



General results for the Marshall and Olkin's family of distributions

WAGNER BARRETO-SOUZA¹, ARTUR J. LEMONTE¹ and GAUSS M. CORDEIRO²

¹Departamento de Estatística, Universidade de São Paulo, Rua do Matão, 1010, 05508-090 São Paulo, SP, Brasil

²Departamento de Estatística, Universidade Federal de Pernambuco, Cidade Universitária, 50740-540 Recife, PE, Brasil

Manuscript received on July 2, 2011; accepted for publication on April 24, 2012

ABSTRACT

Abstract Marshall and Olkin (1997) introduced an interesting method of adding a parameter to a well-established distribution. However, they did not investigate general mathematical properties of their family of distributions. We provide for this family of distributions general expansions for the density function, explicit expressions for the moments and moments of the order statistics. Several especial models are investigated. We discuss estimation of the model parameters. An application to a real data set is presented for illustrative purposes.

Key words: Marshall-Olkin extended distribution, maximum likelihood estimation, Monte Carlo simulation, Rényi entropy.

INTRODUCTION

Adding parameters to a well-established distribution is a time honored device for obtaining more flexible new families of distributions. (Marshall and Olkin 2007) introduced an interesting method of adding a new parameter to an existing distribution. The resulting distribution, known as Marshall-Olkin (M-O) extended distribution, includes the baseline distribution as a special case and gives more flexibility to model various types of data. The M-O family of distributions is also known as the proportional odds family (proportional odds model) or family with tilt parameter (Marshall and Olkin 2007).

Let $\bar{F}(x)=1-F(x)$ denote the survival function of a continuous random variable X which depends on a parameter vector $\beta = (\beta_1, \dots, \beta_q)^T$ of dimension q . Then, the corresponding M-O extended distribution has survival function defined by

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

where $\alpha > 0$ and $\bar{\alpha} = 1 - \alpha$. For $\alpha = 1$, $\bar{G}(x) = \bar{F}(x)$. Marshall and Olkin (1997) have noted that the method has a stability property, i.e., if the method is applied twice, nothing new is obtained in the second time around. Additionally, the extended model is geometrically extremely stable. If $X_i = (i = 1, 2, \dots)$ is a sequence of

Correspondence to: Artur J. Lemonte
E-mail: arturlemonte@gmail.com

independent and identically distributed random variables with cdf $F(x)$ and if N has a geometric distribution taken values $\{1, 2, \dots\}$, then the random variables $U = \min \{X_1, \dots, X_N\}$ and $V = \max \{X_1, \dots, X_N\}$ are distributed as in (1). It implies that the new distribution is geometrically extremely stable. Marshall and Olkin (2007) have called the additional shape parameter “tilt parameter”, since the hazard rate of the new family is shifted below ($\alpha \geq 1$) or above ($0 < \alpha \leq 1$) the hazard rate of the underlying distribution, that is, for all $x \geq 0$, $h(x) \leq r(x)$ when $\alpha \geq 1$, and $h(x) \geq r(x)$ when $0 < \alpha \leq 1$, where $h(x)$ denotes the hazard rate of the transformed distribution and $r(x)$ is that of the original distribution.

Some special cases discussed in the literature include the M-O extensions of the Weibull distribution (Ghitany et al. 2005, Zhang and Xie 2007), Pareto distribution (Ghitany 2005), gamma distribution (Ristić et al. 2007), Lomax distribution (Ghitany et al. 2007) and linear failure-rate distribution (Ghitany and Kotz 2007). More recently, Gómez-Déniz (2010) presented a new generalization of the geometric distribution using the M-O scheme. Economou and Caroni (2007) showed that the M-O extended distributions have a proportional odds property and Caroni (2010) presented some Monte Carlo simulations considering hypothesis testing on the parameter α for the extended Weibull distribution. Maximum likelihood estimation in M-O family is given in Lam and Leung (2001) and Gupta and Peng (2009). Gupta et al. (2010) compared this family and the original distribution with respect to some stochastic orderings and also investigate thoroughly the monotonicity of the failure rate of the resulting distribution when the baseline distribution is taken as Weibull. Nanda and Das (2012) investigated the tilt parameter of the M-O extended family.

The probability density function (pdf) of the M-O extended-F distribution, say $g(x)$, is given by

$$g(x) = \frac{\alpha f(x)}{\{1 - \bar{\alpha} \bar{F}(x)\}^2}, \quad -\infty < x < \infty, \quad (2)$$

where $f(x) = dF(x) / dx$ is the baseline density function corresponding to $F(x)$. Here after, we refer to the family (2) as the M-O extended-F distribution.

General mathematical properties of the M-O extended-F distribution were not derived by Marshall and Olkin (1997) such as moments and moments of order statistics. In this article, we derive some general structural properties of the M-O extended-F distribution including: (i) expansions for the pdf; (ii) general expressions for the moments; (iii) moments of order statistics; (iv) Rényi entropy. We propose several M-O extended-F distributions taken as baseline in the definitions the Weibull, Fréchet, Pareto, generalized exponential, Kumaraswamy and power function distributions. We discuss maximum likelihood estimation of the model parameters.

The article is organized as follows. Section 2 presents expansions for the density function and for the density function of the order statistics. Explicit expressions for the moments and moments of the order statistics of the M-O extended-F distribution are given in Section 3. Rényi entropy is derived in Section 4. Estimation of the model parameters by maximum likelihood is discussed in Section 5. Section 6 presents an alternative method to estimate the model parameters. In Section 7, we propose several M-O extended-F distributions and discuss some of their properties. Simulation results are performed in Section 8. An application of the current family to a real data set is explored in Section 9. Finally, some concluding remarks are presented in Section 10.

EXPANSIONS

Consider the series representation

$$(1-z)^{-k} = \sum_{j=0}^{\infty} \frac{\Gamma(k+j)}{\Gamma(k)j!} z^j \quad (3)$$

which is valid for $|z| < 1$ and $k > 0$, where $\Gamma(\cdot)$ is the gamma function. If $\alpha \in (0,1)$ using (3) in (2), we obtain

$$g(x) = f(x) \sum_{j=0}^{\infty} \sum_{k=0}^j w_{j,k} F(x)^{j-k}, \quad (4)$$

where

$$w_{j,k} = w_{j,k}(\alpha) = \alpha(j+1)(1-\alpha)^j (-1)^{j-k} \binom{j}{k}.$$

The density function (2) can be expressed as

$$g(x) = \frac{f(x)}{\alpha \{1 - (\alpha - 1)F(x)/\alpha\}^2}.$$

Hence, for $\alpha > 1$, using (3) in the last equation yields

$$g(x) = f(x) \sum_{j=0}^{\infty} v_j F(x)^j, \quad (5)$$

where

$$v_j = v_j(\alpha) = \frac{(j+1)(1-1/\alpha)^j}{\alpha}.$$

We now give the pdf of the i th order statistic $X_{i:n}$, say $g_{i:n}(x)$, in a random sample of size n from the M-O extended-F distribution. The pdf of $X_{i:n}$ can be expressed as

$$g_{i:n}(x) = an!f(x) \sum_{j=0}^{n-i} \frac{(-1)^j}{(i-1)!(n-i)!} \frac{F(x)^{l+i-1}}{\{1-\bar{\alpha}\bar{F}(x)\}^{l+i-1}}, \quad -\infty < x < \infty.$$

If $\alpha \in (0,1)$ using expansion (??) in the last equation, we obtain

$$g_{i:n}(x) = f(x) \sum_{j=0}^{\infty} \sum_{l=0}^{n-i} \sum_{k=0}^j u_{j,l,k} F(x)^{j+l-k+i-1}, \quad (6)$$

where

$$u_{j,l,k} = u_{j,l,k}(\alpha) = \frac{an!(-1)^l(1-\alpha)^j(-1)^{j-k}}{(i-1)!(n-i)!} \binom{j}{k} \binom{l+i+j}{j}$$

For $\alpha > 1$, we write $1 - \bar{\alpha}\bar{F}(x) = \alpha \{1 - (\alpha - 1)F(x)/\alpha\}$ and using (Ref: exp), $g_{i:n}(x)$ becomes

$$g_{i:n}(x) = f(x) \sum_{j=0}^{\infty} \sum_{l=0}^{n-i} c_{j,l} F(x)^{j+l+i-1}, \quad (7)$$

where

$$c_{j,l} = c_{j,l}(\alpha) = \frac{(-1)^l (\alpha - 1)^j n!}{\alpha^{l+j+i} (i-1)! (n-i)!} \binom{l+i+j}{j}$$

Equations (4)-(7) reveal that the density functions of the M-O extended-F distribution and of its order statistics can be expressed as the baseline density $f(x)$ multiplied by an infinite power series of $F(x)$. They play an important role and will be used to obtain explicit expressions for the moments of the M-O extended-F distribution and of its order statistics in a general framework and for special models.

MOMENTS

Here after, suppose that X has the density function (2). We derive general expressions for the moments of X and its order statistics in terms of the probability weighted moments (PWMs) of the F distribution. The PWMs, first proposed by Greenwood et al. (1979), are expectations of certain functions of a random variable whose mean exists. A general theory for these moments covers the summarization and description of theoretical probability distributions and observed data samples, non parametric estimation of the underlying distribution of an observed sample, estimation of parameters, quantiles of probability distributions and hypothesis tests. The PWMs method can generally be used for estimating parameters of a distribution whose inverse form cannot be expressed explicitly. The PWMs for the baseline F distribution are formally defined by

$$\Gamma_{p,r} = \int_{-\infty}^{\infty} x^p F(x)^r f(x) dx.$$

Thus, from equations (4) and (5), the s th moment of X for $\alpha \in (0,1)$ and $\alpha > 1$ can be written as

$$E(X^s) = \sum_{j=0}^{\infty} \sum_{k=0}^j w_{j,k} \Gamma_{sj+l-k+i} \quad \text{and} \quad E(X^s) = \sum_{j=0}^{\infty} v_j \Gamma_{sj} \quad (8)$$

respectively.

Now, using equations (6) and (7), we can determine the s th moment of the i th order statistic $X_{i:n}$ in a random sample of size n from X for $\alpha \in (0,1)$ and $\alpha > 1$ as

$$E(X_{i:n}^s) = \sum_{j=0}^{\infty} \sum_{l=0}^{n-j} \sum_{k=0}^j u_{j,l,k} \Gamma_{sj+l-k+i-1} \quad \text{and} \quad E(X_{i:n}^s) = \sum_{j=0}^{\infty} \sum_{l=0}^{n-j} c_{j,l} \Gamma_{sj+l+i-1}, \quad (9)$$

respectively, where the quantities $w_{j,k}$, v_j , $u_{j,l,k}$ and $c_{j,l}$ are defined in Section 2. Thus, the moments of X and $X_{i:n}$ are obtained in terms of infinite weighted sums of PWMs of the baseline F distribution.

RÉNYI ENTROPY

The entropy of a random variable is a measure of uncertainty variation and has been used in various situations in science and engineering. The Rényi entropy is defined by

$$I_R(\delta) = (1 - \delta)^{-1} \log \left(\int_{-\infty}^{\infty} g^\delta(x) dx \right),$$

where $\delta > 0$ and $\delta \neq 1$. For further details, the reader is referred to Song (2001). For $\alpha \in (0,1)$, using expansion (3), we can write

$$g(x)^\delta = \frac{\alpha^\delta f(x)^\delta}{\Gamma(2\delta)} \sum_{j=0}^{\infty} (1-\alpha)^j \Gamma(2\delta+j) \frac{[1-F(x)]^j}{j!}.$$

For $\alpha > 1$, we obtain

$$g(x)^\delta = \frac{f(x)^\delta}{\alpha^\delta \Gamma(2\delta)} \sum_{j=0}^{\infty} (\alpha-1)^j \Gamma(2\delta+j) \frac{F(x)^j}{j!}.$$

Thus, the Rényi entropy of X can be obtained for $\alpha \in (0,1)$ and $\alpha > 1$ as

$$I_R(\delta) = (1-\delta)^{-1} \log \left(\sum_{j=0}^{\infty} e_j \int_{-\infty}^{\infty} f(x)^\delta [1-F(x)]^j dx \right), \quad (10)$$

and

$$I_R(\delta) = (1-\delta)^{-1} \log \left(\sum_{j=0}^{\infty} h_j \int_{-\infty}^{\infty} f(x)^j F(x)^j dx \right), \quad (11)$$

respectively, where

An interesting quantity based on the Rényi entropy is defined by $S_g = -2d I_R(\delta)/d\delta|_{\delta=1}$. It is a location and scale-free positive functional and measures the intrinsic shape of a distribution (see, Song 2001).

$$e_j = e_j(\alpha) = \frac{\alpha^\delta (1-\alpha)^j \Gamma(2\delta+j)}{\Gamma(2\delta)j!}, \quad h_j = h_j(\alpha) = \frac{(1-\alpha)^j \Gamma(2\delta+j)}{\alpha^{\delta+j} \Gamma(2\delta)j!}.$$

For the purpose of illustration, we derive this quantity for the M-O extended exponential (M-O-EE) distribution, whose density function is $g(x) = \alpha \lambda e^{\lambda x} (e^{\lambda x} - \alpha)^{-2}$ (for $x > 0$ and $\lambda > 0$). Consider the case $\alpha \in (0,1)$; the case $\alpha > 1$ follows in a similar way. Define

$$c_{\delta, \alpha} = \sum_{j=0}^{\infty} \frac{(1-\alpha)^j \Gamma(2\delta+j)}{j!(\delta+j)}$$

The Song's measure for the M-O-EE distribution $\alpha \in (0,1)$ takes the form

$$S_g = -4\Psi(2) + \alpha c_1'' - (\alpha c_1')^2,$$

where

$$c_1' = \frac{\log \alpha}{(1-\alpha)} + 2 \sum_{j=0}^{\infty} (1-\alpha)^j \frac{\Gamma'(j+2)}{(j+1)!}$$

and

$$c_1'' = 4 \sum_{j=0}^{\infty} \frac{(1-\alpha)^j}{(j+1)!} \left\{ \Gamma'''(j+2) - \frac{\Gamma'(j+2)}{(j+1)!} \right\} + 2 \sum_{j=0}^{\infty} \frac{(1-\alpha)^j}{(j+1)^3}$$

where $\Psi(\cdot)$ is the digamma function, $\Gamma'(z) = d\Gamma(z)/dz$ and $\Gamma''(z) = d\Gamma'(z)/dz$.

MAXIMUM LIKELIHOOD

The model parameters of the M-O extended-F distribution can be estimated by maximum likelihood. Let $\mathbf{x} = (x_1, \dots, x_n)^\top$ be a random sample of size n from X with unknown parameter vector $\boldsymbol{\theta} = (\alpha, \beta^\top)^\top$, where $\beta = (\beta_1, \dots, \beta_q)^\top$ corresponds to the parameter vector of the baseline distribution. The log-likelihood function for $\boldsymbol{\theta}$ is

$$\ell = \ell(\boldsymbol{\theta}) = n \log \alpha + \sum_{i=1}^n \log f(x_i; \beta) - 2 \sum_{i=1}^n \log \{1 - \bar{\alpha} \bar{F}(x_i; \beta)\}.$$

By taking the partial derivatives of the log-likelihood function with respect to α and β , we obtain the components of the score vector $\mathbf{U}_\theta = (U_\alpha, \mathbf{U}_\beta^\top)^\top$ given by

$$U_\alpha = \frac{\partial \ell}{\partial \alpha} = \frac{n}{\alpha} - 2 \sum_{i=1}^n \frac{\bar{F}(x_i; \beta)}{1 - \bar{\alpha} \bar{F}(x_i; \beta)}, \quad U_\beta = \frac{\partial \ell}{\partial \beta} = \sum_{i=1}^n \frac{\partial \log f(x_i; \beta)}{\partial \beta} + \sum_{i=1}^n \frac{2\bar{\alpha}}{1 - \bar{\alpha} \bar{F}(x_i; \beta)} \frac{\partial \bar{F}(x_i; \beta)}{\partial \beta}.$$

Setting these equations to zero, $\mathbf{U}_\theta = 0$, and solving them simultaneously yields the maximum likelihood estimate (MLE) $\hat{\boldsymbol{\theta}} = (\hat{\alpha}, \hat{\beta}^\top)^\top$ of $\boldsymbol{\theta} = (\alpha, \beta^\top)^\top$. These equations cannot be solved analytically and statistical software can be used to solve them numerically. For example, the BFGS method (see, Nocedal and Wright 1999, Press et al. 2007) with analytical derivatives can be used for maximizing the log-likelihood function $\ell(\boldsymbol{\theta})$.

The normal approximation for the $\boldsymbol{\theta}$ can be used for constructing approximate confidence intervals and confidence regions for the parameters α and β . Under conditions that are fulfilled for the parameters in the interior of the parameter space, we have $\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \stackrel{a}{\sim} \mathcal{N}_{q+1}(\mathbf{0}, \mathbf{K}(\boldsymbol{\theta})^{-1})$, where $\stackrel{a}{\sim}$ means approximately distributed and $\mathbf{K}(\boldsymbol{\theta})$ is the unit expected information matrix given by

$$\mathbf{K}(\boldsymbol{\theta}) = \begin{bmatrix} K_{\alpha,\alpha} & \mathbf{K}_{\alpha,\beta} \\ \mathbf{K}_{\alpha,\beta}^\top & \mathbf{K}_{\beta,\beta} \end{bmatrix},$$

whose elements are given by

$$K_{\alpha,\alpha} = \frac{1}{3\alpha^2}, \quad \mathbf{K}_{\alpha,\beta} = 2\mathbf{E} \left(\frac{1}{[1 - \bar{\alpha} \bar{F}(X; \beta)]^2} \frac{\partial \bar{F}(X; \beta)}{\partial \beta} \right),$$

$$\mathbf{K}_{\beta,\beta} = -\mathbf{E} \left(\frac{\partial^2 \log f(x_i; \beta)}{\partial \beta \partial \beta^\top} \right) - 2\bar{\alpha}^2 \mathbf{E} \left(\frac{1}{[1 - \bar{\alpha} \bar{F}(X; \beta)]^2} \frac{\partial \bar{F}(X; \beta)}{\partial \beta} \frac{\partial \bar{F}(X; \beta)}{\partial \beta^\top} \right)$$

$$- 2\bar{\alpha} \mathbf{E} \left(\frac{1}{1 - \bar{\alpha} \bar{F}(X; \beta)} \frac{\partial^2 \bar{F}(X; \beta)}{\partial \beta \partial \beta^\top} \right).$$

The asymptotic behavior remains valid if $\mathbf{K}(\boldsymbol{\theta}) = \lim_{n \rightarrow \infty} n^{-1} \mathbf{J}_n(\boldsymbol{\theta})$, where $\mathbf{J}_n(\boldsymbol{\theta})$ is the observed information matrix, it is replaced by the average sample information matrix evaluated at $\hat{\boldsymbol{\theta}}$, i.e. $n^{-1} \mathbf{J}_n(\hat{\boldsymbol{\theta}})$. The observed information matrix is given by

$$\mathbf{J}_n(\boldsymbol{\theta}) = \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} = \begin{bmatrix} U_{\alpha\alpha} & U_{\alpha\beta} \\ U_{\alpha\beta}^T & U_{\beta\beta} \end{bmatrix},$$

whose elements are

$$U_{\alpha\alpha} = -\frac{n}{\alpha^2} + 2 \sum_{i=1}^n \left(\frac{\bar{F}(x_i; \beta)}{1 - \bar{\alpha}\bar{F}(x_i; \beta)} \right) \quad U_{\alpha\beta} = -2 \sum_{i=1}^n \frac{\partial \bar{F}(x_i; \beta) / \partial \beta}{[1 - \bar{\alpha}\bar{F}(x_i; \beta)]^2},$$

$$U_{\beta\beta} = \sum_{i=1}^n \frac{\partial^2 \log f(x_i; \beta)}{\partial \beta \partial \beta^T} + 2\bar{\alpha}^2 \sum_{i=1}^n \frac{(\partial \bar{F}(x_i; \beta) / \partial \beta) (\partial \bar{F}(x_i; \beta) / \partial \beta)^T}{[1 - \bar{\alpha}\bar{F}(x_i; \beta)]^2}$$

$$+ 2\bar{\alpha} \sum_{i=1}^n \frac{\partial^2 \bar{F}(x_i; \beta) / \partial \beta \partial \beta^T}{1 - \bar{\alpha}\bar{F}(x_i; \beta)}.$$

We can easily check if the fit of the M-O extended-F model is statistically “superior” to a fit using the F model by testing the null hypothesis $\mathcal{H}_0: \alpha \neq 1$. For testing $\mathcal{H}_0: \alpha \neq 1$ the likelihood ratio (LR) statistic is given by $w = 2 \{ \ell(\hat{\alpha}, \hat{\beta}) - \ell(\hat{\alpha}, \hat{\beta}) \}$, where $\hat{\alpha}$ and $\hat{\beta}$ are the unrestricted MLEs obtained from the maximization of ℓ under \mathcal{H}_1 and $\hat{\beta}$. The limiting distribution of this statistic is χ_1^2 under the null hypothesis. The null hypothesis is rejected if the test statistic exceeds the upper $100(1-\gamma)\%$ quantile of the χ_1^2 distribution.

ESTIMATION-TYPE METHOD OF MOMENTS

We now present an alternative method to estimate the model parameters. Since the moments cannot be obtained in closed form, the estimation by the method of moments is complicated. However, after some algebra, we obtain

$$E \{ [1 - \bar{\alpha}\bar{F}(x_i; \beta)]^v \} = \begin{cases} -\alpha \log(\alpha) / (1-\alpha), & v = 1, \\ \alpha(1 - \alpha^{v-1}) / \{ \bar{\alpha}(v-1) \} & v \in \{2, 3, \dots\}. \end{cases} \quad (12)$$

Thus, we can use (12) to construct a new method of estimation, i.e., if x_1, \dots, x_2 is a random sample with survival function (1), we can estimate the model parameters from the equation

$$\frac{1}{n} \sum_{i=1}^n \{1 - \bar{\alpha}\bar{F}(x_i; \beta)\}^v = \begin{cases} -\alpha \log(\alpha) / (1-\alpha), & v = 1, \\ \alpha(1 - \alpha^{v-1}) / \{ \bar{\alpha}(v-1) \} & v = 2, \dots, q+1. \end{cases}$$

In Section, we apply the two methods (maximum likelihood and estimation-type method of moments) to estimate the model parameters of the M-O extended family.

SPECIAL M-O EXTENDED MODELS

We motivate the study of Marshall and Olkin's distributions by considering some special models to illustrate the applicability of the previous results. Here, we obtain the moments and Rényi entropy for some special M-O extended-F distributions when the baseline F distribution follows the Weibull, Fréchet, Pareto, generalized exponential, Kumaraswamy and power distributions. Some others M-O extended-F distributions could be proposed and our general results applied to them. Clearly, the quantities $\tau_{p,r}$ are determined from the baseline F cdf.

M-O EXTENDED WEIBULL DISTRIBUTION

The M-O extended Weibull (M-O-EW) distribution is defined by taking $F(x) = 1 - e^{-(\lambda x)^\gamma}$ and $f(x) = \gamma \lambda^\gamma x^{\gamma-1} e^{-(\lambda x)^\gamma}$, $x, \lambda, \gamma > 0$. Its associated pdf and cdf are respectively given by

$$g(x) = \frac{\alpha \gamma \lambda^\gamma x^{\gamma-1} e^{-(\lambda x)^\gamma}}{\{1 - \bar{\alpha} e^{-(\lambda x)^\gamma}\}^2} \quad \text{and} \quad G(x) = \frac{-1 e^{-(\lambda x)^\gamma}}{1 - \bar{\alpha} e^{-(\lambda x)^\gamma}}.$$

The M-O-EW distribution was studied by Ghitany et al. (2005); see also Barreto-Souza et al. (2011). We obtain

$$\tau_{p,r} = \frac{\Gamma(1+p/\gamma)}{\lambda^p} \sum_{k=0}^r \binom{r}{k} \frac{(-1)^k}{(k+1)^{1+p/\gamma}}$$

and

$$\int_0^\infty f^\delta(x) F(x)^j dx = (\gamma \lambda)^\delta \Gamma(1+(\delta-1)(1-\gamma^{-1})) \sum_{k=0}^j \frac{(-1)^k \binom{j}{k}}{(\delta+k)^{1+(\delta-1)(1-\gamma^{-1})}},$$

where the last equation holds for $(\delta-1)(\gamma-1) > -1$. From these quantities, we immediately obtain explicit expressions for the moments, moments of the order statistics and Rényi entropy. If $\gamma = 1$, the results correspond to the M-O extended exponential distribution.

M-O EXTENDED FRÉCHET DISTRIBUTION

Here, we consider the Fréchet distribution (for $x, \sigma, \lambda > 0$) with cdf and pdf given by $F(x) = e^{-(\sigma/x)^\lambda}$ and $f(x) = \lambda \sigma^\lambda x^{-(\lambda+1)} e^{-(\sigma/x)^\lambda}$, respectively. The pdf and survival function of the M-O extended Fréchet (M-O-EF) distribution (for $x > 0$) reduce to

$$g(x) = \frac{\alpha \lambda \sigma^\lambda x^{-(\lambda+1)} e^{-(\sigma/x)^\lambda}}{\{1 - \bar{\alpha} [1 - e^{-(\sigma/x)^\lambda}]\}^2} \quad \text{and} \quad G(x) = \frac{\alpha [1 - e^{-(\sigma/x)^\lambda}]}{1 - \bar{\alpha} [1 - e^{-(\sigma/x)^\lambda}]},$$

respectively. After some algebra, we obtain

$$\tau_{p,r} = \sigma^p (r+1)^{p/\lambda-1} \Gamma(1-p/\lambda),$$

which is valid for $p < \lambda$. Applying this result in (8) and (9), it follows simple expressions for the moments and moments of the order statistics of the M-O-EF distribution. An expression for the Rényi entropy of this distribution is obtained by inserting

$$\int_0^\infty f^\delta(x) F(x)^j dx = \frac{\lambda^{\delta-1} \Gamma(\delta(1+\lambda^{-1}))}{\sigma^\delta (\delta+j)^{\delta(\lambda-1)}}$$

in (10) and (11).

M-O EXTENDED EXPONENTIATED PARETO DISTRIBUTION

The pdf and cdf of the exponentiated Pareto distribution, introduced by Nadarajah (2005), for $x > \theta$ and θ , k , $\gamma > 0$, are given by $f(x) = \gamma k \theta^k x^{-(k+1)} \{1 - (\theta/x)^k\}^{\gamma-1}$ and $F(x) = \{1 - (\theta/x)^k\}^\gamma$, respectively. Applying the function $F(x)$ in (1), the M-O extended exponentiated Pareto (M-O-EEP) survival function reduces to

$$\bar{G}(x) = \frac{\alpha [1 - \{1 - (\theta/x)^k\}^\gamma]}{1 - \bar{\alpha} [1 - \{1 - (\theta/x)^k\}^\gamma]}, \quad x > \theta.$$

The corresponding density function becomes

$$g(x) = \frac{\alpha \gamma k \theta^k x^{-(k+1)} \{1 - (\theta/x)^k\}^{\gamma-1}}{(1 - \bar{\alpha} [1 - \{1 - (\theta/x)^k\}^\gamma])^2}, \quad x > \theta.$$

The M-O-EEP distribution has as special sub-models the Pareto, exponentiated Pareto and extended Pareto distributions for the choices $\alpha = \gamma = 1$, $\alpha = 1$ and $\gamma = 1$, respectively. The extended M-O-Pareto ($\gamma = 1$) distribution was studied by Ghitany (2005). Algebraic manipulations lead to

$$\tau_{p,r} = \sigma^p (r+1)^{p/\lambda-1} \Gamma(1-p/\lambda),$$

restricted to $p < k$, and

$$\int_0^\infty f^\delta(x) F(x)^j dx = \gamma^\delta (k/\theta)^{\delta-1} B(\gamma(j+\delta)+1-\delta, (\delta-1)/k+\delta),$$

which is valid for $\delta(\gamma-1) > -1$ and $\delta > (k+1)^{-1}$, where $B(a, b) = \int_0^1 x^{a-1} (1-x)^{b-1} dx$ is the beta function.

M-O EXTENDED GENERALIZED EXPONENTIAL DISTRIBUTION

The pdf and cdf of the generalized exponential distribution, introduced by Gupta and Kundu (1999), for $x > 0$ and λ , $\gamma > 0$, are given by $f^\delta(x) = \gamma \lambda e^{-\lambda x} (1 - e^{-\lambda x})^{\gamma-1}$ and $F(x) = (1 - e^{-\lambda x})^\gamma$, respectively. By replacing these quantities in (1) and (2), the pdf and survival function of the M-O extended generalized exponential (M-O-EGE) distribution reduce respectively to (for $x > 0$)

$$g(x) = \frac{\alpha \gamma \lambda e^{-\lambda x} (1 - e^{-\lambda x})^{\gamma-1}}{\{1 - \bar{\alpha} (1 - e^{-\lambda x})^\gamma\}^2} \quad \text{and} \quad \bar{G}(x) = \frac{\alpha [1 - (1 - e^{-\lambda x})^\gamma]}{1 - \bar{\alpha} [1 - (1 - e^{-\lambda x})^\gamma]}.$$

From (4) and (5), we obtain the moment generating function (mgf) of the M-O-EGE distribution as

$$M(t) = \alpha \sum_{j=0}^{\infty} \sum_{i=0}^j \binom{j}{i} \frac{(-1)^i (j+1) (1-\alpha)^j}{i+1} \frac{\Gamma(1+\gamma(i+1)) \Gamma(1-t/\lambda)}{\Gamma(1+\gamma(i+1)-t/\lambda)},$$

for $\alpha \in (0,1)$ and

$$M(t) = \frac{1}{\alpha} \sum_{j=0}^{\infty} (1-\alpha^{-1})^j \frac{\Gamma(1+\gamma(j+1)) \Gamma(1-t/\lambda)}{\Gamma(1+\gamma(j+1)-t/\lambda)},$$

for $\alpha > 1$. Both formulas hold for $t < \lambda \min \{1, 1 + \gamma\}$. Hence, their moments can be obtained from the derivatives of the mgf at $t = 0$. We also have

$$\int_0^{\infty} f^{\delta}(x) F(x)^j dx = \gamma^{\delta} \lambda^{\delta-1} B(1 + \gamma(j + \delta), \delta),$$

which leads to a simple expression for the Rényi entropy.

M-O EXTENDED KUMARASWAMY DISTRIBUTION

Now, we propose a new distribution for modeling rates and proportions. Consider the Kumaraswamy distribution with pdf and cdf in the forms [for $x \in (0, 1)$ and $a, b > 0$] $f(x) = abx^{a-1}$ and $F(x) = 1 - (1 - x^a)^b$, respectively. This distribution, introduced by Kumaraswamy (1980), was well investigated by Jones (2009); see also Lemonte (2011). The M-O extended Kumaraswamy distribution, for $x \in (0, 1)$, have pdf and cdf given respectively by

$$g(x) = \frac{\alpha abx^{a-1}(1-x^a)^{b-1}}{\{1 - \bar{\alpha}(1-x^a)^b\}^2} \quad \text{and} \quad G(x) = \frac{\alpha(1-x^a)^b}{1 - \bar{\alpha}(1-x^a)^b}.$$

Hence,

$$\tau_{p,r} = b \sum_{k=0}^r \binom{r}{k} \frac{(-1)^k}{k+1} B(1+p/\alpha, b(k+1))$$

and

$$\int_0^1 f^{\delta}(x) F(x)^j dx = \alpha^{\delta-1} b^{\delta} \sum_{k=0}^r \binom{j}{k} (-1)^k B(1 + (\delta-1)(1-\alpha^{-1}), 1+b(k+\delta) - \delta),$$

restricted to $(\delta-1)(1-\alpha^{-1}) > -1$.

M-O EXTENDED POWER DISTRIBUTION

Our final special case concentrates on the M-O extended power (M-O-EPo) distribution defined (for $x \in (0, 1/\theta)$ and $\theta > 0$) by taking $F(x) = (\theta x)^k$ in (1). For $x \in (0, 1/\theta)$, the pdf and cdf of the M-O-EPo distribution are given respectively by

$$g(x) = \frac{\alpha k \theta^k x^{k-1}}{\{1 - \bar{\alpha} [1 - (\theta x)^k]\}^2} \quad \text{and} \quad G(x) = \frac{\alpha [1 - (\theta x)^k]}{1 - \bar{\alpha} [1 - (\theta x)^k]}.$$

Their moments and moments of the order statistics can be obtained by setting

$$\tau_{p,r} = \frac{k\theta^{-p}}{p+k(r+1)}$$

in (8) and (9). We also have

$$\int_0^{1/\theta} f^{\delta}(x) F(x)^j dx = \frac{k\delta\theta^{\delta-1}}{k(\delta+j) - \delta + 1},$$

valid for $\delta (1-k)^{-1}$. Replacing the last expression in (10) and (11), it follows simple expressions for the Rényi entropy.

SIMULATION RESULTS

In what follows, we shall present Monte Carlo simulation results. All the Monte Carlo simulation experiments are performed using the Ox matrix programming language (Doornik 2006). Ox is freely distributed for academic purposes and available at <http://www.doornik.com>. The number of Monte Carlo replications was $R = 20,000$.

We apply the estimation methods before discussed in order to estimate the model parameters of the M-O extended-F distribution. We adopt the M-O-EE distribution with pdf and hazard rate given by

$$g(x) = \frac{\alpha \lambda e^{-\lambda x}}{(1 - \bar{\alpha} e^{-\lambda x})^2} \quad \text{and} \quad h(x) = \frac{\lambda}{1 - \bar{\alpha} e^{-\lambda x}} \quad x > 0,$$

respectively, where $\alpha > 0$ and $\lambda > 0$. According to Marshall and Olkin (1997), this distribution may sometimes be a competitor to the Weibull and gamma distributions. The authors derived several properties of the M-O-EE distribution. For example, they showed that $h(x)$ is decreasing in x for $0 < \alpha \leq 1$ and that $h(x)$ is increasing in x for $\alpha \geq 1$. Additionally, $\lambda \leq h(x) \leq \lambda / \alpha$ for $0 \leq \alpha \leq 1$, $\lambda / \alpha \leq h(x) \leq \lambda$ for $\alpha \leq 1$, $e^{-\lambda x / \alpha} \leq \bar{G}(x) \leq e^{-\lambda x}$ for $0 \leq \alpha \leq 1$ and $e^{-\lambda x} \leq \bar{G}(x) \leq e^{-\lambda x / \alpha}$ for $\alpha \geq 1$. Rao et al. (2009) developed a reliability test plan for acceptance/rejection of a lot of products submitted for inspection with lifetimes governed by the M-O-EE distribution.

For a random sample of size n from this distribution, the total log-likelihood function for the parameter vector $\theta = (\alpha, \lambda)^T$ is given by

$$\ell(\theta) = n \log(\alpha \lambda) - \lambda \sum_{i=1}^n x_i - 2 \sum_{i=1}^n \log(1 - \bar{\alpha} e^{-\lambda x_i}).$$

The components of the score vector $U_\theta = (U_\alpha, U_\lambda)^T$ are

$$U_\alpha = \frac{n}{\alpha} - 2 \sum_{i=1}^n \frac{e^{-\lambda x_i}}{1 - \bar{\alpha} e^{-\lambda x_i}}, \quad U_\lambda = \frac{n}{\lambda} - \sum_{i=1}^n x_i - 2 \bar{\alpha} \sum_{i=1}^n \frac{x_i e^{-\lambda x_i}}{1 - \bar{\alpha} e^{-\lambda x_i}}.$$

Setting $U_\theta = \mathbf{0}$ and solving them simultaneously yields the MLEs $\hat{\alpha}$ and $\hat{\lambda}$ of α and λ , respectively. The observed information matrix for the parameter vector $\theta = (\alpha, \lambda)^T$ is given in the Appendix. The estimates of $\check{\alpha}$ and $\check{\lambda}$, come from the simultaneous solution of the non linear equations

$$\frac{1}{n} \sum_{i=0}^n (1 - \bar{\alpha} e^{-\lambda x_i}) + \frac{\alpha \log(\alpha)}{1 - \alpha} = 0, \quad \frac{1}{n} \sum_{i=1}^n e^{-\lambda x_i} (2 - \bar{\alpha} e^{-\lambda x_i}) - 1 = 0.$$

The simulation of the M-O-EE independent deviates can be performed using

$$X = -\frac{1}{\lambda} \log \left(\frac{1 - U}{1 - \bar{\alpha} U} \right),$$

where $U \sim \mathcal{U}(0,1)$.

The evaluation of point estimation was performed based on the following quantities for each sample size: (i) mean; (ii) relative bias (the relative bias of an estimator $\hat{\theta}$ of a parameter θ is defined as $\{E(\hat{\theta} - \theta)\} / \theta$, its estimate being obtained by estimating $E(\hat{\theta})$ by Monte Carlo); (iii) root mean squared error, \sqrt{MSE} , where MSE is the mean squared error estimated from R Monte Carlo replications. The sample sizes are taken as $n = 50$ and 150 . The values of the parameters were set at $\lambda = 0.5$ and $\alpha = 0.2, 0.6, 1.5, 1.8, 2.0, 2.5, 3.0, 4.0, 5.0, 7.0, 10$ and 20 .

The point estimates are presented in Tables I and II for $n = 50$ and $n = 150$, respectively. From these tables, note that the root mean squared errors of $\hat{\alpha}$ and $\check{\alpha}$ increase with α whereas the root mean squared error of $\hat{\lambda}$ decreases with α . The relative bias of $\hat{\lambda}$ also decreases as α increases. Additionally, in most of the cases, the relative bias of $\hat{\alpha}$ is smaller than the relative bias (in absolute value) of $\check{\alpha}$. Also, the relative bias and the root mean squared error of $\hat{\lambda}$ is smaller than the relative bias (in absolute value) and the root mean squared error of $\check{\lambda}$ in most of the cases. Also noteworthy is that as the sample size increases, the root mean squared error decreases. In general, the maximum likelihood method yields better estimates of α and λ than the estimation-type method of moments.

Tables III and IV evaluate the overall performance of each of the two different estimators, for each value of n . Each entry in Table III corresponds to a specific estimator and a specific value of n . We obtain what we call the integrated relative bias squared norm (Cribari-Neto and Vasconcellos 2002). This is computed as

$$\text{IRBSQ} = \sqrt{\frac{1}{12} \sum_{h=1}^{12} r_h^2},$$

where the r_h 's ($h = 1, \dots, 12$) correspond to the twelve different values of the relative bias of each estimator. Similarly, Table IV provides the average root mean squared error expressed as

$$\text{ARMSE} = \frac{1}{12} \sum_{h=1}^{12} \sqrt{MSE}_h,$$

where \sqrt{MSE}_h is the root mean squared error corresponding to the twelve different values of the \sqrt{MSE} of each estimator. For each value of n , Table III provides a measure of the overall performance of the estimators regarding the bias, and Table IV gives a measure of their overall performance regarding the root mean squared error. In short, the figures in these tables reveal that the maximum likelihood method should be preferred than the estimation-type method of moments in order to estimate the model parameters.

APPLICATION

As Marshall and Olkin (1997) pointed out the M-O-EE distribution considered in the previous section may be an alternative to the Weibull and gamma distributions. In what follows, we shall present an empirical application to a real data set in which the M-O-EE distribution may be preferred than the Weibull and gamma models. We consider the active repair times (hours) for an airborne communication transceiver given in Jørgensen (1982). All the computations were done using the Ox matrix programming language (Doornik 2006).

Table V lists the MLEs of the parameters (standard errors in parentheses) and the values of the log-likelihood functions. The M-O-EE distribution yields the highest value of the log-likelihood function. Plots

TABLE I
Point estimates for $n=50$ and different values of α .

Maximum likelihood						
α	Estimates of α			Estimates of λ		
	Mean	Rel. Bias	$\sqrt{\text{MSE}}$	Mean	Rel. Bias	$\sqrt{\text{MSE}}$
0.2	0.2815	0.4073	0.2041	0.6324	0.2649	0.3495
0.6	0.7503	0.2505	0.4355	0.5572	0.1144	0.1953
1.5	1.8160	0.2107	0.9866	0.5329	0.0658	0.1358
1.8	2.1746	0.2081	1.1797	0.5300	0.0600	0.1278
2.0	2.4145	0.2073	1.3108	0.5285	0.0571	0.1236
2.5	3.0171	0.2068	1.6465	0.5258	0.0516	0.1156
3.0	3.6235	0.2078	1.9928	0.5239	0.0478	0.1099
4.0	4.8465	0.2116	2.7143	0.5215	0.0429	0.1021
5.0	6.0819	0.2164	3.4705	0.5199	0.0399	0.0970
7.0	8.5851	0.2264	5.0745	0.5181	0.0362	0.0905
10.0	12.4085	0.2409	7.6766	0.5166	0.0332	0.0849
20.0	25.6060	0.2803	17.6683	0.5147	0.0294	0.0771
Method of moments						
α	Estimates of α			Estimates of λ		
	Mean	Rel. Bias	$\sqrt{\text{MSE}}$	Mean	Rel. Bias	$\sqrt{\text{MSE}}$
0.2	0.3781	0.8906	0.3438	0.8216	0.6432	0.6851
0.6	0.8682	0.4469	0.3256	0.6601	0.3201	0.2320
1.5	1.1207	-0.2528	0.4437	0.4197	-0.1607	0.1129
1.8	1.2927	-0.2818	0.6802	0.4064	-0.1872	0.1328
2.0	1.4299	-0.2850	0.8329	0.4033	-0.1935	0.1429
2.5	1.8539	-0.2584	1.2343	0.4058	-0.1884	0.1582
3.0	2.3946	-0.2018	1.7244	0.4160	-0.1680	0.1671
4.0	3.6432	-0.0892	3.0690	0.4338	-0.1324	0.1843
5.0	5.0381	0.0076	4.8922	0.4446	-0.1108	0.1971
7.0	8.2577	0.1797	11.7298	0.4591	-0.0818	0.3087
10.0	13.4647	0.3465	25.7848	0.4603	-0.0794	0.3387
20.0	28.3708	0.4185	53.4771	0.4343	-0.1315	0.2879

MSE: mean squared error; Rel. Bias: relative bias.

of the estimated density of all fitted models are given in Figure 1. Note that the M-O-EE model provides a better fit than the other models. In Figure 2, we plot the estimated hazard rate for the M-O-EE, Weibull and gamma distributions. Notice that the estimated hazard ratio of the M-O-EE distribution is decreasing and belongs to the interval $[\hat{\lambda}; \hat{\lambda} \hat{\alpha}] = [0.1618; 0.3935]$, expected.

Now, we shall apply formal goodness-of-fit tests in order to verify which distribution fits better to these data. We apply the Cramér-von Mises (W^*) and Anderson-Darling (A^*) statistics. The statistics W^* and A^* are described in details in Chen and Balakrishnan (1995). In general, the smaller the values of the statistics W^* and A^* , the better the fit to the data. Let $H(x; \theta)$ be the cdf, where the form of H is known but θ (a k -dimensional parameter vector, say) is unknown. To obtain the statistics W^* and A^* , one can proceed as

TABLE II
Point estimates for $n=50$ and different values of α .

Maximum likelihood						
α	Estimates of α			Estimates of $\hat{\lambda}$		
	Mean	Rel. Bias	$\sqrt{\text{MSE}}$	Mean	Rel. Bias	$\sqrt{\text{MSE}}$
0.2	0.2254	0.1268	0.0893	0.5438	0.0877	0.1697
0.6	0.6470	0.0784	0.2028	0.5194	0.0388	0.1030
1.5	1.5980	0.0653	0.4625	0.5113	0.0225	0.0740
1.8	1.9158	0.0643	0.5517	0.5103	0.0205	0.0700
2.0	2.1278	0.0639	0.6119	0.5098	0.0195	0.0678
2.5	2.6588	0.0635	0.7647	0.5088	0.0176	0.0637
3.0	3.1906	0.0635	0.9207	0.5082	0.0163	0.0607
4.0	4.2570	0.0642	1.2413	0.5073	0.0146	0.0566
5.0	5.3265	0.0653	1.5720	0.5068	0.0135	0.0538
7.0	7.4736	0.0677	2.2596	0.5061	0.0122	0.0503
10.0	10.7113	0.0711	3.3452	0.5056	0.0112	0.0473
20.0	21.6129	0.0806	7.3070	0.5049	0.0098	0.0429

Method of moments						
α	Estimates of α			Estimates of $\hat{\lambda}$		
	Mean	Rel. Bias	$\sqrt{\text{MSE}}$	Mean	Rel. Bias	$\sqrt{\text{MSE}}$
0.2	0.2881	0.4407	0.2331	0.6595	0.3191	0.4515
0.6	0.7899	0.3165	0.2714	0.6090	0.2180	0.1732
1.5	1.2103	-0.1932	0.3948	0.4364	-0.1272	0.0936
1.8	1.4839	-0.1756	0.5785	0.4397	-0.1206	0.1080
2.0	1.6932	-0.1534	0.6985	0.4453	-0.1095	0.1133
2.5	2.2637	-0.0945	1.0076	0.4595	-0.0811	0.1192
3.0	2.8963	-0.0346	1.3633	0.4731	-0.0539	0.1212
4.0	4.2494	0.0624	2.3092	0.4919	-0.0162	0.1262
5.0	5.6095	0.1219	3.5487	0.5003	0.0005	0.1330
7.0	8.5945	0.2278	8.6495	0.5117	0.0235	0.1461
10.0	13.3775	0.3378	15.3995	0.5182	0.0365	0.1606
20.0	32.1990	0.6099	53.5464	0.5267	0.0534	0.1878

MSE: mean squared error; Rel. Bias: relative bias.

TABLE III
Integrated relative bias squared norm.

n	Estimates of α		Estimates of $\hat{\lambda}$	
	MLE	MME	MLE	MME
50	0.2458	0.3720	0.0940	0.2480
150	0.0749	0.2818	0.0314	0.1312

MLE: maximum likelihood estimate.

MME: modified moment estimate.

TABLE IV
Average root mean squared error.

n	Estimates of α		Estimates of λ	
	MLE	MME	MLE	MME
50	3.6967	8.7115	0.1341	0.2456
150	1.6107	7.3334	0.0716	0.1611

MLE: maximum likelihood estimate.
MME: modified moment estimate.

TABLE V
MLEs of the model parameters and log-likelihood functions.

Distribution	Estimates			$\ell(\hat{\theta})$
	α	λ	γ	
M-O-EE	0.4112 (0.2407)	0.1618 (0.0691)	0.9066 (0.1002)	-98.4665
Weibull		0.3169 (0.0751)	0.9453 (0.1777)	-99.4840
Gamma		0.2517 (0.0618)		-99.8571

MLEs: maximum likelihood estimate.
M-O-EE: Marshall-Olkin extended exponential.

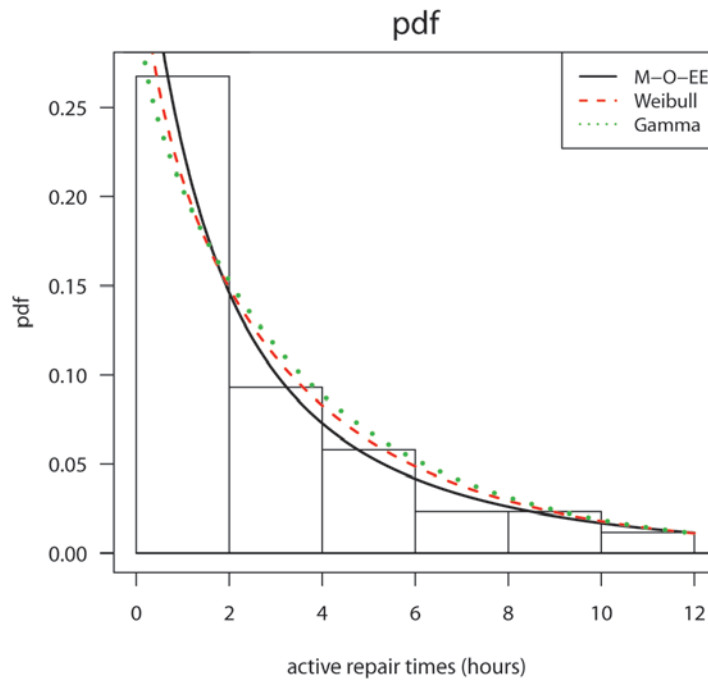


Figure 1 - Estimated pdf of the Marshall-Olkin extended exponential (M-O-EE), Weibull and gamma models for a real data set.

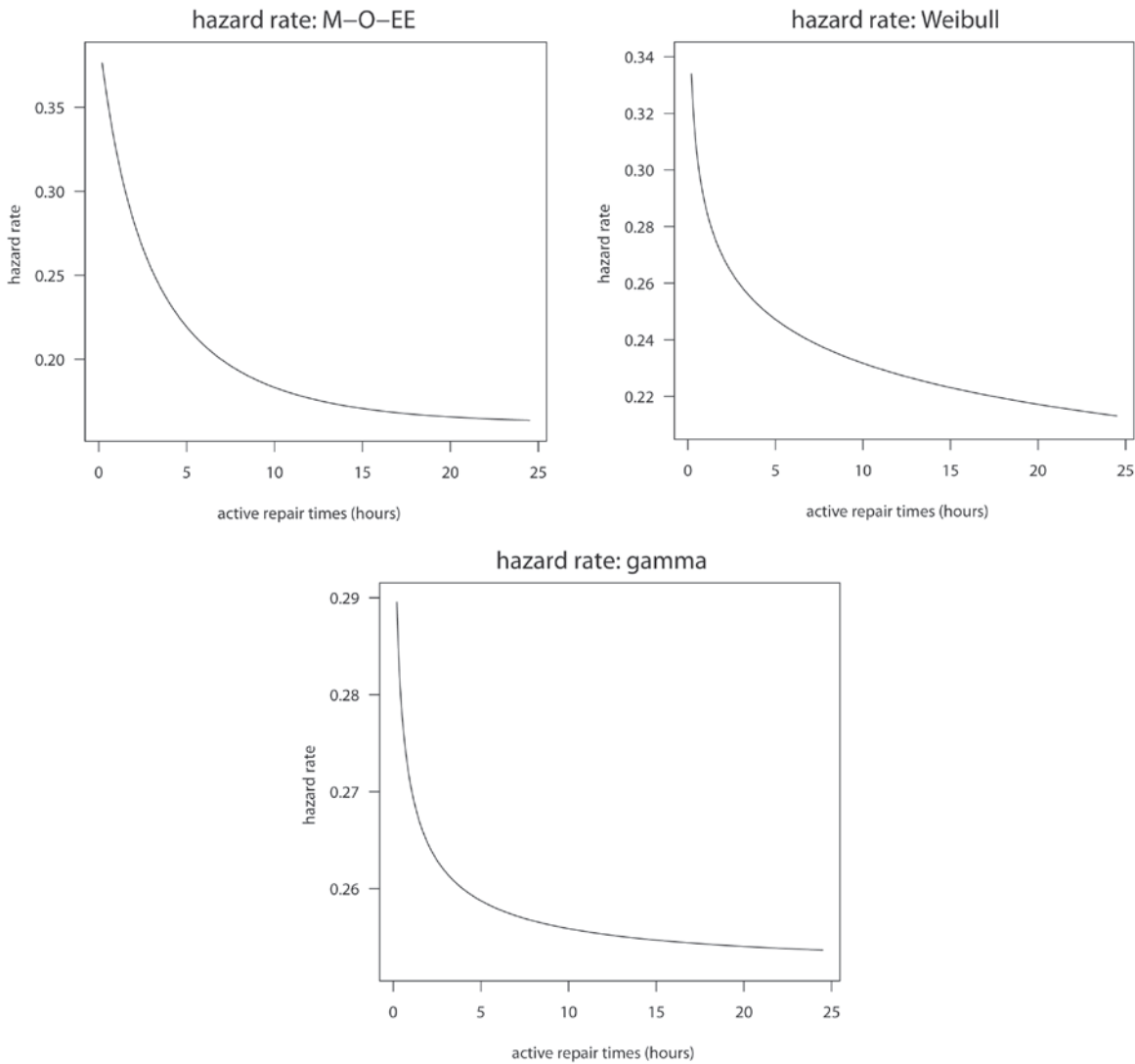


Figure 2 - Estimated hazard rate of the Marshall-Olkin extended exponential (M-O-EE), Weibull and gamma models for a real data set.

follows: (i) Compute $v_i H(x; \hat{\theta})$, where the x_i 's are in ascending order; (ii) Compute $y_i = \Phi^{-1}(v_i)$, where $\Phi(\cdot)$ is the standard normal cdf and $\Phi^{-1}(\cdot)$ its inverse; (iii) Compute $u_i = \Phi \{(y_i - \bar{y})/ s_y\}$, where $\bar{y} = (1/n)\sum_{i=1}^n y_i$ and $s_y^2 = (n - 1)^{-1} \sum_{i=1}^n (y_i - \bar{y})^2$; (iv) Calculate

$$W^2 \sum_{i=1}^n \left\{ u_i - \frac{(2_i - 1)^2}{2n} \right\} + \frac{1}{12n}$$

and

$$A^2 = -n - \frac{1}{n} \sum_{i=1}^n \{ (2_i - 1) \ln(u_i) + (2n + 1 - 2_i) \ln(1 - u_i) \};$$

(v) Modify W^2 into $W^* = (1 + 0.5 / n)$ and A^2 into $A^* = A^2(1 + 0.75 / n + 2.25 / n^2)$. The values of the statistics W^* and A^* for all models are given in Table VI. According to these statistics, the M-O-EE model fits the current data better than the other models.

TABLE VI
MLEs of the model parameters and log-likelihood functions.

Distribution	Statistics	
	W^*	A^*
M-O-EE	0.08525	0.58468
Weibull	0.12316	0.83717
Gamma	0.22368	1.48280

MLEs: maximum likelihood estimate.

M-O-EE: Marshall-Orkin extended exponential.

CONCLUDING REMARKS

We study some mathematical properties of the Marshall and Olkin's family of distributions (Marshall and Olkin 1997). This family is defined by adding a parameter to a baseline distribution, giving more flexibility to model various type of data. In the last few years, several authors proposed Marshall-Olkin models to extend well-known distributions (see, for example, Ghitany 2005, Zhang and Xie 2007, Ristić et al. 2007, Ghitany et al. 2007, Ghitany and Kotz 2007, Gómez-Déniz 2010). In this article, we derive various structural properties of the Marshall-Olkin extended-F distribution not explored before including simple expansions for the density function and explicit expressions for the moments, moments of the order statistics and Rényi entropy. Our formulas related with this class of distributions are manageable, and with the use of modern computer resources with analytic and numerical capabilities, may turn into adequate tools comprising the arsenal of applied statisticians. Several Marshall-Olkin extended-F distributions are proposed and some of their mathematical properties are given. We discuss maximum likelihood estimation of the model parameters and propose an alternative estimation method. Monte Carlo simulation experiments are also considered. Finally, an empirical application to a real data set is presented.

ACKNOWLEDGMENTS

We gratefully acknowledge grants from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) Brazil. We thank an anonymous referee for helpful comments that lead to several improvements in this paper.

RESUMO

Marshall and Olkin (1997) introduziram um método interessante de adicionar um parâmetro a uma distribuição especificada para obter uma família ampla de distribuições. Entretanto, eles não investigaram propriedades matemáticas gerais dessa família. Nós deduzimos para essa classe de distribuições, expansões gerais para a função densidade, expressões explícitas para os momentos e momentos das estatísticas de ordem. Vários modelos especiais são investigados. Nós discutimos a estimação dos parâmetros do modelo. Uma aplicação de um conjunto de dados reais é apresentada para fins ilustrativos.

Palavras-chave: distribuição estendida Marshall-Olkin, estimação por máxima verossimilhança, simulação de Monte Carlo, entropia de Rényi.

REFERENCES

- BARRETO-SOUZA W, DE MORAIS AL AND CORDEIRO GM. 2011. The Weibull-geometric distribution. *J Stat Comput Sim* 40: 798-811.
- CARONI C. 2010. Testing for the Marshall-Olkin extended form of the Weibull distribution. *Stat Pap* 51: 325-336.
- CHEN G AND BALAKRISHNAN N. 1995. A general purpose approximate goodness-of-fit test. *J Qual Technol* 27: 154-161.
- CRIBARI-NETO F AND VASCONCELLOS KLP. 2002. Nearly unbiased maximum likelihood estimation for the beta distribution. *J Stat Comput Sim* 72: 107-118.
- DOORNIK JA. 2006. *An Object-Oriented Matrix Language - Ox 4: 5th ed.* Timberlake Consultants Press: London.
- ECONOMOU P AND CARONI C. 2007. Parametric proportional odds frailty models. *Commun Stat Simulat* 36: 579-592.
- GHITANY ME. 2005. Marshall-Olkin extended Pareto distribution and its application. *Int J Appl Math* 18: 17-32.
- GHITANY ME, AL-HUSSAINI EK AND AL-JARALLAH RA. 2005. Marshall-Olkin extended Weibull distribution and its application to censored data. *J Appl Stat* 32: 1025-1034.
- GHITANY ME, AL-AWADHI FA AND ALKHALFAN LA. 2007. Marshall-Olkin extended Lomax distribution and its application to censored data. *Commun Stat Theory M* 36: 1855-1866.
- GHITANY ME AND KOTZ S. 2007. Reliability properties of extended linear failure-rate distributions. *Probab Eng Inform Sc* 21: 441-450.
- GÓMEZ-DÉNIZ E. 2010. Another generalization of the geometric distribution. *Test* 19: 399-415.
- GREENWOOD JA, LANDWEHR JM, MATALAS NC AND WALLIS JR. 1979. Probability weighted moments: definition and relation to parameters of several distributions expressible in inverse form. *Water Resour Res* 15: 1049-1054.
- GUPTA RD AND KUNDU D. 1999. Generalized exponential distributions. *Aust N Z J Stat* 41: 173-188.
- GUPTA RC, LVIN S AND PENG C. 2010. Estimating turning points of the failure rate of the extended Weibull distribution. *Comput Stat Data An* 54: 924-934.
- GUPTA RD AND PENG C. 2009. Estimating reliability in proportional odds ratio models. *Comput Stat Data An* 53: 1495-1510.
- JONES MC. 2009. Kumaraswamy's distribution: A beta-type distribution with some tractability advantages. *Stat Methodol* 6: 70-81.
- JØRGENSEN B. 1982. *Statistical Properties of the Generalized Inverse Gaussian Distribution.* Springer-Verlag, New York.
- KUMARASWAMY P. 1980. A generalized probability density function for double-bounded random processes. *J Hydrol* 46: 79-88.
- LAM KF AND LEUNG TL. 2001. Marginal likelihood estimation for proportional odds models with right censored data. *Lifetime Data Anal* 7: 39-54.
- LEMONTE AJ. 2011. Improved point estimation for the Kumaraswamy distribution. *J Stat Comput Sim* 81: 1971-1982.
- MARSHALL AW AND OLKIN I. 1997. A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families. *Biometrika* 84: 641-652.
- MARSHALL AW AND OLKIN I. 2007. *Life Distributions. Structure of Nonparametric, Semiparametric and Parametric Families.* Springer, New York.
- NADARAJAH S. 2005. Exponentiated Pareto distributions. *Statistics* 39: 255-260.
- NANDA AK AND DAS S. 2012. Stochastic orders of the Marshall-Olkin extended distribution. *Stat Probabil Lett* 82: 295-302.
- NOCEDAL J AND WRIGHT SJ. 1999. *Numerical Optimization.* Springer: New York.
- PRESS WH, TEULOSKY SA, VETTERLING WT AND FLANNERY BP. 2007. *Numerical Recipes in C: The Art of Scientific Computing: 3rd ed.,* Cambridge University Press.
- RAO GS, GHITANY ME AND KANTAM RRL. 2009. Reliability test plans for Marshall-Olkin extended exponential distribution. *Appl Math Sci* 55: 2745-2755.
- RISTIĆ MM, JOSE KK AND ANCY J. 2007. A Marshall-Olkin gamma distribution and minification process. *STARS: Stress and Anxiety Research Society* 11: 107-117.
- SONG KS. 2001. Rényi information, loglikelihood and an intrinsic distribution measure. *em J Stat Plan Infer* 93: 51-69.
- ZHANG T AND XIE M. 2007. Failure data analysis with extended Weibull distribution. *Commun Stat Simulat* 36: 579-592.

APPENDIX

The observed information matrix for $\theta = (\alpha, \lambda)^T$, $J_n(\theta)$, is given by

$$J_n(\theta) = \begin{bmatrix} U_{\alpha,\alpha} & U_{\alpha,\lambda} \\ U_{\lambda,\alpha} & U_{\lambda,\lambda} \end{bmatrix},$$

where the elements are

$$U_{\alpha\alpha} = -\frac{n}{\alpha^2} + 2 \sum_{i=1}^n \left(\frac{e^{-\lambda x_i}}{1 - \alpha e^{-\lambda x_i}} \right)^2, \quad U_{\alpha\lambda} = U_{\lambda\alpha} = 2 \sum_{i=1}^n \frac{x_i e^{-\lambda x_i}}{(1 - \alpha e^{-\lambda x_i})^2},$$

$$U_{\lambda\lambda} = \frac{n}{\lambda^2} + 2 \sum_{i=1}^n \frac{x_i^2 e^{-\lambda x_i}}{(1 - \alpha e^{-\lambda x_i})^2}.$$