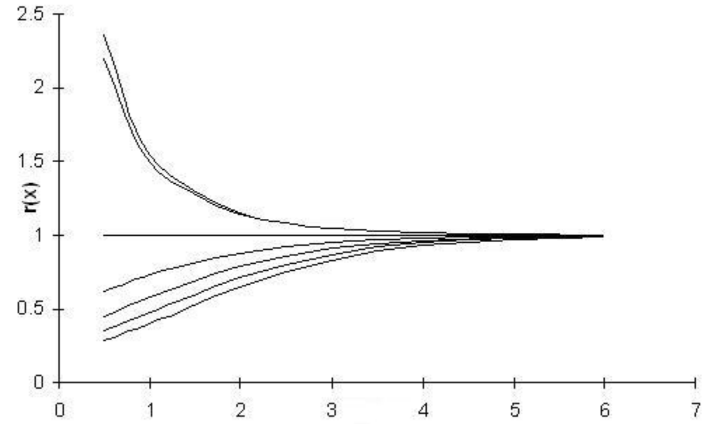

Fig. 1. Hazard rate function of MOEE (α, λ) for various values of α and λ

3. Record Value Theory

Chandler (1952) introduced the concept of records and laid the foundations of the mathematical theory related to records. Record values and associated statistics are of greater importance in many real life situations involving data relating to sports, weather, economics, life testing etc. Galambos (1978), Galambos and Kotz (1987), Arnold and Balakrishnan (1989), Balakrishnan and Ahsanullah (1994), Ahsanullah (1995), Sultan *et al.* (2003), etc. have made significant contributions to the theory of records. Arnold *et al.* (1998) provide an excellent discussion on various results in the theory of record values.

Let $X_1, X_2, \ldots$ be an infinite sequence of i.i.d. random variables having the same distribution as the (population) random variable X. An observation X_j will be called an upper record value (or simply a record) if its value exceeds that of all previous observations. Then X_j is a record if $X_j > X_i$ for every $i < j$. The time at which records appear are of interest. Let X_j be observed at time j. Then the record time sequence $\{T_n,\ n \geqslant 0\}$ is defined as $T_0 = 1$ with probability 1 and for $n \geqslant 1$, $T_n = \min\{j : X_j > X_{T_{n-1}}\}$.

The record value sequence $\{R_n\}$ is then defined by $R_n = X_{T_n}$, $n = 1, 2, \ldots, n$. Then R_n is called the n^{th} record.

3.1. *Moments of Record values*

Let $g_{R_n}(x)$ denote the p.d.f. of the n^{th} record then

$$g_{R_n}(x) = \frac{g(x)[-\log(1 - G(x))]^{n-1}}{(n-1)!}, \quad -\infty < x < \infty \tag{3.1}$$

The joint p.d.f. of a pair of records say R_m, R_n is given by

$$g_{R_m,R_n}(x,y) = \frac{[-\log \overline{G}(x)]^{m-1}}{(m-1)!} \frac{\left[-\log \frac{\overline{G}(y)}{\overline{G}(x)}\right]^{n-m-1}}{(n-m-1)!} \frac{g(x)g(y)}{1-G(x)}, \quad -\infty < x < y < \infty \tag{3.2}$$

(see Arnold *et al.*, 1998). By (3.1) the density function of the n^{th} record for MOEE(α) distribution is given by

$$g_{R_n}(x) = \frac{\alpha e^x}{(n-1)![e^x-(1-\alpha)]^2}\left[-\ln\left(\frac{\alpha}{e^x-(1-\alpha)}\right)\right]^{n-1}, \quad 0<x<\infty \tag{3.3}$$

The single moment of n^{th} record statistic can be written as

$$\beta_n = \int_0^\infty \ln(\overline{\alpha}+\alpha e^u)\frac{u^{n-1}}{(n-1)!}e^{-u}du \tag{3.4}$$

Theorem 3.1. *The single moment of n^{th} upper record value for $\alpha > 0.5$ is given by*

$$\beta_n = \ln(\alpha)+n-\sum_{i=1}^{\infty}\frac{k^i}{i(i+1)^n}, \quad \textit{where} \ \ k = 1-\frac{1}{\alpha}. \tag{3.5}$$

And consequently, for $n \geqslant 2$

$$\beta_n = \beta_{n-1}+\sum_{i=0}^{\infty}\frac{k^i}{(i+1)^n} \tag{3.6}$$

Proof. From (3.4) and using the fact that $\ln[1-ke^{-u}] = -\sum_{i=1}^{\infty}\frac{k^i e^{-iu}}{i}$

$$\beta_n = \ln(\alpha)\int_0^\infty \frac{u^{n-1}e^{-u}}{(n-1)!}du+\int_0^\infty \frac{u^n e^{-u}}{(n-1)!}du-\sum_{i=1}^{\infty}\frac{k^i}{i}\int_0^\infty \frac{e^{-(i+1)u}u^{n-1}}{(n-1)!}du \tag{3.7}$$

which on evaluation directly gives (3.5). □

Now

$$\beta_n = \ln(\alpha)+n-\sum_{i=1}^{\infty}\frac{k^i}{(i+1)^{n-1}}\left[\frac{1}{i}-\frac{1}{i+1}\right]$$

simplifying we get the recurrence relation (3.6).
Using the result (3.5) the mean of record values from MOEE(α) for $\alpha = 1.0(0.5)4.0$ are evaluated and presented in Table 1.

Table 1. Mean of upper record values

n	$\alpha = 1$	$\alpha = 1.5$	$\alpha = 2$	$\alpha = 2.5$	$\alpha = 3$	$\alpha = 3.5$	$\alpha = 4$
1	1	1.2164	1.3863	1.5272	1.6479	1.7539	1.8484
2	2	2.3150	2.5508	2.7398	2.8978	3.0337	3.1530
3	3	3.3615	3.6252	3.8331	4.0049	4.1513	4.2789
4	4	4.3839	4.6602	4.8762	5.0537	5.2043	5.3352
5	5	5.3948	5.677	5.8967	6.0767	6.2292	6.3615
6	6	6.4002	6.6852	6.9066	7.0879	7.2412	7.3741
7	7	7.4028	7.6892	7.9115	8.0933	8.2471	8.3803

Theorem 3.2. *The second single moment of n^{th} upper record value is*

$$\beta_n^2 = \ln(\alpha)^2 + n(n+1+2\ln(\alpha)) - 2n\sum_{i=1}^{\infty}\frac{k^i}{i(i+1)^{n+1}} - 2\ln(\alpha) \times \sum_{i=1}^{\infty}\frac{k^i}{i(i+1)^n} + \sum_{i=1}^{\infty}\sum_{j=1}^{\infty}\frac{k^{i+j}}{ij(i+j+1)^n} \tag{3.8}$$

Proof. From (3.4) the 2nd single moment of n^{th} record value is given by

$$\begin{aligned}\beta_n^2 &= \int_0^{\infty}\left\{\ln[\alpha e^u(1-ke^{-u})]\right\}^2\frac{u^{n-1}e^{-u}}{(n-1)!}du, \ \ k = 1-\frac{1}{\alpha} \\ &= (\ln\alpha)^2 + n(n+1) + 2n\ \ln\alpha - 2n\sum_{i=1}^{\infty}\frac{k^i}{i(i+1)^{(n+1)}} - 2 \\ &\quad \times \ln\alpha\sum_{i=1}^{\infty}\frac{k^i}{i(i+1)^n} + \sum_{i=1}^{\infty}\sum_{j=1}^{\infty}\frac{k^{i+j}}{ij}\int_0^{\infty}e^{-(i+j+1)u}\frac{u^{n-1}}{(n-1)!}du\end{aligned} \tag{3.9}$$

On simplification using the fact that $(a_1+a_2)^2 = \sum_{i=1}^{2}\sum_{j=1}^{2}a_ia_j$ we get (3.8). By (3.2) the joint p.d.f. of m^{th} and n^{th} record values of MOEE(α) distribution is given by

$$\begin{aligned}g_{R_m,R_n}(x) &= \frac{\alpha^2\left[-\ln\left\{\frac{\alpha}{e^x-(1-\alpha)}\right\}\right]^{m-1}}{(m-1)!}\frac{1}{[e^x-(1-\alpha)]} \\ &\quad \times\frac{\left[-\ln\left\{\frac{e^x-(1-\alpha)}{e^y-(1-\alpha)}\right\}\right]^{n-m-1}}{(n-m-1)!}\times\frac{e^y}{[e^y-(1-\alpha)]^2}, \quad 0<x<y<\infty\end{aligned} \tag{3.10}$$

□

Theorem 3.3. *For $1 \leqslant m \leqslant n$ the product moment*

$$\begin{aligned}\beta_{m,n} &= (\ln\alpha)^2 + \ln\alpha(m+n) + m(n+1) - [\ln\alpha + (n-m)] \\ &\quad \times\sum_{i=1}^{\infty}\frac{k^i}{i(i+1)^m} - m\sum_{i=1}^{\infty}\frac{k^i}{i(i+1)^{m+1}} - \ln\alpha\sum_{i=1}^{\infty}\frac{k^j}{j(j+1)^n} - m \\ &\quad \times\sum_{j=1}^{\infty}\frac{k^j}{j(j+1)^{n+1}} + \sum_{i=1}^{\infty}\sum_{j=1}^{\infty}\frac{k^{(i+j)}}{ij(j+1)^{n-m}(i+j+1)^m}\end{aligned} \tag{3.11}$$

Proof.

$$\beta_{m,n} = \frac{\alpha}{(m-1)!}\int_0^{\infty}x\left[-\ln\left(\frac{\alpha}{e^x-\overline{\alpha}}\right)\right]^{m-1}\frac{e^x}{e^x-\overline{\alpha}}\,I_x\,dx \tag{3.12}$$

where

$$I_x = \frac{1}{(n-m-1)!}\int_x^{\infty}\frac{ye^y}{(e^y-\overline{\alpha})^2}\left[-\ln\left(\frac{e^x-\overline{\alpha}}{e^y-\overline{\alpha}}\right)\right]^{(n-m-1)}dy$$

now making use of the transformation $u = -\ln\left(\frac{e^x-\overline{\alpha}}{e^y-\overline{\alpha}}\right)$ and writing $\ln\left[1-\left(\frac{\alpha-1}{e^x-\overline{\alpha}}\right)e^{-u}\right] = -\sum_{i=1}^{\infty}\left(\frac{\alpha-1}{e^x-\overline{\alpha}}\right)^i \frac{e^{-iu}}{i}$ we get

$$I_x = \frac{1}{(e^x-\overline{\alpha})}\left[\ln(e^x-\overline{\alpha}) + (n-m) - \sum_{i=1}^{\infty}\left(\frac{\alpha-1}{e^x-\overline{\alpha}}\right)^i \frac{1}{i(i+1)^{n-m}}\right]$$

substituting the expression of I_x in (3.12) and using the transformation $t = -\ln\left(\frac{\alpha}{e^x-\overline{\alpha}}\right)$ yields (3.11). $\square$

3.2. *Limiting distribution of n^{th} record*

In this context we derive a limit theorem which follows from Resnick (1987). According to (Resnick (1987)). If ψ_F is such that

$$\lim_{s\to\infty} \frac{\psi_F(s+x\sqrt{s}) - \psi_F(s)}{\psi_F(s+\sqrt{s}) - \psi_F(s)} = x, \quad \text{for all } x$$

then

$$\frac{R_n - \psi_F(n)}{\psi_F(n+\sqrt{n}) - \psi_F(n)} \xrightarrow{d} N(0,1) \tag{3.13}$$

where $\psi_F(u) = F^{-1}(1-e^{-u})$.

Theorem 3.4. *MOEE records are asymptotically distributed as normal.*

Proof. Substituting $\overline{G}(x) = \dfrac{\alpha e^{-x}}{1-(1-\alpha)e^{-x}}, \quad x \geqslant 0, \ \alpha > 0.$
We have $\psi_G(u) = \log[\alpha(e^u-1)+1]$ and

$$\lim_{s\to\infty} \frac{\psi_G(s+x\sqrt{s}) - \psi_G(s)}{\psi_G(s+\sqrt{s} - \psi_G(s)} = \lim_{s\to\infty} \frac{\log\left(\frac{\alpha(e^{s+x\sqrt{s}}-1)+1}{\alpha(e^s-1)+1}\right)}{\log\left(\frac{\alpha(e^{s+\sqrt{s}}-1)+1}{\alpha(e^s-1)+1}\right)} = x.$$

$\square$

Then the result follows from Resnick (1987).

3.3. *Entropy of Record Value Distribution*

Entropy provides an excellent tool to quantify the amount of information (or uncertainty) contained in a random observation regarding its parent distribution. Shannon's (1948) entropy of an absolutely continuous random variable X with probability density function $f(x)$ is given by

$$H_x[f(x)] = -\int_{-\infty}^{\infty} f(x)\ln[f(x)]dx \tag{3.14}$$

The entropy is always non-negative in the case of a discrete random variable X and is also invariant under a one-to-one transformation of X. For a continuous random variable, entropy is not invariant under a one-to-one transformation of X and it takes values in $(-\infty, +\infty)$. Now we discuss the entropy for the record values of MOEE(α,λ). Let $H_{(R_n)}$ be the entropy of the n^{th} record value. Then by Shakil (2005)

$$H_{(R_n)} = \ln(\Gamma n) - (n-1)\psi(n) - \frac{1}{\Gamma(n)}\int_{-\infty}^{\infty}[-\ln(1-G(x))]^{n-1}g(x)\ln(g(x)) \tag{3.15}$$

where $\int_0^\infty t^{j-1}\mathrm{e}^{-t}dt = \Gamma(j)$ and $\int_0^\infty t^{j-1}\mathrm{e}^{-t}\ln(t)dt = \Gamma(j)\psi(j)$, $\psi(j)$ is the digamma function. For $n=1$ entropy of the first record value is same as the entropy of parent distribution. Comparison of the entropy of parent distribution and n^{th} record value $n \geqslant 2$ is same as comparison of entropy of first record value with entropy of a given n^{th} record value, $n \geqslant 2$. Since the first observation from the parent distribution is always considered as a record value, entropy of the first non-trivial record value is obtained when $n \geqslant 2$.

Theorem 3.5. *For MOEE(α,λ) distribution if $H_{(j)}$ represents the entropy corresponding to j^{th} record, then*

$$H_{(j)} = \ln(j) - (j-1)\psi(j) + j - \ln(\lambda) + \sum_{i=1}^{\infty}\frac{k^i}{i(i+1)^j} \tag{3.16}$$

Proof. By (3.15) the entropy of j^{th} record for MOEE (α,λ) is

$$H_{(j)} = \ln(\Gamma j) - (j-1)\psi(j) - \frac{1}{\Gamma(j)}\int_0^\infty \left[-\ln\left(\frac{\alpha}{\mathrm{e}^{\lambda x}-\overline{\alpha}}\right)\right]^{j-1} v(x)\ln v(x)\, dx$$

where $v(x) = \frac{\alpha\lambda\mathrm{e}^{\lambda x}}{(\mathrm{e}^{\lambda x}-\overline{\alpha})^2}$ By the transformation $t = -\ln\frac{\alpha}{\mathrm{e}^{\lambda x}-\overline{\alpha}}$ and writing
$\ln(1-k\mathrm{e}^{-t}) = -\sum_{i=1}^{\infty}\frac{k^i\mathrm{e}^{-it}}{i}$ where $k = \frac{\alpha-1}{\alpha}$ the result (3.16) can be easily obtained. □

Using (3.16) the entropy of MOEE (α,λ) for $\alpha = 0.8$ and for various record values and various values of λ are tabulated and presented in Table 2.

Table 2. Entropy of MOEE (α,λ)

Record	$\lambda=0.5$	$\lambda=1$	$\lambda=2$	$\lambda=5$
2	2.2113	1.5182	0.8250	−0.0913
4	2.0090	2.7021	1.3159	0.3996
6	2.2545	2.9476	1.5613	0.6450
8	2.4167	3.1098	1.7235	0.8073
10	2.5384	3.2315	1.8452	0.9289

4. Stress-Strength Analysis and Estimation of Reliability

Sankaran and Jayakumar (2006) discussed the physical interpretation of Marshall–Olkin family of distributions using proportionate odds model. Bennet (1983) introduced the proportional odds(PO) model to analyse the life time data with covariates as the odds ratio. Let X be a random variable with cdf $F(x)$ and pdf $f(x)$. Then the PO model with covariates can be written as

$$\frac{\overline{G}(x;\alpha(x))}{1-\overline{G}(x,\alpha(x))} = \alpha(x)\frac{\overline{F}(x)}{1-\overline{F}(x)}$$

Then

$$\overline{G}(x,\alpha(x)) = \frac{\alpha(x)\ F(x)}{1-(1-\alpha(x))\overline{F}(x)}$$

where $\alpha(x)$ is a non negative function of the covariates, and $\overline{G}(x;\alpha(x))$ is the survival function incorporating covariates.
Treating $\alpha(x)$ as a constant α, we get

$$\overline{G}(x,\alpha) = \frac{\alpha\overline{F}(x)}{1-(1-\alpha)\overline{F}(x)}$$

The above family is known as the Marshall–Olkin family of distributions introduced by Marshall–Olkin (1997). Thus the MO family is very closely related to the PO model in survival analysis.
Gupta *et al.* (2009) showed that for two independent random variables represent strength (X) and stress (Y) follow the same Marshall–Olkin extended distributions with tilt parameters α_1 and α_2 then the Reliability of the system given by $P(X>Y)$ denoted by R is

$$R = \frac{\frac{\alpha_1}{\alpha_2}}{(\frac{\alpha_1}{\alpha_2}-1)^2}\left[-\ln\frac{\alpha_1}{\alpha_2}+\frac{\alpha_1}{\alpha_2}-1\right] \tag{4.1}$$

To estimate R it is enough if we estimate α_1,α_2 by the method of m.l.e. The log likelihood equation here is

$$LL \propto m\ln(\alpha_1)+n\log(\alpha_2)-2\sum_{i=1}^{m}\log(\mathrm{e}^{\lambda x_i}-(1-\alpha_1))-2\sum_{i=1}^{n}\log(\mathrm{e}^{\lambda y_i}-(1-\alpha_2))$$

Then the mle of α_1 and α_2 are the solutions of the non-linear equations

$$\frac{\partial LL}{\partial\alpha_1} = \frac{m}{\alpha_1}-2\sum_{i=1}^{m}\frac{1}{(\mathrm{e}^{\lambda x_i}-(1-\alpha_1))}$$
$$\frac{\partial LL}{\partial\alpha_2} = \frac{n}{\alpha_2}-2\sum_{i=1}^{m}\frac{1}{(\mathrm{e}^{\lambda y_i}-(1-\alpha_2))}$$

By the property of m.l.e. for $m\to\infty$, $n\to\infty$

$$\sqrt{m}(\widehat{\alpha_1}-\alpha_1),\sqrt{n}(\widehat{\alpha_2}-\alpha_2)\xrightarrow{d} N_2\left(\mathbf{0},\mathrm{diag}\left\{\frac{1}{a_{11}},\frac{1}{a_{22}}\right\}\right)$$

where $a_{11} = \lim_{m,n\to\infty}\frac{1}{m}I_{11} = \frac{1}{3\alpha_1^2}$ and $a_{22} = \lim_{m,n\to\infty}\frac{1}{n}I_{22} = \frac{1}{3\alpha_2^2}$

Now the information matrix has the elements

$$\begin{aligned} I_{11} &= -E\left(\frac{\partial^2 LL}{\partial \alpha_1^2}\right) \\ &= -E\left(\frac{-m}{\alpha_1^2} + 2\sum_{i=1}^{m} \frac{1}{(\mathrm{e}^{\lambda x_i} - (1-\alpha_1))^2}\right) \\ &= \frac{m}{\alpha_1^2} - 2\alpha_1 m \int_{\alpha}^{\infty} \frac{dt}{t^4} \\ &= \frac{m}{3\alpha_1^2} \end{aligned}$$

similarly $I_{22} = -E\left(\frac{\partial^2 LL}{\partial \alpha_1^2}\right) = -\frac{n}{3\alpha_2^2}$ and $I_{12} = I_{21} = -E\left(\frac{\partial^2 LL}{\partial \alpha_1 \alpha_2}\right) = 0$.

Now from Gupta *et al.* (2009) the 95% confidence interval for R is given by

$$\widehat{R} \mp 1.96\, \widehat{\alpha}_1 b_1(\widehat{\alpha}_1, \widehat{\alpha}_2)\sqrt{\frac{3}{m} + \frac{3}{n}},$$

where

$$b_1(\alpha_1, \alpha_2) = \frac{\partial R}{\partial \alpha_1} = \frac{\alpha_2}{(\alpha_1 - \alpha_2)^3}\left[-2(\alpha_1 - \alpha_2) + (\alpha_1 + \alpha_2)\ln\frac{\alpha_1}{\alpha_2}\right]$$

and

$$b_2(\alpha_1, \alpha_2) = \frac{\partial R}{\partial \alpha_2} = \frac{\alpha_1}{(\alpha_1 - \alpha_2)^3}\left[2(\alpha_1 - \alpha_2) - (\alpha_1 + \alpha_2)\ln\frac{\alpha_1}{\alpha_2}\right] = -\frac{\alpha_1}{\alpha_2} b_1(\alpha_1, \alpha_2).$$

4.1. *Simulation Study*

We generate $N = 10{,}000$ sets of X-samples and Y-samples from the Marshall–Olkin extended exponential distribution with parameters $\alpha_1,\ \lambda$ and $\alpha_2,\ \lambda$ respectively. The combinations of samples of sizes $m = 20,\ 25,\ 30$ and $n = 20,\ 25,\ 30$ along with $m = 40$, $n = 40$ are considered. The validity of the estimate of R is discussed by the measures namely average bias of the estimate $(\bar{b})$,average mean square error of the estimate (AMSE), average confidence interval of the estimate and coverage probability.

The numerical values obtained for the measures listed above are presented in Tables 3–6. For $\alpha_1 < \alpha_2$ the average bias is positive and for $\alpha_1 > \alpha_2$ the average bias is negative but in both cases the average bias decreases as the sample size increases. The average MSE is almost symmetric with respect to (α_1, α_2). This symmetric property can also be observed in the case of average confidence interval and its performance is quite good. The coverage probability is very close to 0.95 and approaches to the nominal value as the sample size increases. The simulation study indicates that the average bias, average MSE, average confidence interval and coverage probability do not show much variability for various parameter combinations.

Copyright: the authors

Table 3. Average bias and average MSE of the simulated estimates of R for $\lambda = 0.5$

(α_1, α_2)								
	Average bias $(\bar{b})$				**Average Mean Square Error AMSE**			
(m,n)	(0.5,0.8)	(0.8,1.5)	(0.8,0.5)	(1.5,0.8)	(0.5,0.8)	(0.8,1.5)	(0.8,0.5)	(1.5,0.8)
(20,20)	0.0433	0.0586	−0.0432	−0.0582	0.0063	0.0065	0.0063	0.0064
(20,25)	0.0432	0.0578	−0.0455	−0.0582	0.0061	0.0061	0.0062	0.0061
(20,30)	0.0424	0.0574	−0.0481	−0.0599	0.0057	0.0059	0.0062	0.0060
(25,20)	0.0455	0.0593	−0.0423	−0.0579	0.0062	0.0063	0.0060	0.0061
(25,25)	0.0451	0.0584	−0.0468	−0.0585	0.0058	0.0058	0.0059	0.0060
(25,30)	0.0438	0.0576	−0.0478	−0.0593	0.0055	0.0057	0.0056	0.0058
(30,20)	0.0475	0.0596	−0.0430	−0.0573	0.0061	0.0060	0.0057	0.0059
(30,25)	0.0473	0.0587	−0.0450	−0.0585	0.0056	0.0058	0.0056	0.0058
(30,30)	0.0463	0.0580	−0.0465	−0.0596	0.0054	0.0056	0.0053	0.0056
(40,40)	0.0458	0.0575	−0.0468	−0.0597	0.0048	0.0052	0.0048	0.0052

Table 4. Average confidence length and coverage probability of the simulated 95% confidence intervals of R for $\lambda = 0.5$

(α_1, α_2)								
	Average confidence length				**coverage probability**			
(m,n)	(0.5,0.8)	(0.8,1.5)	(0.8,0.5)	(1.5,0.8)	(0.5,0.8)	(0.8,1.5)	(0.8,0.5)	(1.5,0.8)
(20,20)	0.3506	0.3516	0.3508	0.3516	0.9748	0.9763	0.9793	0.9740
(20,30)	0.3327	0.3336	0.3432	0.3432	0.9780	0.9719	0.9696	0.9685
(20,25)	0.3207	0.3209	0.3216	0.3223	0.9805	0.9716	0.9672	0.9624
(25,20)	0.3335	0.3344	0.3329	0.3337	0.9706	0.9693	0.9788	0.9726
(25,25)	0.3146	0.3152	0.3145	0.3154	0.9739	0.9620	0.9742	0.9618
(25,30)	0.3010	0.3019	0.3017	0.3023	0.9787	0.9572	0.9699	0.9545
(30,20)	0.3199	0.3222	0.3204	0.3211	0.9700	0.9600	0.9782	0.9718
(30,25)	0.3017	0.3023	0.3015	0.3019	0.9655	0.9537	0.9738	0.9582
(30,30)	0.2880	0.2883	0.2879	0.2883	0.9738	0.9469	0.9763	0.9468
(40,40)	0.2500	0.2502	0.2499	0.2502	0.9664	0.9108	0.9676	0.9164

Table 5. Average bias and average MSE of the simulated estimates of R for $\lambda = 3$

(α_1, α_2)								
	Average bias ($\bar{b}$)				**Average Mean Square Error AMSE**			
(m,n)	(0.5,0.8)	(0.8,1.5)	(0.8,0.5)	(1.5,0.8)	(0.5,0.8)	(0.8,1.5)	(0.8,0.5)	(1.5,0.8)
(20,20)	0.0433	0.0586	−0.0432	−0.0582	0.0063	0.0065	0.0063	0.0064
(20,25)	0.0432	0.0578	−0.0455	−0.0582	0.0061	0.0061	0.0062	0.0061
(20,30)	0.0424	0.0574	−0.0481	−0.0599	0.0057	0.0059	0.0062	0.0060
(25,20)	0.0455	0.0593	−0.0423	−0.0579	0.0062	0.0063	0.0060	0.0061
(25,25)	0.0451	0.0584	−0.0468	−0.0585	0.0058	0.0058	0.0059	0.0060
(25,30)	0.0438	0.0576	−0.0478	−0.0593	0.0055	0.0057	0.0056	0.0058
(30,20)	0.0475	0.0596	−0.0430	−0.0573	0.0061	0.0060	0.0057	0.0059
(30,25)	0.0473	0.0587	−0.0450	−0.0585	0.0056	0.0058	0.0056	0.0058
(30,30)	0.0463	0.0580	−0.0465	−0.0596	0.0054	0.0056	0.0053	0.0056
(40,40)	0.0458	0.0575	−0.0468	−0.0597	0.0048	0.0052	0.0048	0.0052

Table 6. Average confidence length and coverage probability of the simulated 95% confidence intervals of R for $\lambda = 3$

(α_1, α_2)								
	Average confidence length				**coverage probability**			
(m,n)	(0.5,0.8)	(0.8,1.5)	(0.8,0.5)	(1.5,0.8)	(0.5,0.8)	(0.8,1.5)	(0.8,0.5)	(1.5,0.8)
(20,20)	0.3506	0.3512	0.3505	0.3512	0.9699	0.9799	0.9705	0.9819
(20,25)	0.3328	0.3334	0.3331	0.3339	0.9660	0.9785	0.9612	0.9794
(20,30)	0.3206	0.3210	0.3210	0.3214	0.9641	0.9764	0.9606	0.9739
(25,20)	0.3333	0.3339	0.3329	0.3334	0.9643	0.9762	0.9674	0.9799
(25,25)	0.3145	0.3150	0.3146	0.3150	0.9625	0.9728	0.9583	0.9733
(25,30)	0.3012	0.3016	0.3015	0.3017	0.9607	0.9713	0.9581	0.9677
(30,20)	0.3212	0.3215	0.3205	0.3211	0.9629	0.9753	0.9676	0.9762
(30,25)	0.2991	0.3018	0.3013	0.3017	0.9604	0.9692	0.9571	0.9704
(30,30)	0.2877	0.2880	0.2877	0.2880	0.9503	0.9660	0.9535	0.9655
(40,40)	0.2499	0.2498	0.2498	0.2498	0.9356	0.9485	0.9383	0.9479

5. Applications in Autoregressive Time Series Modeling

One of the simplest and widely used time series models is the autoregressive models and it is well known that autoregressive process of appropriate orders is extensively used for modeling time series data. The p^{th} order autoregressive model is defined by

$$X_n = a_1 X_{n-1} + a_2 X_{n-2} + \cdots + a_p X_{n-p} + \varepsilon_n$$

where $\{\varepsilon_n\}$ is a sequence of independent and identically distributed random variables and $a_1, a_2, \ldots, a_n$ are autoregressive parameters. In particular the first order autoregressive model is

$$X_n = a_1 X_{n-1} + \varepsilon_n, \;\; n = 1, 2, \ldots, |a_1| < 1$$

The need for non-Gaussian autoregressive models have been long felt from the fact that many naturally arising time series are clearly non-Gaussian with Markovian dependence structure. Many non-Gaussian autoregressive processes were introduced and studied during the past two decades (see Jayakumar *et al.* (1995), Jose and Pillai (1995), Seethalakshmi and Jose (2004). Jayakumar and Pillai (1993) introduced and studied first order autoregressive Mittag-Leffler process. Pillai and Jayakumar (1995) characterized a p^{th} order autoregressive Mittag-Leffler process using specialized class L property. Jose and Pillai (1995) developed generalized autoregressive time series models in Mittag-Leffler variables. Alice and Jose (2003, 2004 a,b) developed autoregressive minification processes and studied their properties.

Now we discuss some application of MOEE distribution in autoregressive time series modeling.

Lewis and McKenzie (1991) introduced and discussed various minification processes having structure

$$X_n = \min(aX_{n-1}, \varepsilon_n), \quad n = 1, 2, \ldots, |a_1| < 1$$

A more general structure is given by

$$X_n = \begin{cases} \varepsilon_n & \text{w.p. } p \\ a\min(X_{n-1}, \varepsilon_n) & \text{w.p. } (1-p);\ 0 \leqslant p \leqslant 1 \end{cases} \tag{5.1}$$

5.1. *An* AR(1) *Model with MOEE Marginal Distribution*

We construct a first order autoregressive minification process with structure given by (5.2). The model is developed as follows. Consider an AR (1) structure

$$X_n = \begin{cases} \varepsilon_n & \text{w.p. } p \\ \min(X_{n-1}, \varepsilon_n) & \text{w.p. } (1-p);\ 0 \leqslant p \leqslant 1 \end{cases} \tag{5.2}$$

where $\{\varepsilon_n\}$ is a sequence of i.i.d. r.v.s with exponential distribution with unit mean and is independent of $\{X_n\}$. Here w.p. means ‘with probability’. This is a special case of the model considered in (5.1).

Theorem 5.1. *Consider the* AR(1) *structure given by* (5.2). *Then* $\{X_n\}$ *is stationary Markovian with MOEE marginal distribution if* $\{\varepsilon_n\}$ *is distributed as exponential distribution with unit mean.*

Proof. From (5.2) it follows that

$$\overline{F}_{X_n}(x) = p\overline{F}_{\varepsilon_n}(x) + (1-p)\overline{F}_{X_{n-1}}(x)\overline{F}_{\varepsilon_n}(x) \tag{5.3}$$

Under stationary equilibrium

$$\overline{F}_X(x) = \frac{p\overline{F}_\varepsilon(x)}{1-(1-p)\overline{F}_\varepsilon(x)} \quad \text{and hence} \quad \overline{F}_\varepsilon(x) = \frac{\overline{F}_X(x)}{p+(1-p)\overline{F}_X(x)}.$$

If $\varepsilon_n \sim$ Exp (1), $\overline{F}_\varepsilon(x) = \mathrm{e}^{-x}$, then it easily follows that,

$$\overline{F}_X(x) = \frac{p\mathrm{e}^{-x}}{1-(1-p)\mathrm{e}^{-x}},$$

which is the survival function of MOEE(p).
Conversely, if we take,

$$\overline{F}_{X_n}(x) = \frac{pe^{-x}}{1-(1-p)e^{-x}},$$

In order to establish stationarity, we proceed as follows. Assume $X_{n-1} \stackrel{d}{=}$ MOEE (p) and $\varepsilon_n \stackrel{d}{=}$ Exp(1), then from (11),

$$\overline{F}_{X_n}(x) = \frac{pe^{-x}}{1-(1-p)e^{-x}}.$$

This establishes that $\{X_n\}$ is distributed as MOEE(p). □

Even if X_0 is arbitrary, it is easy to establish that $\{X_n\}$ is stationary and is asymptotically marginally distributed as MOEE(p).

In order to study the behavior of the process we simulate the sample paths for various values of p. From the sample path properties it follows that the MOEE AR(1) minification process can be used for modeling a rich variety of real data from various contexts such as financial modeling, reliability modeling, hydrological modeling etc. Now we consider some properties of MOEE AR(1) minification processes we start with the joint survival function of the random variables X_{n+1} and X_n. Let $\overline{S}(x,y) = P(X_{n+1} > x, X_n > y)$ be the joint survival function of the random variables X_{n+1} and X_n. Then we have

$$\begin{aligned}
\overline{S}(x,y) &= p\overline{F}_\varepsilon(x)\overline{F}_X(y) + (1-p)\overline{F}_\varepsilon(x)\overline{F}_X(\max(x,y)) \\
&= \begin{cases} \overline{F}_\varepsilon(x)\overline{F}_X(y), & y > x \\ \overline{F}_\varepsilon(x)\big(p\overline{F}_X(y) + (1-p)\overline{F}_X(x)\big), & y < x \end{cases} \\
&= \begin{cases} \dfrac{pe^{-x-y}}{1-(1-p)e^{-y}}, & y > x \\ \dfrac{pe^{-x}(pe^{-y} + (1-p)e^{-x} - 2p(1-p)e^{-x-y})}{(1-(1-p)e^{-x})(1-(1-p)e^{-y})}, & y < x \end{cases}
\end{aligned}$$

The joint survival function $\overline{S}$ is not absolutely continuous since the probability $P(X_{n+1} = X_n)$ is positive. Namely, it is easy to show that

$$P(X_{n+1} = X_n) = \frac{-p(1-p+\log p)}{(1-p)^2} \in (0, 0.5)$$

Consider now the probability of the event $\{X_{n+1} > X_n\}$. From (5.2) it follows that

$$P(X_{n+1} > X_n) = pP(\varepsilon_{n+1} > x_n) = \frac{p(1-p+p\log p)}{(1-p)^2} \in (0, 0.5)$$

Also, we can show that

$$P(X_{n+2} > X_n) = \frac{p(2-p-p^2+3p\log p)}{(1-p)^2} \in (0, 0.5)$$

We can use these probabilities to estimate the unknown parameter p. Define the random variables $U_n = I(X_{n+1} > X_n)$ and $V_n = I(X_{n+2} > X_n)$. It is easy to show that $E(U_n) = P(X_{n+1} > X_n)$ and

$E(V_n) = P(X_{n+2} > X_n)$. Now we consider the equations

$$\frac{1}{N}\sum_{i=1}^{N} U_i = \frac{p(1-p+p\log p)}{(1-p)^2}$$

$$\frac{1}{N-1}\sum_{i=1}^{N-1} V_i = \frac{p(2-p-p^2+3p\log p)}{(1-p)^2}$$

Solving these equations, we will obtain that the estimator of the unknown parameter p is given by

$$\widehat{p} = \frac{3}{N}\sum_{i-1}^{N} U_i - \frac{1}{N-1}\sum_{i=1}^{N-1} V_i$$

Since the MOEE AR(1) minification process$\{X_n\}$ is ergodic, it follows that $\widehat{p}$ is consistent estimator for p.

In Table 7 we give some numerical results of the estimation. We estimate 10 000 realizations of the MOEE AR (1) minification process for the true values $p = 0.2$, $p = 0.4$, $p = 0.6$ and $p = 0.8$. The simulations are repeated 100 times. We computed the sample means and the standard errors of the estimate of $\widehat{p}$.

Let us consider the autocovariance function at lag 1. After some calculations we obtain that

$$E(X_{n+1}X_n) = p\int_0^{\infty} \frac{x\mathrm{e}^{-x}dx}{1-(1-p)\mathrm{e}^{-x}} = \frac{p}{1-p}.Li_2(1-p),$$

where

$$Li_2(z) = z\int_0^{\infty} \frac{x\mathrm{e}^{-x}dx}{1-z\mathrm{e}^{-x}}$$

is dilogarithm. Now, autocovariance function at lag 1 is

$$\operatorname{cov}(X_{n+1}, X_n) = \frac{p}{1-p}.Li_2(1-p) - \frac{p^2\log p}{(1-p)^2}$$

the autocorrelation function at lag1 is

$$\operatorname{Corr}(X_{n+1}, Xn) = \frac{p(1-p)Li_2(1-p) - p^2\log p}{2p(1-p)Li_2(1-p) - p^2\log p}$$

5.2. *Extension to K^{th} Order Processes*

In this section we develop a k^{th} order autoregressive model. Consider an autoregressive model of order k with structure as

$$X_n = \begin{cases} \varepsilon_n & \text{w.p. } p_0 \\ \min(X_{n-1}, \varepsilon_n) & \text{w.p. } p_1 \\ \vdots & \\ \min(X_{n-k}, \varepsilon_n) & \text{w.p. } p_k. \end{cases}$$

such that $0 < p_i < 1$, $p_1 + p_2 + \cdots + p_k = 1 - p_0$; where $\{\varepsilon_n\}$ is a sequence of i.i.d. r.v.s following MOEE distribution independent of $\{X_{n-1},\ X_{n-2},\ \ldots\}$.

$$\overline{F}_{X_n}(x) = p_0\overline{F}_{\varepsilon_n}(x) + p_1\overline{F}_{X_{n-1}}(x)\overline{F}_{\varepsilon_n}(x) + \cdots + p_k\overline{F}_{X_{n-k}}(x)\overline{F}_{\varepsilon_n}(x)$$

Under stationary equilibrium,

$$\overline{F}_X(x) = p_0\overline{F}_\varepsilon(x) + p_1\overline{F}_X(x)\overline{F}_\varepsilon(x) + \cdots + p_k\overline{F}_X(x)\overline{F}_\varepsilon(x)$$

This reduces to

$$\overline{F}_X(x) = \frac{p_0\overline{F}_\varepsilon(x)}{1-(1-p_0)\overline{F}_\varepsilon(x)}$$

This shows that Theorem 5.1 can be suitably extended to this case also.

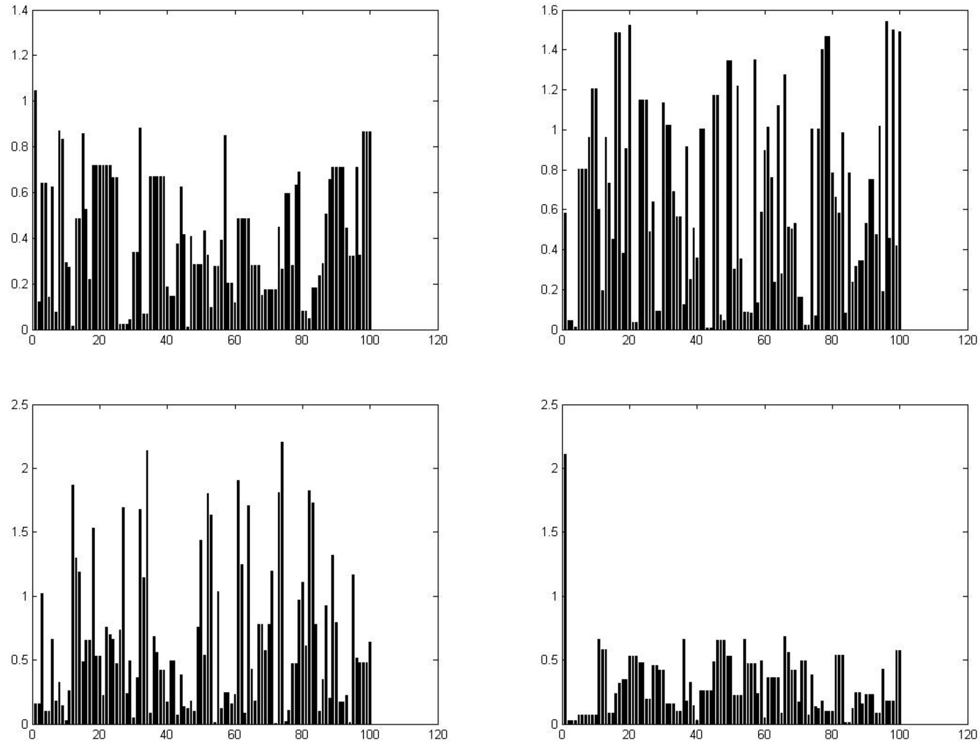

Fig. 2. Sample paths of MOEE AR (1) process $p = 0.6$, 0.8, 0.9 and 0.5

Table 7. Some numerical results of the estimation

n	$\widehat{p}$(True $p = 0.2$)	SE($\widehat{p}$)	$\widehat{p}$(True $p = 0.4$)	SE($\widehat{p}$)
100	0.205328	0.050357	0.391892	0.067521
500	0.202402	0.023738	0.397680	0.029513
1000	0.200810	0.017016	0.396480	0.022201
5000	0.200076	0.006890	0.398344	0.011454
10000	0.200169	0.004981	0.399264	0.007092
n	$\widehat{p}$(True $p = 0.6$)	SE($\widehat{p}$)	$\widehat{p}$(True $p = 0.8$)	SE($\widehat{p}$)
100	0.596091	0.084100	0.796522	0.097800
500	0.595670	0.034729	0.801667	0.038394
1000	0.595555	0.025793	0.802864	0.030903
5000	0.599764	0.010598	0.801462	0.015817
10000	0.598863	0.007853	0.800391	0.011084

Acknowledgement

The first author wishes to acknowledge the financial support in the form of a Major Research Grant from UGC, New Delhi for carrying out this research.

References

[1] Ahsanullah, M. (1995). Record values, in the exponential distributions: Theory, Methods and applications (Eds. N. Balakrishnan and A.P. Basu), Gordon and Breach publishers New York, New Jersey.

[2] Alice, T. and Jose, K.K. (2003). Marshall–Olkin Pareto Processes, *Far East Journal of Theoretical Statistics*, **9**(2), 117–132.

[3] Alice, T. and Jose, K.K. (2004 a). Bivariate semi-Pareto minification Processes, **Metrika**, **59**, 305–313.

[4] Alice, T. and Jose, K.K. (2004 b). Marshall-Olkin bivariate Pareto distribution and reliability applications, *IAPQR Trans.*, **29**, 1, 1–9.

[5] Arnold B.C. and Balakrishnan, N. (1989). Relations, Bounds and approximations for order statistics. Lecture notes in Statistics 53. Springer-Verlag, New York.

[6] Arnold B.C. and Balakrishnan, N. and Nagaraja, H.N. (1998). Records, John Wiley, New York.

[7] Balakrishnan, N. and Ahsanullah, M. (1994). Relations for single and product moments of record values from Lomax distributions. *Sankhya*: Vol. **56**, series B, 140–146.

[8] Bennet, S. (1983). Analysis of survival data by the proportional odds model. *Statist.Med.*, **2**, 273–277.

[9] Chandler K.N. (1952). The distribution and frequency of record values. *J. Roy. Statist. Soc.*, B, **14**, 220–228.

[10] Galambos, J. (1987). The Asymptotic Theory of Extreme Order Statistics, Second edition, Krieger, Malabar, Florida.

[11] Galambos, J. and Kotz, S. (1978). Characterizations of Probability Distributions, Lecture Notes in Mathematics 675, Springer-Verlag. Berlin.

[12] Gupta, R.C., Ghitany, M.E and Mutairi, D.K. (2009). Estimation of reliability from Marshall-Olkin extended Lomax distributions. *Journal of Statistical Computation and Simulation*, 1–11.

[13] Jayakumar, K., Kalyanaraman, K. and Pillai, R.N. (1995). α-Laplace processes, *Mathl. Comput. Modeling.*, **22**, 109–116.

[14] Jayakumar,K. and Pillai, R.N. (1993). The first order autoregressive Mittag-Leffler process, *J. Appl. Prob.*, **30**, 462–466.

[15] Johnson, N.L., Kotz, S., Balakrishnan, N. (2004). *Continuous Univariate Distributions*. Vol. 1, John Wiley and Sons, New York.

[16] Jose, K.K. and Pillai, R.N. (1995). Geometric Infinite divisibility and its applications in autoregressive time series modeling. Stochastic Process and its Applications, Wiley Eastern Ltd. New Delhi.

[17] Lewis, P.A.W. and McKenzie, E. (1991). Minification processes and their Transformations, *J. Appl. Prob.*, **28**, 45–57.

[18] Marshall, A.W. and Olkin, I. (1997). A new method for adding a parameter to a family of distributions with applications to the exponential and Weibull families. *Biometrika*, **84**, 641–652.

[19] Pillai, R.N. and Jayakumar, K. (1995). Discrete Mittag-Leffler Distributions. *Statist. and Prob. Letters*, **23**, 271–274.

[20] Resnick, S.I. (1987). Extreme Values, Regular Variation, and Point Processes, New York, Springer Verlag.

[21] Sankaran, P.G., Jayakumar, K., (2006). On Proportional odds models, Statistical Papers.

[22] Seetha Lekshmi, V. and Jose, K.K. (2004). An autoregressive process with geometric α-Laplace Marginals. *Statistical Papers*, **45**, 337–350.

[23] Shakil, M,(2005). Entropies of record values obtained from the normal distribution and some of their properties. *Journal of Statistical Theory and Applications*, **4**, No. 4.

[24] Shanon, C. (1948). A mathematical theory of communications. *Bell. Syst. Tech. J.*, **27**, 379–432.

[25] Sultan, K.S. Moshref, M.E. and Childs, A. (2003). Records Values from Generalized Power Function Distribution and Associated Inference. Applied Statistics at the Leading Edge, Ed. M. Ahsanullah, 107–121.

Copyright: the authors