Santos-Neto *et al. Journal of Statistical Distributions and Applications* 2014, **1**:9
http://www.jsdajournal.com/content/1/1/9

Journal of Statistical Distributions and Applications
a SpringerOpen Journal

RESEARCH Open Access

The Marshall-Olkin extended Weibull family of distributions

Manoel Santos-Neto[1*], Marcelo Bourguignon[2*], Luz M Zea[3], Abraão DC Nascimento[4] and Gauss M Cordeiro[4]

*Correspondence:
mn.neco@gmail.com;
m.p.bourguignon@gmail.com
[1] Departamento de Estatística, Universidade Federal de Campina Grande, Bodocongó, 58429-970, Campina Grande, PB, Brazil
[2] Departamento de Estatística, Universidade Federal do Piauí, Ininga, 64049-550, Teresina, PI, Brazil
Full list of author information is available at the end of the article

Abstract
We introduce a new class of models called the Marshall-Olkin extended Weibull family of distributions based on the work by Marshall and Olkin (Biometrika 84:641–652, 1997). The proposed family includes as special cases several models studied in the literature such as the Marshall-Olkin Weibull, Marshall-Olkin Lomax, Marshal-Olkin Fréchet and Marshall-Olkin Burr XII distributions, among others. It defines at least twenty-one special models and thirteen of them are new ones. We study some of its structural properties including moments, generating function, mean deviations and entropy. We obtain the density function of the order statistics and their moments. Special distributions are investigated in some details. We derive two classes of entropy and one class of divergence measures which can be interpreted as new goodness-of-fit quantities. The method of maximum likelihood for estimating the model parameters is discussed for uncensored and multi-censored data. We perform a simulation study using Markov Chain Monte Carlo method in order to establish the accuracy of these estimators. The usefulness of the new family is illustrated by means of two real data sets.

Mathematics Subject Classification (2010): 60E05; 62F03; 62F10; 62P10

Keywords: Extended Weibull distribution; Hazard rate function; Marshall-Olkin distribution; Maximum likelihood estimation; Survival function

1 Introduction

The Weibull distribution has assumed a prominent position as statistical model for data from reliability, engineering and biological studies (McCool 2012). This model has been exaustively used for describing *hazard rates* – an important quantity of survival analysis. In the context of monotone hazard rates, some results from the literature suggest that the Weibull law is a reasonable choice due to its negatively and positively skewed density shapes. However, this distribution is not a good model for describing phenomenon with non-monotone failure rates, which can be found on data from applications in reliability and biological studies. Thus, extended forms of the Weibull model have been sought in many applied areas. As a solution for this issue, the inclusion of additional parameters to a well-defined distribution has been indicated as a good methodology for providing more flexible new classes of distributions.

Marshall and Olkin (1997) derived an important method of including an extra shape parameter to a given baseline model thus defining an extended distribution. The Marshall and Olkin ("$\mathcal{MO}$" for short) transformation furnishes a wide range of behaviors with respect to the baseline distribution. The geometrical and inferential properties associated with the generated distribution depend on the values of the extra parameter. These characteristics provide more flexibility to the $\mathcal{MO}$ generated distributions. Considering the proportional odds model, Sankaran and Jayakumar (2008) presented a detailed discussion about the physical interpretation of the $\mathcal{MO}$ family.

This family has a relationship with the odds ratio associated with the baseline distribution. Let X be a distributed $\mathcal{MO}$ random variable which describes the lifetime relative to each individual in the population with a vector of p-covariates $\boldsymbol{z} = (z_1, \ldots, z_p)^\top$, where $(\cdot)^\top$ denotes the transposition operator. Then, the cumulative distribution function (cdf) of X is given by

$$\overline{F}(x\,;\,\boldsymbol{z}) = \frac{k(\boldsymbol{z})\,\overline{G}(x)}{1-[\,1-k(\boldsymbol{z})]\;\overline{G}(x)}, \tag{1}$$

where $k(\boldsymbol{z}) = \lambda_G(x)/\lambda_F(x\,;\,\boldsymbol{z})$ is a non-negative function such that $\boldsymbol{z}$ is independent of the time x, $\lambda_F(x\,;\,\boldsymbol{z})$ is the proportional odds model [for a discussion about such modeling, see Sankaran and Jayakumar (2008)] and $\lambda_G(x) = G(x)/\overline{G}(x)$ represents an arbitrary odds for the baseline distribution.

In this paper, we consider $k(\boldsymbol{z}) = \delta$. Before, however, it is important to highlight two important properties of the $\mathcal{MO}$ transformation: (i) the stability and (ii) geometric extreme stability (Marshall and Olkin 1997). In other words, the $\mathcal{MO}$ distribution possesses a stability property in the sense that if the method is applied twice, it returns to the same distribution. In addition, the following stochastic behavior can also be verified: let $\{X_1, \ldots, X_N\}$ be a random sample from the population random variable equipped with the survival function (1) at $k(\boldsymbol{z}) = \delta$. Suppose that N has the geometric distribution with probability p and that this quantity is independent of X_i, for $i = 1, \ldots, N$. Then, $U = min(X_1, \ldots, X_N)$ and $V = max(X_1, \ldots, X_N)$ are random variables having survival functions (1) such that $k(\boldsymbol{z})$ can be equal to p and p^{-1}, respectively, i.e., the $\mathcal{MO}$ transform satisfies the geometric extreme stability property.

Due to these advantages, many papers have employed the $\mathcal{MO}$ transformation. In Marshall and Olkin work, the exponential and Weibull distributions were generalized. Subsequently, the $\mathcal{MO}$ extension was applied to several well-known distributions: Weibull (Ghitany *et al.* 2005, Zhang and Xie 2007), Pareto (Ghitany 2005), gamma (Ristić *et al.* 2007), Lomax (Ghitany *et al.* 2007) and linear failure-rate (Ghitany and Kotz 2007) distributions. More recently, general results have been addressed by Barreto-Souza *et al.* (2013) and Cordeiro and Lemonte (2013). In this paper, we aim to apply the $\mathcal{MO}$ generator to the extended Weibull ($\mathcal{EW}$) class of distributions to obtain a new more flexible family to describe reliability data. The proposed family can also be applied to other fields including business, environment, informatics and medicine in the same way as it was originally done with the Birnbaum-Saunders and other lifetime distributions.

Let $\overline{G}(x) = 1-G(x)$ and $g(x) = dG(x)/dx$ be the survival and density functions of a continuous random variable Y with baseline cdf $G(x)$. Then, the $\mathcal{MO}$ extended distribution has survival function given by

$$\overline{F}(x;\delta) = \frac{\delta\overline{G}(x)}{1-\overline{\delta}\,\overline{G}(x)} = \frac{\delta\overline{G}(x)}{G(x)+\delta\overline{G}(x)}, \quad x \in \mathcal{X} \subseteq \mathbb{R},\ \delta > 0, \tag{2}$$

where $\overline{\delta} = 1-\delta$.

Clearly, $\delta = 1$ implies $\overline{F}(x) = \overline{G}(x)$. The family (2) has probability density function (pdf) given by

$$f(x;\delta) = \frac{\delta g(x)}{[1 - \overline{\delta}\,\overline{G}(x)]^2}, \quad x \in \mathcal{X} \subseteq \mathbb{R},\ \delta > 0.$$

Its hazard rate function (hrf) becomes

$$\tau(x;\delta) = \frac{g(x)}{\overline{G}(x)[1 - \overline{\delta}\,\overline{G}(x)]}, \quad x \in \mathcal{X} \subseteq \mathbb{R},\ \delta > 0.$$

Further, the class of extended Weibull ($\mathcal{EW}$) distributions pioneered by Gurvich *et al.* (1997) has achieved a prominent position in lifetime models. Its cdf is given by

$$G(x;\alpha,\boldsymbol{\xi}) = 1 - \exp[-\alpha H(x;\boldsymbol{\xi})], \quad x \in \mathcal{D} \subseteq \mathbb{R}_+,\ \alpha > 0, \tag{3}$$

where $H(x;\boldsymbol{\xi})$ is a non-negative monotonically increasing function which depends on the parameter vector $\boldsymbol{\xi}$. The corresponding pdf is given by

$$g(x;\alpha,\boldsymbol{\xi}) = \alpha \exp[-\alpha H(x;\boldsymbol{\xi})]\ h(x;\boldsymbol{\xi}), \tag{4}$$

where $h(x;\boldsymbol{\xi})$ is the derivative of $H(x;\boldsymbol{\xi})$.

Different expressions for $H(x;\boldsymbol{\xi})$ in Equation (3) define important models such as:

(i) $H(x;\boldsymbol{\xi}) = x$ gives the exponential distribution;
(ii) $H(x;\boldsymbol{\xi}) = x^2$ leads to the Rayleigh (Burr type-X) distribution;
(iii) $H(x;\boldsymbol{\xi}) = \log(x/k)$ leads to the Pareto distribution;
(iv) $H(x;\boldsymbol{\xi}) = \beta^{-1}[\exp(\beta x) - 1]$ gives the Gompertz distribution.

In this paper, we derive a new family of distributions by compounding the $\mathcal{MO}$ and $\mathcal{EW}$ classes. We define a new generated family in order to provide a "better fit" in certain practical situations. The compounding procedure follows by taking the $\mathcal{EW}$ class (3) as the baseline model in Equation (2). The *Marshall-Olkin extended Weibull* ($\mathcal{MOEW}$) *family* of distributions contains some special models as those listed in Table 1 with the corresponding $H(\cdot;\cdot)$ and $h(\cdot;\cdot)$ functions and the parameter vectors.

The paper unfolds as follows. Section 2 presents the cdf and pdf of the proposed distribution and some expansions for the density function. The main statistical properties of the new family are derived in Section 3 including the moments, moment generating function (mgf) and incomplete moments, quantile function (qf), random number generator, skewness and kurtosis measures, order statistics, mean deviations and average lifetime functions. In Section 4, we derive four measures of information theory: Shannon and Rényi entropies, cross entropy and Kullback-Leibler divergence. The maximum likelihood method to estimate the model parameters is adopted in Section 5. Two special models are studied in some details in Section 6. We perform a simulation study using Monte Carlo's experiments in order to assess the accuracy of the maximum likelihood estimators (MLEs) in Section 7.1 and two applications to real data in Section 7.2. Conclusions and some future lines of research are addressed in Section 8.

2 The $\mathcal{MOEW}$ family

The cdf of the new family of distributions is given by

$$F(x;\delta,\alpha,\boldsymbol{\xi}) = \frac{1 - \exp[-\alpha H(x;\boldsymbol{\xi})]}{1 - \overline{\delta}\exp[-\alpha H(x;\boldsymbol{\xi})]}, \quad x \in \mathcal{D}, \tag{5}$$

Table 1 Special models and the corresponding functions $H(x;\,\xi)$ and $h(x;\,\xi)$

Distribution	$H(x;\xi)$	$h(x;\xi)$	α	ξ	References
Exponential ($x \geq 0$)	x	1	α	$\emptyset$	Johnson *et al.* (1994)
Pareto ($x \geq k$)	$\log(x/k)$	$1/x$	α	k	Johnson *et al.* (1994)
Burr XII ($x \geq 0$)	$\log(1+x^c)$	$cx^{c-1}/(1+x^c)$	α	c	Rodriguez (1977)
Lomax ($x \geq 0$)	$\log(1+x)$	$1/(1+x)$	α	$\emptyset$	Lomax (1954)
Log-logistic ($x \geq 0$)	$\log(1+x^c)$	$cx^{c-1}/(1+x^c)$	1	c	Fisk (1961)
Rayleigh ($x \geq 0$)	x^2	$2x$	α	$\emptyset$	Rayleigh (1880)
Weibull ($x \geq 0$)	x^γ	$\gamma x^{\gamma-1}$	α	γ	Johnson *et al.* (1994)
Fréchet ($x \geq 0$)	$x^{-\gamma}$	$-\gamma x^{-(\gamma+1)}$	α	γ	Fréchet (1927)
Linear failure rate($x \geq 0$)	$ax+bx^2/2$	$a+bx$	1	$[a,b]$	Bain (1974)
Modified Weibull ($x \geq 0$)	$x^\gamma\exp(\lambda x)$	$x^{\gamma-1}\exp(\lambda x)(\gamma+\lambda x)$	α	$[\gamma,\,\lambda]$	Lai *et al.* (2003)
Weibull extension ($x \geq 0$)	$\lambda[\exp(x/\lambda)^\beta-1]$	$\beta\exp(x/\lambda)^\beta(x/\lambda)^{\beta-1}$	α	$[\gamma,\,\lambda,\,\beta]$	Xie *et al.* (2002)
Phani ($0<\mu<x<\sigma<\infty$)	$[(x-\mu)/(\sigma-x)]^\beta$	$\beta[(x-\mu)/(\sigma-x)]^{\beta-1}[(\sigma-\mu)/(\sigma-t)^2]$	α	$[\mu,\,\sigma,\,\beta]$	Phani (1987)
Weibull Kies ($0<\mu<x<\sigma<\infty$)	$(x-\mu)^{\beta_1}/(\sigma-x)^{\beta_2}$	$(x-\mu)^{\beta_1-1}(\sigma-x)^{-\beta_2-1}[\beta_1(\sigma-x)+\beta_2(x-\mu)]$	α	$[\mu,\,\sigma,\,\beta_1,\,\beta_2]$	Kies (1958)
Additive Weibull ($x \geq 0$)	$(x/\beta_1)^{\alpha_1}+(x/\beta_2)^{\alpha_2}$	$(\alpha_1/\beta_1)(x/\beta_1)^{\alpha_1-1}+(\alpha_2/\beta_2)(x/\beta_2)^{\alpha_2-1}$	1	$[\alpha_1,\,\alpha_2,\,\beta_1,\,\beta_2]$	Xie and Lai (1995)
Traditional Weibull ($x \geq 0$)	$x^b[\exp(cx^d-1)]$	$bx^{b-1}[\exp(cx^d)-1]+cdx^{b+d-1}\exp(cx^d)$	α	$[b,\,c,\,d]$	Nadarajah and Kotz (2005)
Gen. power Weibull ($x \geq 0$)	$[1+(x/\beta)^{\alpha_1}]^\theta-1$	$(\theta\alpha_1/\beta)[1+(x/\beta)^{\alpha_1}]^{\theta-1}(x/\beta)^{\alpha_1}$	1	$[\alpha_1,\,\beta,\,\theta]$	Nikulin and Haghighi (2006)
Flexible Weibull extension($x \geq 0$)	$\exp(\gamma x-\beta/x)$	$\exp(\gamma x-\beta/x)(\gamma+\beta/x^2)$	1	$[\gamma,\,\beta]$	Bebbington *et al.* (2007)
Gompertz ($x \geq 0$)	$\beta^{-1}[\exp(\beta x)-1]$	$\exp(\beta x)$	α	β	Gompertz (1825)
Exponential power ($x \geq 0$)	$\exp[(\lambda x)^\beta]-1$	$\beta\lambda\exp[(\lambda x)^\beta](\lambda x)^{\beta-1}$	1	$[\lambda,\,\beta]$	Smith and Bain (1975)
Chen ($x \geq 0$)	$\exp(x^b)-1$	$bx^{b-1}\exp(x^b)$	α	b	Chen (2000)
Pham ($x \geq 0$)	$(a^x)^\beta-1$	$\beta(a^x)^\beta\log(a)$	1	$[a,\,\beta]$	Pham (2002)

where $\alpha > 0$ and $\delta > 0$. Using (5), we can express its survival function as

$$\overline{F}(x;\delta,\alpha,\boldsymbol{\xi}) = \frac{\delta \exp[-\alpha H(x;\boldsymbol{\xi})]}{1-\overline{\delta}\exp[-\alpha H(x;\boldsymbol{\xi})]}, \quad x \in \mathcal{D} \tag{6}$$

and the associated hrf reduces to

$$\tau(x;\delta,\alpha,\boldsymbol{\xi}) = \frac{\alpha\, h(x;\boldsymbol{\xi})}{1-\overline{\delta}\exp[-\alpha H(x;\boldsymbol{\xi})]}, \quad x \in \mathcal{D}. \tag{7}$$

The corresponding pdf is given by

$$f(x;\delta,\alpha,\boldsymbol{\xi}) = \frac{\delta\,\alpha\, h(x;\boldsymbol{\xi})\exp[-\alpha H(x;\boldsymbol{\xi})]}{\{1-\overline{\delta}\exp[-\alpha H(x;\boldsymbol{\xi})]\}^2}, \tag{8}$$

where $H(x;\boldsymbol{\xi})$ can be any special distribution listed in Table 1.

Hereafter, let X be a random variable having the $\mathcal{MOEW}$ pdf (8) with parameters δ, α and $\boldsymbol{\xi}$, say $X \sim \mathcal{MOEW}(\delta,\alpha,\boldsymbol{\xi})$. Equation (8) extends several distributions which have been studied in the literature.

The $\mathcal{MO}$ Pareto (Ghitany 2005) is obtained by taking $H(x;\,\boldsymbol{\xi}) = \log(x/k)$ $(x \geq k)$. Further, for $H(x;\,\boldsymbol{\xi}) = x^{\gamma}$ we obtain the $\mathcal{MO}$ Weibull (Ghitany *et al.* 2005, Zhang and Xie 2007). The $\mathcal{MO}$ Lomax (Ghitany *et al.* 2007) and $\mathcal{MO}$ log-logistic are derived from (8) by taking $H(x;\,\boldsymbol{\xi}) = \log(1+x^c)$ with $c = 1$ and $H(x;\,\boldsymbol{\xi}) = \log(1+x^c)$ with $\alpha = 1$, respectively. For $H(x;\,\boldsymbol{\xi}) = a\,x + b\,x^2/2$ and $\alpha = 1$, Equation (8) reduces to the $\mathcal{MO}$ linear failure rate (Ghitany and Kotz 2007). In the same way, for $H(x;\,\boldsymbol{\xi}) = \log(1+x^c)$, we have the $\mathcal{MO}$ Burr XII (Jayakumar and Mathew 2008). Finally, we obtain the $\mathcal{MO}$ Fréchet (Krishna *et al.* 2013) from Equation (8) by setting $H(x;\,\boldsymbol{\xi}) = x^{-\gamma}$. Table 1 displays some useful quantities and corresponding parameter vectors for special distributions.

A general approximate goodness-of-fit test for the null hypothesis $H_0 : X_1, \ldots, X_n$ with X_i following $F(x;\boldsymbol{\theta})$, where the form of F is known but the p-vector $\boldsymbol{\theta} = (\delta,\alpha,\boldsymbol{\xi})^\top$ is unknown, was proposed by Chen and Balakrishnan (1995). This method is based on the Cramér-von Mises (CM) and Anderson-Darling (AD) statistics and, in general, the smaller the values of these statistics, the better the fit. In this paper, such methodology is applied to provide goodness-of-fit tests for the distributions under study.

Some results in the following sections can be obtained numerically in any software such as MAPLE (Garvan 2002), MATLAB (Sigmon and Davis 2002), MATHEMATICA (Wolfram 2003), Ox (Doornik 2007) and R (R Development Core Team 2009). The Ox (for academic purposes) and R are freely available at `http://www.doornik.com` and `http://www.r-project.org`, respectively. The results can be computed by taking in the sums a large positive integer value in place of ∞.

2.1 Expansions for the density function

For any positive real number a, and for $|z| < 1$, we have the generalized binomial expansion

$$(1-z)^{-a} = \sum_{k=0}^{\infty} \frac{(a)_k}{k!}\, z^k, \tag{9}$$

where $(a)_k = \Gamma(a+k)/\Gamma(a) = a(a+1)\ldots(a+k-1)$ is the ascending factorial and $\Gamma(\cdot)$ is the gamma function. Applying (9) to (8), for $0 < \delta < 1$, gives

$$f(x;\delta,\alpha,\boldsymbol{\xi}) = \sum_{j=0}^{\infty} \eta_j\, g(x;(j+1)\alpha,\boldsymbol{\xi}), \tag{10}$$

where $\eta_j = \delta\bar{\delta}^j$ and $g(x;(j+1)\alpha,\boldsymbol{\xi})$ denotes the $\mathcal{EW}$ density function with parameters $(j+1)\alpha$ and $\boldsymbol{\xi}$. Otherwise, for $\delta > 1$, after some algebra, we can express (8) as

$$f(x;\delta,\alpha,\boldsymbol{\xi}) = \frac{g(x;\alpha,\boldsymbol{\xi})}{\delta\,\{1-(1-1/\delta)\,[\,1-\exp(-\alpha H(x;\boldsymbol{\xi}))]\,\}^2}. \tag{11}$$

In this case, we can verify that $|(1-1/\delta)\,[\,1-\exp(-\alpha H(x;\boldsymbol{\xi}))]\,| < 1$. Then, applying twice the expansion (9) in Equation (11), we obtain

$$f(x;\delta,\alpha,\boldsymbol{\xi}) = \sum_{j=0}^{\infty} \nu_j\, g(x;(j+1)\alpha,\boldsymbol{\xi}), \tag{12}$$

where

$$\nu_j = \nu_j(\delta) = \frac{(-1)^j}{\delta(j+1)!}\sum_{k=j}^{\infty}(k+1)!\,(1-1/\delta)^k.$$

We can verify that $\sum_{j=0}^{\infty}\eta_j = \sum_{j=0}^{\infty}\nu_j = 1$. Then, the $\mathcal{MOEW}$ density function can be expressed as an infinite linear combination of $\mathcal{EW}$ densities. Equations (10) and (12) have the same form except for the coefficients $\eta_j's$ in (10) and $\nu_j's$ in (12). They depend only on the generator parameter δ. For simplicity, we can write

$$f(x;\delta,\alpha,\boldsymbol{\xi}) = \sum_{j=0}^{\infty} w_j\, g(x;(j+1)\alpha,\boldsymbol{\xi}), \tag{13}$$

where

$$w_j = \begin{cases} \eta_j, & \text{if } 0 < \delta < 1,\\ \nu_j, & \text{if } \delta > 1, \end{cases}$$

and η_j and ν_j are given by (10) and (12), respectively. Thus, some mathematical properties of (13) can be obtained directly from those $\mathcal{EW}$ properties. For example, the ordinary, incomplete, inverse and factorial moments and the mgf of X follow immediately from those quantities of the $\mathcal{EW}$ distribution.

3 General properties

3.1 Moments, generating function and incomplete moments

The nth ordinary moment of X can be obtained from (13) as

$$\mathrm{E}(X^n) = \sum_{j=0}^{\infty} w_j\, E(Y_j^n),$$

where from now on $Y_j \sim \mathcal{EW}((j+1)\alpha,\boldsymbol{\xi})$ denotes a random variable having the $\mathcal{EW}$ density function $g(y;(j+1)\alpha,\boldsymbol{\xi})$.

The mgf and the kth incomplete moment of X follow from (13) as

$$M_X(t) = E\left(\mathrm{e}^{tX}\right) = \sum_{j=0}^{\infty} w_j\, M_j(t)$$

and

$$T_k(z) = \sum_{j=0}^{\infty} w_j\, T_k^{(j)}(z), \tag{14}$$

where $M_j(t)$ is the mgf of Y_j and $T_k^{(j)}(z) = \int_{-\infty}^{z} x^k\, g(x;(j+1)\alpha,\boldsymbol{\xi})\mathrm{d}x$ comes directly from the $\mathcal{EW}$ model.

3.2 Quantile function and random number generator

The qf of X follows by inverting (5) and it can be expressed in terms of $H^{-1}(\cdot)$ as

$$Q(u) = H^{-1}\left(\frac{1}{\alpha}\log\left(\frac{1-\overline{\delta}\,u}{1-u}\right), \boldsymbol{\xi}\right). \tag{15}$$

In Table 2, we provide the function $H^{-1}(x;\boldsymbol{\xi})$ for some special models.

Hence, the generator for X can be given by the algorithm:

Algorithm 1 Random number generator for the $\mathcal{MOEW}$ distribution

1: Generate $U \sim U(0,1)$.
2: Specify a function $H^{-1}(\cdot;\cdot)$ such as anyone in Table 2 and use (15).
3: Obtain an outcome of X by $X = Q(U)$.

The $\mathcal{MOEW}$ distributions can be very useful in modeling lifetime data and practitioners may be interested in fitting one of these models. We provide a script using the R language to generate the density, distribution function, hrf, qf, random numbers, Anderson-Darling test, Cramer-von Mises test and likelihood ratio (LR) tests. This script can be be obtained from the authors upon requested.

3.3 Mean deviations

The mean deviations of X about the mean and the median are given by

$$\delta_1 = \int_{\mathcal{D}} |x-\mu| f(x;\delta,\alpha,\boldsymbol{\xi})\,\mathrm{d}x \quad\text{and}\quad \delta_2 = \int_{\mathcal{D}} |x-M| f(x;\delta,\alpha,\boldsymbol{\xi})\,\mathrm{d}x,$$

respectively, where $\mu = E(X)$ denotes the mean and $M = Median(X)$ the median. The median follows from the nonlinear equation $F(M;\delta,\alpha,\boldsymbol{\xi}) = 1/2$. So, these quantities reduce to

$$\delta_1 = 2\,\mu\,F(\mu;\delta,\alpha,\boldsymbol{\xi}) - 2\,T_1(\mu) \quad\text{and}\quad \delta_2 = \mu - 2\,T_1(M),$$

where $T_1(z)$ is the first incomplete moment of X obtained from (14) as

$$T_1(z) = \sum_{j=0}^{\infty} w_j\, T_1^{(j)}(z),$$

and $T_1^{(j)}(z) = \int_{-\infty}^{z} x\, g(x;(j+1)\alpha,\boldsymbol{\xi})\,\mathrm{d}x$ is the first incomplete moment of Y_j.

An important application of the mean deviations is related to the Bonferroni and Lorenz curves. These curves are useful in economics, reliability, demography, medicine

Table 2 The $H^{-1}(x;\xi)$ function

Distribution	$H^{-1}(x;\xi)$
Exponential power	$\frac{[\log(x+1)]^{1/\beta}}{\lambda}$
Chen	$\left[\log(x+1)\right]^{1/\beta}$
Weibull extension	$\lambda\left[\log\left(\frac{x}{\lambda}+1\right)\right]^{1/\beta}$
Log-Weibull	$\sigma\log(x)+\mu$
Kies	$\frac{x^{1/\beta}\sigma+\mu}{x^{1/\beta}+1}$
Gen. Power Weibull	$\beta\left[(x+1)^{1/\theta}-1\right]^{1/\alpha_1}$
Gompertz	$\frac{\log(\beta x+1)}{\beta}$
Pham	$\left[\frac{\log(x+1)}{\log(a)}\right]^{1/\beta}$

and other fields. For a given probability p, they are defined by $B(p) = T_1(q)/(p\,\mu)$ and $L(p) = T_1(q)/\mu$, respectively, where $q = Q(p)$ is the qf of X given by (15) at $u = p$.

3.4 Average lifetime and mean residual lifetime functions

The average lifetime is given by

$$t_m = \int_0^\infty [1 - F(x;\delta,\alpha,\boldsymbol{\xi})]\,\mathrm{d}x = \sum_{j=0}^\infty w_j \int_0^\infty \overline{G}(x;(j+1)\alpha,\boldsymbol{\xi})\mathrm{d}x.$$

In fields such as actuarial sciences, survival studies and reliability theory, the mean residual lifetime has been of much interest; see, for a survey, Guess and Proschan (1988). Given that there was no failure prior to x_0, the residual life is the period from time x_0 until the time of failure. The mean residual lifetime is given by

$$\begin{aligned} m(x_0;\delta,\alpha,\boldsymbol{\xi}) &= \mathrm{E}\,(X - x_0 | X \geq x_0;\delta,\alpha,\boldsymbol{\xi}) = \int\limits_{\{x:x>x_0\}} \frac{(x-x_0)f(x;\delta,\alpha,\boldsymbol{\xi})}{\Pr(X>x_0)}\mathrm{d}x \\ &= [\Pr(X>x_0)]^{-1}\int_0^\infty y f(x_0+y;\delta,\alpha,\boldsymbol{\xi})\mathrm{d}y \\ &= \left[\overline{F}(x_0;\delta,\alpha,\boldsymbol{\xi})\right]^{-1}\sum_{j=0}^\infty w_j \int_0^\infty y\, g(x_0+y;(j+1)\alpha,\boldsymbol{\xi})\mathrm{d}y. \end{aligned}$$

The last integral can be computed from the baseline $\mathcal{EW}$ distribution. Further, $m(x_0;\delta,\alpha,\boldsymbol{\xi}) \to E(X)$ as $x_0 \to 0$.

4 Information theory measures

The seminal idea about information theory was pioneered by Hartley (1928), who defined a logarithmic measure of information for communication. Subsequently, Shannon (1948) formalized this idea by defining the entropy and mutual information concepts. The relative entropy notion (which would later be called *divergence*) was proposed by Kullback and Leibler (1951). The Kullback-Leibler's measure can be understood like a comparison criterion between two distributions. In this section, we derive two classes of entropy measures and one class of divergence measures which can be understood as new goodness-of-fit quantities such those discussed by Seghouane and Amari (2007). All these measures are defined for one element or between two elements in the $\mathcal{MOEW}$ family.

4.1 Rényi entropy

The Rényi entropy of X with pdf (8) is given by

$$H_{\mathrm{R}}^s(X) = \frac{1}{1-s}\log\left(\int_{\mathcal{D}} f(x;\delta,\alpha,\boldsymbol{\xi})^s\mathrm{d}x\right),$$

where $s \in (0,1) \cup (1,\infty)$.

It is a difficult problem to obtain $H_{\mathrm{R}}^s(X)$ in closed-form for the $\mathcal{MOEW}$ family. So, we derive an expansion for this quantity.

By using (9), $f(x;\delta,\alpha,\boldsymbol{\xi})^s$ can be expanded as

$$f(x;\delta,\alpha,\boldsymbol{\xi})^s = \sum_{j=0}^\infty w_j' \exp[-(j+s)\alpha H(x;\boldsymbol{\xi})]\ h(x;\boldsymbol{\xi})^s, \tag{16}$$

where

$$w_j' = \begin{cases} \eta_j'(\alpha,\delta) = \frac{\alpha^s\delta^s(2s)_j\,\bar{\delta}^j}{j!}, & \text{for } 0<\delta<1,\\ \nu_j'(\alpha,\delta) = \frac{\alpha^s\delta^{-s}}{j!}\sum_{k=0}^{\infty}\frac{(2s)_k(k)_j}{k!}\,(1-1/\delta)^k, & \text{for } \delta>1. \end{cases}$$

The proof of this expansion is given in Appendix 8.

Finally, based on Equation (16), the Rényi entropy can be expressed as

$$H_{\mathrm{R}}^s(X) = \frac{1}{1-s}\log\left\{\sum_{j=0}^{\infty} w_j' \int_{\mathcal{D}} \exp[\,-(j+s)\alpha H(x;\boldsymbol{\xi})]\,h(x;\boldsymbol{\xi})^s \mathrm{d}x\right\}.$$

An advantage of this expansion is its dependence of an integral which has closed-form for some $\mathcal{EW}$ distributions.

4.2 Shannon entropy

The Shannon entropy of X is given by

$$H_{\mathrm{S}}(X) = \mathrm{E}_X\left\{-\log[f(X;\delta,\alpha,\boldsymbol{\xi})]\right\},$$

where the log-likelihood function corresponding to one observation follows from (8) as

$$\log[f(x;\delta,\alpha,\boldsymbol{\xi})] = \log(\delta\alpha) + \log[\,h(x;\boldsymbol{\xi})] - \alpha H(x;\boldsymbol{\xi}) - 2\log\left\{1-\bar{\delta}\exp[\,-\alpha H(x;\boldsymbol{\xi})]\,\right\}.$$

Thus, it can be reduced to

$$H_{\mathrm{S}}(X) = -\log(\alpha\delta) + 2\mathrm{E}\left\{\log\left[1-\bar{\delta}\bar{G}(X;\boldsymbol{\xi})\right]\right\} - \mathrm{E}\left\{\log\left[h(X;\boldsymbol{\xi})\right]\right\} + \alpha\mathrm{E}\left[H(X;\boldsymbol{\xi})\right].$$

4.3 Cross entropy and Kullback-Leibler divergence and distance

Let X and Y be two random variables with common support $\mathbb{R}_+$ whose densities are $f_X(x;\boldsymbol{\theta}_1)$ and $f_Y(y;\boldsymbol{\theta}_2)$, respectively. Cover and Thomas (1991) defined the *cross entropy* as

$$C_X(Y) = \mathrm{E}_X\left\{-\log\left[f_Y(X;\boldsymbol{\theta}_2)\right]\right\} = -\int_0^{\infty} f_X(z;\boldsymbol{\theta}_1)\log\left[f_Y(z;\boldsymbol{\theta}_2)\right]\mathrm{d}z.$$

We consider that $X \sim \mathcal{MOEW}(\delta_x,\alpha_x,\boldsymbol{\xi}_x)$ and $Y \sim \mathcal{MOEW}(\delta_y,\alpha_y,\boldsymbol{\xi}_y)$. After some algebraic manipulations, we obtain

$$\begin{aligned} C_X(Y) &= -\int_{\mathcal{D}} f_X(z;\delta_x,\alpha_x,\boldsymbol{\xi}_x)\log\left[f_Y(z;\delta_y,\alpha_y,\boldsymbol{\xi}_y)\right]\mathrm{d}z\\ &= -\log\left(\delta_y\alpha_y\right) - \mathrm{E}_X\left\{\log\left[h(X;\boldsymbol{\xi}_y)\right]\right\} + \alpha_y\,\mathrm{E}_X\left[H(X;\boldsymbol{\xi}_y)\right]\\ &\quad + 2\mathrm{E}_X\left\{\log\left[1-\bar{\delta}\bar{G}(X;\boldsymbol{\xi}_y)\right]\right\}. \end{aligned} \tag{17}$$

An important measure in information theory is the Kullback-Leibler divergence given by

$$D(X||Y) = C_X(Y) - H_{\mathrm{S}}(X) = \mathrm{E}_X\left\{\log\left[\frac{f_X(X;\delta_x,\alpha_x,\boldsymbol{\xi}_x)}{f_Y(X;\delta_y,\alpha_y,\boldsymbol{\xi}_y)}\right]\right\}. \tag{18}$$

Applying (4.2) and (17) in Equation (18) gives

$$\begin{aligned} D(X||Y) =& \log\left(\frac{\delta_x\alpha_x}{\delta_y\alpha_y}\right) + \mathrm{E}_X\left\{\log\left[\frac{h(X;\boldsymbol{\xi}_x)}{h(X;\boldsymbol{\xi}_y)}\right]\right\} + 2\mathrm{E}_X\left\{\log\left[\frac{1-\bar{\delta}\bar{G}(X;\boldsymbol{\xi}_y)}{1-\bar{\delta}\bar{G}(X;\boldsymbol{\xi}_x)}\right]\right\}\\ &+ \alpha_y E_X[\,H(X;\boldsymbol{\xi}_y)] - \alpha_x E_X[\,H(X;\boldsymbol{\xi}_x)]\,. \end{aligned} \tag{19}$$

According to Cover and Thomas (1991), the Kullback-Leibler measure $D(X||Y)$ is the quantification of the error considering that the Y model is true when the data follow the X distribution. For example, this measure has been proposed as essential parts of test statistics, which has seen strongly applied to contexts of radar synthetic aperture image processing in both univariate (Nascimento et al. 2010) and polarimetric (or multivariate) (Nascimento et al. 2014) perspectives.

In order to work with measures that satisfied the non-negativity, symmetry and definiteness properties, Nascimento et al. (2010) considered the symmetrization of (19)

$$d_{\mathrm{KL}}(X,Y) = \frac{1}{2}\,[D(X||Y) \,+\, D(Y||X)] = \int_{\mathcal{D}} \underbrace{\left(f_X\left(x;\delta_x,\alpha_x,\boldsymbol{\xi}_x\right) - f_Y\left(x;\delta_y,\alpha_y,\boldsymbol{\xi}_y\right)\right)\,\log\left(\frac{f_X(x;\delta_x,\alpha_x,\boldsymbol{\xi}_x)}{f_Y(x;\delta_x,\alpha_x,\boldsymbol{\xi}_x)}\right)}_{\equiv\, IntegrandKL(x,y)}\,\mathrm{d}x,$$

which is given by

$$\begin{aligned} 2\,d_{\mathrm{KL}}(X,Y) = {} & \alpha_y\,\left\{\mathrm{E}_X\left[H(X;\boldsymbol{\xi}_y)\right] - \mathrm{E}_Y\left[H(Y;\boldsymbol{\xi}_y)\right]\right\} + \alpha_x\,\left\{\mathrm{E}_Y\left[H(Y;\boldsymbol{\xi}_x)\right] - \mathrm{E}_X[H(X;\boldsymbol{\xi}_x)]\right\} \\ & + \mathrm{E}_X\left\{\log\left[\frac{h\left(X;\boldsymbol{\xi}_x\right)}{h\left(X;\boldsymbol{\xi}_y\right)}\right]\right\} \,+\, \mathrm{E}_Y\left\{\log\left[\frac{h\left(Y;\boldsymbol{\xi}_y\right)}{h\left(Y;\boldsymbol{\xi}_x\right)}\right]\right\} \\ & + 2\,\mathrm{E}_X\left\{\log\left[\frac{1-\bar{\delta}\bar{G}\left(X;\boldsymbol{\xi}_y\right)}{1-\bar{\delta}\bar{G}\left(X;\boldsymbol{\xi}_x\right)}\right]\right\} + 2\,\mathrm{E}_Y\left\{\log\left[\frac{1-\bar{\delta}\bar{G}\left(Y;\boldsymbol{\xi}_x\right)}{1-\bar{\delta}\bar{G}\left(Y;\boldsymbol{\xi}_y\right)}\right]\right\}. \end{aligned} \tag{20}$$

Although this measure does not satisfy the triangle inequality, it is usually called the *Kullback-Leibler* distance (*Jensen-Shannon* divergence). The new measure can be used to answer questions like "how could one quantify the difference in selecting the Phani model with three parameters as the baseline distribution instead of the Weibull Kies distribution which has four parameters?".

As an illustration for (20), we initially consider two distinct elements of the generated special model from the specifications: $H(x;\beta) = \beta^{-1}[\,\exp(\beta x) - 1]$ and $h(x;\beta) = \exp(\beta x)$ in (8). This model will be presented with more details in future sections and its parametric space is represented by the vector (δ,α,β). Suppose that we are interested in quantifying the influence of a nuisance degree ϵ in the parameter α over the distance between two distinct elements, $(2,1,3)$ and $(2,1+\epsilon,3)$, at such parametric space. Figure 1(a) displays

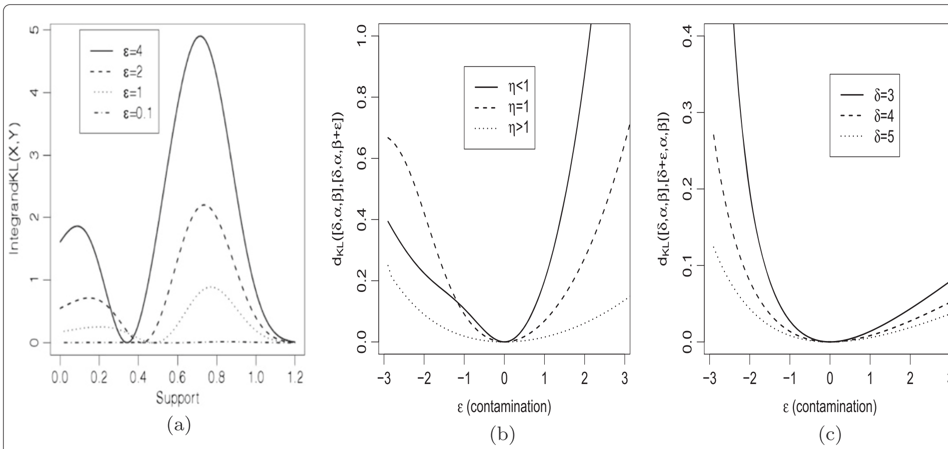


Figure 1 MSE curves for $\delta \in \{0.3, 0.8, 1, 2, 4\}$, $\eta = 0.5, 1, 2$ and $n = 200$. (a) Behavior of the function IntegrandKL. **(b)** Influence of $\eta = \alpha/\beta$ under $\beta = \delta = 3$ and $\alpha \in \{1, 3, 9\}$. **(c)** Influence of δ under (α,β)=(3,3).

the integrand of (20) for $\epsilon = 0.1, 1, 2$ and 4 for which the distances (or areas) associated with $d_{\mathrm{KL}}(X, Y)$ are 6.50×10^{-3}, 3.56×10^{-1}, 9.46×10^{-1} and 2.25, respectively. It is notable that $d_{\mathrm{KL}}(X, Y)$ takes smaller values for more closer points (or, equivalently, for more closer fits) and, therefore, (20) consists of new goodness-of-fit measures. In Figures 1(b) and 1(c), we show the influence of $\eta = \alpha/\beta$ on $d_{\mathrm{KL}}([\,\delta, \alpha, \beta\,], [\,\delta, \alpha, \beta + \epsilon\,])$ (for $\beta = \delta = 3$ and $\alpha \in \{1, 3, 9\}$) and of δ on $d_{\mathrm{KL}}([\,\delta, \alpha, \beta\,], [\,\delta + \epsilon, \alpha, \beta\,])$ (for $\beta = \alpha = 3$ and $\delta \in \{3, 4, 5\}$). For all cases, the contamination ϵ takes values in the interval $(-2.9, 2.9)$.

5 Estimation

Here, we present a general procedure for estimating the $\mathcal{MOEW}$ parameters from one observed sample and from multi-censored data. Additionally, we provide a discussion about how one can test the significance of additional parameter at the proposed class. Let $x_1, \ldots, x_n$ be a sample of size n from X. The log-likelihood function for the vector of parameters $\boldsymbol{\theta} = (\delta, \alpha, \boldsymbol{\xi}^\top)^\top$ can be expressed as

$$\ell(\boldsymbol{\theta}) = n \log(\delta\alpha) + \sum_{i=1}^{n} \log\left[h(x_i; \boldsymbol{\xi})\right] - \alpha \sum_{i=1}^{n} H(x_i; \boldsymbol{\xi}) - 2 \sum_{i=1}^{n} \log\left\{1 - \overline{\delta} \exp[-\alpha H(x_i; \boldsymbol{\xi})]\right\}.$$

From the above log-likelihood, the components of the score vector, $\mathbf{U}(\theta) = (U_\delta, U_\alpha, U_{\boldsymbol{\xi}}^\top)^\top$, are given by

$$U_\delta(\boldsymbol{\theta}) = \frac{\partial \ell(\theta)}{\partial \delta} = \frac{n}{\delta} - 2 \sum_{i=1}^{n} \frac{\exp[\,-\alpha H(x_i; \boldsymbol{\xi})]}{1 - \overline{\delta} \exp[\,-\alpha H(x_i; \boldsymbol{\xi})]},$$

$$U_\alpha(\boldsymbol{\theta}) = \frac{\partial \ell(\theta)}{\partial \alpha} = \frac{n}{\alpha} - \sum_{i=1}^{n} H(x_i; \boldsymbol{\xi}) - 2\overline{\delta} \sum_{i=1}^{n} \frac{H(x_i; \boldsymbol{\xi}) \exp[\,-\alpha H(x_i; \boldsymbol{\xi})]}{1 - \overline{\delta} \exp[\,-\alpha H(x_i; \boldsymbol{\xi})]} \quad \text{and}$$

$$U_{\boldsymbol{\xi}_k}(\boldsymbol{\theta}) = \frac{\partial \ell(\theta)}{\partial \boldsymbol{\xi}_k} = \sum_{i=1}^{n} \frac{1}{h(x_i; \boldsymbol{\xi})} \frac{\partial h(x_i; \boldsymbol{\xi})}{\partial \boldsymbol{\xi}_k} - \alpha \sum_{i=1}^{n} \frac{\partial H(x_i; \boldsymbol{\xi})}{\partial \boldsymbol{\xi}_k} - 2\overline{\delta}\alpha \sum_{i=1}^{n} \frac{\partial H(x_i; \boldsymbol{\xi})}{\partial \boldsymbol{\xi}_k} \frac{\exp[\,-\alpha H(x_i; \boldsymbol{\xi})]}{1 - \overline{\delta} \exp[\,-\alpha H(x_i; \boldsymbol{\xi})]}.$$

Finally, the partitioned observed information matrix for the $\mathcal{MOEW}$ family is

$$J(\boldsymbol{\theta}) = -\begin{pmatrix} U_{\delta\delta} & U_{\delta\alpha} & | & U_{\delta\boldsymbol{\xi}}^\top \\ U_{\alpha\delta} & U_{\alpha\alpha} & | & U_{\alpha\boldsymbol{\xi}}^\top \\ -- & -- & -- & -- \\ U_{\delta\boldsymbol{\xi}} & U_{\alpha\boldsymbol{\xi}} & | & U_{\boldsymbol{\xi}\boldsymbol{\xi}} \end{pmatrix},$$

whose elements are

$$U_{\delta\delta}(\boldsymbol{\theta}) = -n\delta^{-2}, U_{\delta\alpha}(\boldsymbol{\theta}) = 2 \sum_{i=1}^{n} \frac{H(x_i; \boldsymbol{\xi}) \exp[\,-\alpha H(x_i; \boldsymbol{\xi})]}{\left\{1 - \overline{\delta} \exp[\,-\alpha H(x_i; \boldsymbol{\xi})]\right\}^2},$$

$$U_{\delta\boldsymbol{\xi}_k}(\boldsymbol{\theta}) = 2\alpha \sum_{i=1}^{n} \frac{\partial H(x_i; \boldsymbol{\xi})}{\partial \boldsymbol{\xi}_k} \frac{\exp[\,-\alpha H(x_i; \boldsymbol{\xi})]}{\left\{1 - \overline{\delta} \exp[\,-\alpha H(x_i; \boldsymbol{\xi})]\right\}^2},$$

$$U_{\alpha\alpha}(\boldsymbol{\theta}) = -\frac{n}{\alpha^2} + 2\overline{\delta} \sum_{i=1}^{n} \frac{H(x_i; \boldsymbol{\xi})^2 \exp[\,-\alpha H(x_i; \boldsymbol{\xi})]}{\left\{1 - \overline{\delta} \exp[\,-\alpha H(x_i; \boldsymbol{\xi})]\right\}^2},$$

$$U_{\alpha\xi_k}(\boldsymbol{\theta}) = -2\overline{\delta}\sum_{i=1}^{n}\frac{\partial H(x_i;\boldsymbol{\xi})}{\partial \boldsymbol{\xi}_k}\frac{\exp[\,-\alpha H(x_i;\boldsymbol{\xi})]}{1-\overline{\delta}\exp[\,-\alpha H(x_i;\boldsymbol{\xi})]}\left[1-\frac{\alpha H(x_i;\boldsymbol{\xi})}{1-\overline{\delta}\exp[\,-\alpha H(x_i;\boldsymbol{\xi})]}\right] + \sum_{i=1}^{n}\frac{\partial H(x_i;\boldsymbol{\xi})}{\partial \boldsymbol{\xi}_k} \quad \text{and}$$

$$U_{\xi_k\xi_j}(\boldsymbol{\theta}) = \sum_{i=1}^{n}\frac{1}{h(x_i;\boldsymbol{\xi})}\left[\frac{\partial^2 h(x_i;\boldsymbol{\xi})}{\partial \boldsymbol{\xi}_k\boldsymbol{\xi}_j} - \frac{1}{h(x_i;\boldsymbol{\xi})}\frac{\partial h(x_i;\boldsymbol{\xi})}{\partial \boldsymbol{\xi}_k}\frac{\partial h(x_i;\boldsymbol{\xi})}{\partial \boldsymbol{\xi}_j}\right] - \alpha\sum_{i=1}^{n}\frac{\partial^2 H(x_i;\boldsymbol{\xi})}{\partial \boldsymbol{\xi}_k\boldsymbol{\xi}_j}$$
$$-2\alpha\overline{\delta}\sum_{i=1}^{n}\frac{\exp[-\alpha H(x_i;\boldsymbol{\xi})]}{1-\overline{\delta}\exp[-\alpha H(x_i;\boldsymbol{\xi})]}\left[\frac{\partial^2 H(x_i;\boldsymbol{\xi})}{\partial \boldsymbol{\xi}_k\boldsymbol{\xi}_j} - \frac{\partial H(x_i;\boldsymbol{\xi})}{\partial \boldsymbol{\xi}_k}\frac{\alpha H(x_i;\boldsymbol{\xi})}{1-\overline{\delta}\exp[-\alpha H(x_i;\boldsymbol{\xi})]}\right].$$

When some standard regularity conditions are satisfied (Cox and Hinkley 1974), one can verify that $\sqrt{n}\left(\left[\widehat{\alpha},\,\widehat{\delta},\,\widehat{\boldsymbol{\xi}}\right]^\top - [\alpha,\,\delta,\,\boldsymbol{\xi}]^\top\right)$ converges *in distribution* to the multivariate $N_{p+2}\left(\mathbf{0},\mathcal{K}([\,\alpha,\,\delta,\,\boldsymbol{\xi}]\,)^{-1}\right)$ distribution, where p denotes the dimension of $\boldsymbol{\xi}$ and $\mathcal{K}([\,\alpha,\,\delta,\,\boldsymbol{\xi}]\,)$ is the expected information matrix for which the limit identity $\lim_{n\to\infty} J_n([\,\alpha,\,\delta,\,\boldsymbol{\xi}]\,) = \mathcal{K}([\,\alpha,\,\delta,\,\boldsymbol{\xi}]\,)$ is satisfied. Based on this result, one can compute confidence regions for the $\mathcal{MOEW}$ parameters. Such regions can be used as decision criteria in several practical situations.

For checking if δ is statistically different from one, i.e. for testing the null hypothesis $H_0 : \delta = 1$ against $H_1 : \delta \neq 1$, we use the LR statistic given by $\text{LR} = 2\left\{\ell(\widehat{\boldsymbol{\theta}}) - \ell(\widetilde{\boldsymbol{\theta}})\right\}$, where $\widehat{\boldsymbol{\theta}}$ is the vector of unrestricted MLEs under H_1 and $\widetilde{\boldsymbol{\theta}}$ is the vector of restricted MLEs under H_0. Under the null hypothesis, the limiting distribution of LR is a χ^2_1 distribution. If the test statistic exceeds the upper $100(1-\alpha)\%$ quantile of the χ^2_1 distribution, then we reject the null hypothesis.

Censored data occur very frequently in lifetime data analysis. Some mechanisms of censoring are identified in the literature as, for example, types I and II censoring (Lawless 2003). Here, we consider the general case of multi-censored data: there are $n = n_0 + n_1 + n_2$ subjects of which n_0 is known to have failed at the times $x_1, \ldots, x_{n_0}$, n_1 is known to have failed in the interval $[s_{i-1}, s_i]$, $i = 1, \ldots, n_1$, and n_2 survived to a time r_i , $i = 1, \ldots, n_2$, but not observed any longer. Note that type I censoring and type II censoring are contained as particular cases of multi-censoring. The log-likelihood function of $\boldsymbol{\theta} = (\delta, \alpha, \boldsymbol{\xi}^\top)^\top$ for this multi-censoring data reduces to

$$\ell(\boldsymbol{\theta}) = n_0\log(\delta\alpha) + \sum_{i=1}^{n_0}\log\left[h(x_i;\boldsymbol{\xi})\right] - \alpha\sum_{i=1}^{n_0}H(x_i;\boldsymbol{\xi}) - 2\sum_{i=1}^{n_0}\log\left\{1-\overline{\delta}\exp[\,-\alpha H(x_i;\boldsymbol{\xi})]\right\}$$
$$+\sum_{i=1}^{n_1}\log\left\{\frac{1-\exp[\,-\alpha\, H(s_i;\boldsymbol{\xi})]}{1-\overline{\delta}\exp[\,-\alpha\, H(s_i;\boldsymbol{\xi})]} - \frac{1-\exp[\,-\alpha\, H(s_{i-1};\boldsymbol{\xi})]}{1-\overline{\delta}\exp[\,-\alpha\, H(s_{i-1};\boldsymbol{\xi})]}\right\}$$
$$+\,n_2\log(\delta) - \alpha\sum_{i=1}^{n_2}H(r_i;\boldsymbol{\xi}) - 2\sum_{i=1}^{n_2}\log\left\{1-\overline{\delta}\exp[\,-\alpha H(r_i;\boldsymbol{\xi})]\right\}. \tag{21}$$

The score functions and the observed information matrix corresponding to (21) is too complicated to be presented here.

6 Two special models

In this section, we study two special $\mathcal{MOEW}$ models, namely the Marshall-Olkin modified Weibull ($\mathcal{MOMW}$) and Marshall-Olkin Gompertz ($\mathcal{MOG}$) distributions. We provide plots of the density and hazard rate functions for some parameters to illustrate the flexibility of these distributions.

6.1 The $\mathcal{MOMW}$ model

For $H(x;\lambda,\gamma) = x^{\gamma}\exp(\lambda x)$ and $h(x;\lambda,\gamma) = x^{\gamma-1}\exp(\lambda x)(\gamma+\lambda x)$, we obtain the $\mathcal{MOMW}$ distribution. Its density function is given by

$$f(x;\alpha,\delta,\lambda,\gamma) = \delta\alpha(\gamma+\lambda x)x^{\gamma-1}\frac{\exp[\,\lambda x - \alpha x^{\gamma}\exp(\lambda x)]}{\left\{1-\bar{\delta}\exp[\,-\alpha x^{\gamma}\exp(\lambda x)]\,\right\}^{2}}, \quad x>0,$$

where $\lambda,\gamma \geq 0$. If $\delta = 1$, it leads to the special case of the modified Weibull ($\mathcal{MW}$) distribution (Lai *et al.* 2003). In addition, when $\lambda = 0$, it gives the Weibull distribution. Its cdf and hrf are given by

$$F(x;\alpha,\delta,\lambda,\gamma) = \frac{1-\exp[\,-\alpha x^{\gamma}\exp(\lambda x)]}{1-\bar{\delta}\exp[\,-\alpha x^{\gamma}\exp(\lambda x)]}$$

and

$$\tau(x;\alpha,\delta,\lambda,\gamma) = \frac{\alpha x^{\gamma-1}\exp(\lambda x)(\gamma+\lambda x)}{1-\bar{\delta}\exp[\,-\alpha x^{\gamma}\exp(\lambda x)]},$$

respectively. In Figures 2(a), 2(b), 2(c) and 2(d), we note some different shapes of the $\mathcal{MOMW}$ pdf. Further, Figures 3(a), 3(b), 3(c) and 3(d) display plots of the $\mathcal{MOMW}$ hrf, which can have increasing, decreasing, non-monotone and bathtub forms.

The rth raw moment of the $\mathcal{MOMW}$ distribution comes from (13) as

$$\mathrm{E}\left(X^{r}\right) = \sum_{j=1}^{\infty} w_j\,\mu_r(j), \tag{22}$$

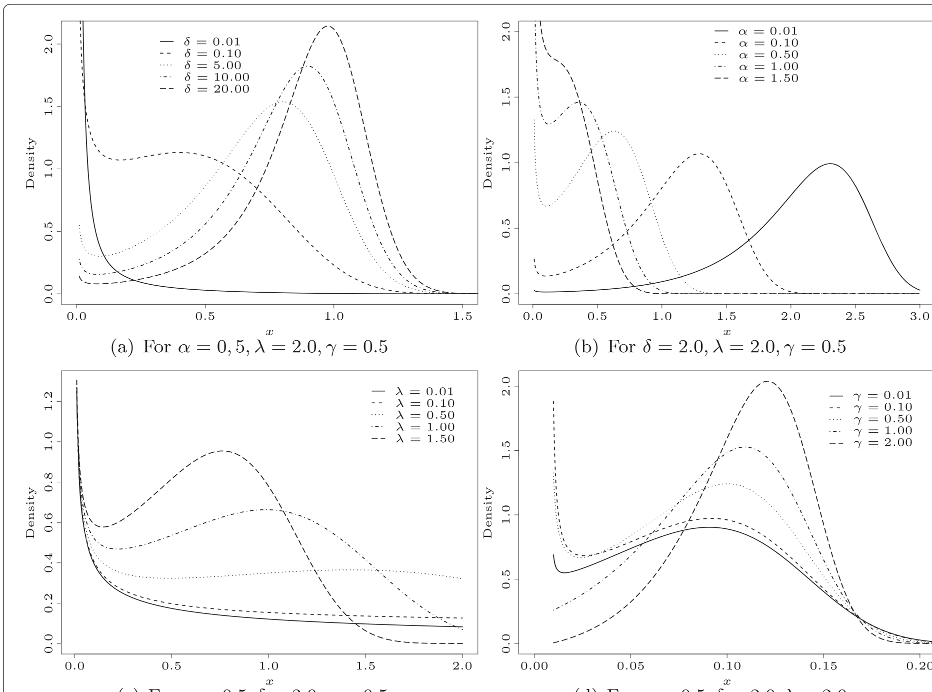
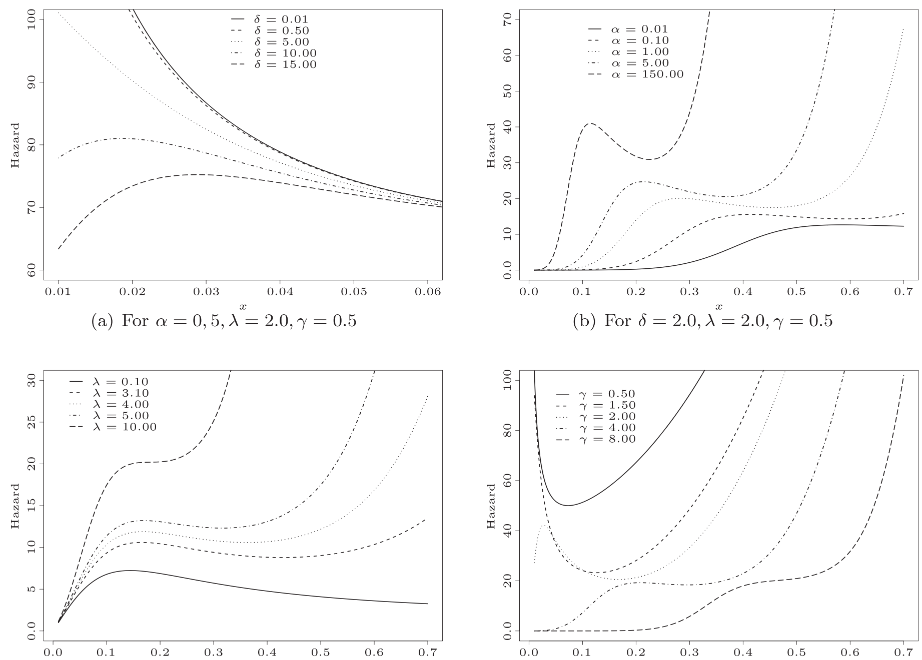


Figure 2 The $\mathcal{MOMW}$ density functions. (a) For $\alpha = 0.5, \lambda = 2.0, \gamma = 0.5$. **(b)** For $\delta = 2.0, \lambda = 2.0, \gamma = 0.5$. **(c)** For $\delta = 5.0, \delta = 2.0, \gamma = 0.5$. **(d)** For $\alpha = 0.5, \delta = 2.0, \lambda = 2.0$.



(a) For $\alpha = 0, 5, \lambda = 2.0, \gamma = 0.5$ (b) For $\delta = 2.0, \lambda = 2.0, \gamma = 0.5$

(c) For $\alpha = 0.5, \delta = 2.0, \gamma = 0.5$ (d) For $\alpha = 0.5, \delta = 2.0, \lambda = 2.0$

Figure 3 The $\mathcal{MOWN}$ hrfs. (a) For $\alpha = 0.5, \lambda = 2.0, \gamma = 0.5$. **(b)** For $\delta = 2.0, \lambda = 2.0, \gamma = 0.5$. **(c)** For $\delta = 5.0, \delta = 2.0, \gamma = 0.5$. **(d)** For $\alpha = 0.5, \delta = 2.0, \lambda = 2.0$.

where $\mu_r(j) = \int_0^\infty x^r g(x;\ (j+1)\alpha, \gamma, \lambda))\mathrm{d}x$ denotes the rth raw moment of the $\mathcal{MW}$ distribution with parameters $(j+1)\alpha, \gamma$ and λ. Carrasco *et al.* (2008) determined an infinite representation for $\mu_r(j)$ given by

$$\mu_r(j) = \sum_{i_1,\ldots,i_r=1}^{\infty} \frac{A_{i_1,\ldots,i_r}\,\Gamma(s_r/\gamma + 1)}{[\,(j+1)\alpha]^{s_r/\gamma}}, \tag{23}$$

where

$$A_{i_1,\ldots,i_r} = a_{i_1}, \ldots, a_{i_r} \quad \text{and} \quad s_r = i_1, \ldots, i_r,$$

and

$$a_i = \frac{(-1)^{i+1} i^{i-2}}{(i-1)!} \left(\frac{\lambda}{\gamma}\right)^{i-1}.$$

Hence, the $\mathcal{MOMW}$ moments can be obtained directly from (22) and (23).

Let $x_1, \ldots, x_n$ be a sample of size n from $X \sim \mathcal{MOMW}(\alpha, \delta, \lambda, \gamma)$. The log-likelihood function for the vector of parameters $\boldsymbol{\theta} = (\alpha, \delta, \lambda, \gamma)^\top$ can be expressed as

$$\ell(\boldsymbol{\theta}) = n\log(\delta\alpha) + \sum_{i=1}^{n}\log(\gamma + \lambda\, x_i) + (\gamma - 1)\sum_{i=1}^{n}\log(x_i) + \lambda\sum_{i=1}^{n} x_i - \alpha\sum_{i=1}^{n} x_i^{\lambda}\exp(\lambda\, x_i)$$
$$-2\sum_{i=1}^{n}\log\left(1 - \overline{\delta}\exp\left[-\alpha x_i^{\gamma}\exp(\lambda\, x_i)\right]\right).$$

6.2 The $\mathcal{MOG}$ model

For $H(x;\beta) = \beta^{-1}[\exp(\beta x) - 1]$ and $h(x;\beta) = \exp(\beta x)$, we obtain the $\mathcal{MOG}$ distribution. Its pdf is given by

$$f(x;\alpha,\delta,\beta) = \frac{\delta\alpha \exp\{\beta x - \alpha/\beta[\exp(\beta x) - 1]\}}{\left\{1 - \overline{\delta}\exp\{-\alpha/\beta[\exp(\beta x) - 1]\}\right\}^2}, \quad x > 0,$$

where $-\infty < \beta < \infty$. For $\delta = 1$, it follows the Gompertz distribution as a special case. The $\mathcal{MOG}$ model is a special case of the Marshall-Olkin Makeham distribution (EL-Bassiouny and Abdo 2009). The cdf and hrf of the $\mathcal{MOG}$ distribution are given by

$$F(x;\alpha,\delta,\beta) = \frac{1 - \exp\{-\alpha/\beta[\exp(\beta x) - 1]\}}{1 - \overline{\delta}\exp\{-\alpha/\beta[\exp(\beta x) - 1]\}}$$

and

$$\tau(x;\alpha,\delta,\beta) = \frac{\alpha\exp(\beta x)}{1 - \overline{\delta}\exp\{-\alpha/\beta[\exp(\beta x) - 1]\}}.$$

Figures 4(a), 4(b) and 4(c) display some plots of the density functions for some values of α, δ and β. The hrf of the Gompertz distribution is increasing ($\beta > 0$) and decreasing ($\beta < 0$). Besides these two forms, Figures 5(a), 5(b) and 5(c) indicate that the $\mathcal{MOG}$ hrf can be bathtub shaped.

From Equation (15), the $\mathcal{MOG}$ qf becomes

$$Q(u) = \beta^{-1}\log\left[\frac{\beta}{\alpha}\log\left(\frac{1-\overline{\delta}u}{1-u}\right) + 1\right].$$

Let $x_1, \ldots, x_n$ be a sample of size n from the $\mathcal{MOG}$ model. The log-likelihood function for the vector of parameters $\boldsymbol{\theta} = (\delta,\alpha,\beta)^\top$ can be expressed as

$$\ell(\boldsymbol{\theta}) = n\log(\delta\alpha) + \beta\sum_{i=1}^{n} x_i - \frac{\alpha}{\beta}\sum_{i=1}^{n}\left[\exp(\beta x_i) - 1\right] - 2\sum_{i=1}^{n}\log\left(1 - \overline{\delta}\exp\{-\alpha[\exp(\beta x_i) - 1]/\beta\}\right).$$

7 Simulation and applications

This section is divided in two parts. First, we perform a simulation study in order to assess the performance of the MLEs on some points at the parametric space of one of the special

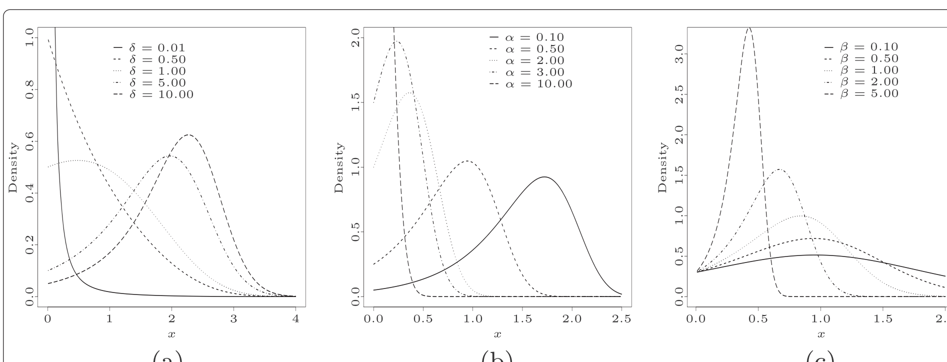


Figure 4 The $\mathcal{MOG}$ density functions. (a) For $\alpha = 0.5, \beta = 0.7$. **(b)** For $\delta = 2.0, \beta = 2.0$. **(c)** For $\delta = 5.0, \alpha = 1.5$.

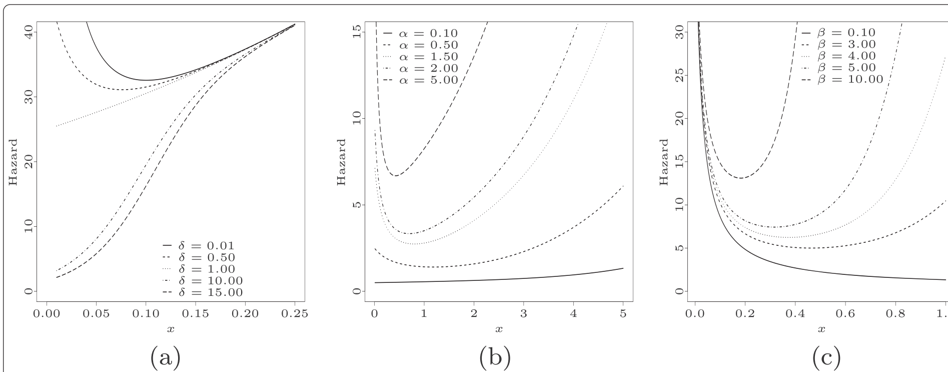


Figure 5 The $\mathcal{MOG}$ hrfs. (a) For $\alpha = 25, \beta = 2.0$. **(b)** For $\delta = 0.2, \beta = 0.5$. **(c)** For $\alpha = 0.01, \delta = 0.5$.

models. Second, an application to real data provides evidence in favor of one distribution in the $\mathcal{MOEW}$ class.

7.1 Simulation study

We present a simulation study by means of Monte Carlo's experiments in order to assess the performance of the MLEs described in Section 5. To that end, we work with the $\mathcal{MOG}$ distribution. One of advantages of this model is that its cdf has tractable analytical form. This fact implies in a simple random number generation (RNG) determined by the $\mathcal{MOG}$ qf given in Section 6.2. The $\mathcal{MOG}$ generator is illustrated in Figure 6.

The simulation study is conducted in order to quantify the influence of $\eta = \alpha/\beta$ over the estimation of the extra parameter δ. It is known that $\eta > 1$ gives the Gompertz distribution which presents mode at zero or, for $\eta < 1$, having their modes at $x^* = \beta^{-1}[1 - \log(\eta)]$. An initial discussion using the Kullback-Leibler distance derived in Section 4.3 points out that increasing the contamination (or the bias of the estimates) can affect the quality of fit.

In this study, the following scenarios are taken into account. For the sample size $n = 50, 100, 150, 200$, we adopt as the true parameters the following cases:

(i) Scenario $\eta < 1$: $(\alpha, \beta) = (1, 2)$ and $\delta \in \{0.3, 1, 4\}$;
(ii) Scenario $\eta = 1$: $(\alpha, \beta) = (2, 2)$ and $\delta \in \{0.3, 1, 4\}$;
(iii) Scenario $\eta > 1$: $(\alpha, \beta) = (4, 2)$ and $\delta \in \{0.3, 1, 4\}$.

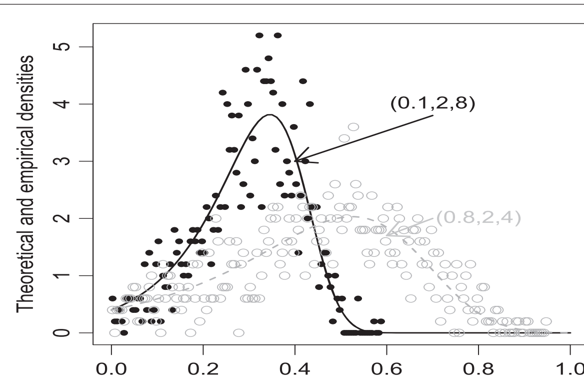


Figure 6 Illustration of the $\mathcal{MOG}$ generator for two points at the parametric space.

Also, we use 10,000 Monte Carlo's replications and, at each one of them, we quantify (i) the average of the MLEs and (ii) the mean square error (MSEs).

Table 3 gives the results of the simulation study. In general, the MLEs present smaller values of the biases and MSEs when the sample size increases. It is important to highlight the following atypical case: for the MLEs of α at the scenarios $(\alpha, \delta, \beta) \in \{(1, 4, 2), (2, 1, 2), (4, 0.3, 2), (4, 1, 2)\}$ and of δ at $(4, 0.3, 2)$, the associated biases do not have an inverse monotonic relationship with sample sizes, as expected.However, based on the fact

Table 3 Performance of the MLEs for the $\mathcal{MOG}$ distribution

(α, δ, β)	n	$\overline{\widehat{\theta}}_i(MSE(\widehat{\theta}_i))$					
		$\overline{\widehat{\alpha}}$ (MSE ($\widehat{\alpha}$))		$\overline{\widehat{\delta}}$ (MSE ($\widehat{\delta}$))		$\overline{\widehat{\beta}}$ (MSE ($\widehat{\beta}$))	
		For $\eta < 1$					
(1, 0.3, 2)	50	1.201	(2.837)	0.478	(0.883)	2.502	(1.698)
·	100	1.181	(1.745)	0.406	(0.290)	2.320	(1.238)
·	150	1.156	(1.299)	0.385	(0.195)	2.249	(1.015)
·	200	1.103	(1.008)	0.358	(0.134)	2.244	(0.899)
(1, 1, 2)	50	1.202	(1.965)	1.620	(5.938)	2.425	(1.630)
·	100	1.134	(1.199)	1.361	(2.690)	2.305	(1.145)
·	150	1.079	(0.884)	1.231	(1.638)	2.288	(0.979)
·	200	1.063	(0.735)	1.180	(1.244)	2.250	(0.845)
(1, 4, 2)	50	0.965	(0.810)	4.764	(26.798)	2.544	(1.561)
·	100	0.958	(0.544)	4.398	(14.813)	2.390	(1.025)
·	150	0.959	(0.443)	4.283	(11.454)	2.328	(0.831)
·	200	0.970	(0.369)	4.246	(8.953)	2.262	(0.653)
		For $\eta = 1$					
(2, 0.3, 2)	50	2.246	(7.571)	0.426	(0.473)	2.787	(3.543)
·	100	2.137	(4.502)	0.361	(0.172)	2.561	(2.473)
·	150	2.073	(3.279)	0.341	(0.116)	2.471	(1.981)
·	200	2.011	(2.596)	0.324	(0.083)	2.434	(1.698)
(2, 1, 2)	50	2.161	(5.462)	1.481	(4.886)	2.687	(3.051)
·	100	2.012	(3.115)	1.199	(1.798)	2.543	(2.157)
·	150	1.947	(2.277)	1.100	(1.062)	2.483	(1.763)
·	200	1.923	(1.874)	1.056	(0.787)	2.430	(1.507)
(2, 4, 2)	50	1.805	(2.404)	4.534	(21.279)	2.785	(2.783)
·	100	1.817	(1.681)	4.202	(12.456)	2.572	(1.869)
·	150	1.828	(1.390)	4.097	(9.474)	2.487	(1.527)
·	200	1.861	(1.153)	4.075	(7.495)	2.388	(1.184)
		For $\eta > 1$					
(4, 0.3, 2)	50	3.770	(13.137)	0.336	(0.191)	3.400	(6.701)
·	100	3.737	(8.129)	0.304	(0.072)	2.951	(4.152)
·	150	3.731	(6.119)	0.298	(0.051)	2.764	(3.184)
·	200	3.685	(4.865)	0.289	(0.038)	2.676	(2.613)
(4, 1, 2)	50	3.845	(13.615)	1.272	(3.153)	3.149	(6.239)
·	100	3.735	(7.757)	1.076	(1.043)	2.833	(4.060)
·	150	3.717	(5.760)	1.024	(0.634)	2.689	(3.150)
·	200	3.721	(4.759)	1.000	(0.472)	2.588	(2.601)
(4, 4, 2)	50	3.608	(8.172)	4.605	(21.140)	3.036	(5.150)
·	100	3.677	(5.234)	4.262	(11.467)	2.668	(2.989)
·	150	3.737	(4.039)	4.172	(8.228)	2.510	(2.169)
·	200	3.796	(3.247)	4.138	(6.370)	2.389	(1.588)

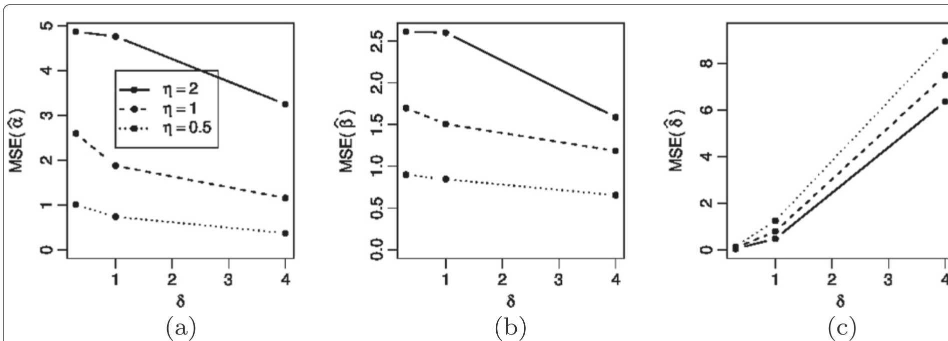


Figure 7 MSE curves for $\delta \in \{0.3, 1, 4\}$, $\eta = 0.5, 1, 2$ and $n = 200$. (a) $MSE(\widehat{\alpha})$, **(b)** $MSE(\widehat{\beta})$, **(c)** $MSE(\widehat{\delta})$.

that their MSEs tend to zero, we can expect that there exists a sample size n_0 such that biases of the MLEs decrease when the sample sizes increase from n_0.

The results provide evidence that the scenarios under the condition $\eta > 1$ yield a hard estimation (having larger variation ranges of the MSEs than those obtained for the cases when $\eta < 1$) for α and β parameters, and that the MLEs present smaller values of the MSEs under such conditions. Figure 7 illustrates the above behavior for the cases $\delta \in \{0.3, 0.8, 1, 2, 4\}$ and $n = 200$. In summary, the scenario with less numerical problems is $(\eta, \delta) = (2, 0.1)$, whereas that one which requires more attention for estimating the $\mathcal{MOG}$ parameters is $(\eta, \delta) = (0.5, 4)$.

7.2 Applications

Here, the usefulness of the $\mathcal{MOEW}$ distribution is illustrated by means of two real data sets.

7.2.1 Uncensored data

Here, we compare the fits of some special models of the $\mathcal{MOEW}$ family using a real data set. The estimation of the model parameters is performed by the maximum likelihood method discussed in Section 5. We use the `maxLik` function of the `maxLik` package in `R` language. In this function, if the argument "method" is not specified, a suitable method is selected automatically. For this application, we use the Newton-Raphson method. The data represent the percentage of body fat determined by underwater weighing for 250 men. For more details about the data see http://lib.stat.cmu.edu/datasets/bodyfat.

Table 4 provides some descriptive measures. They suggest an empirical distribution which is slightly asymmetric and platykurtic.

We compare the classical models and generalized models within the $\mathcal{MO}$ family. The null hypothesis $H_0 : \delta = 1$ is tested against $H_1 : \delta \neq 1$ using the LR statistic. The comparisons are presented in Table 5. For the $\mathcal{MOW}$ and $\mathcal{MOEP}$ models, one cannot say that the parameter δ is statistically different from one at the 10% significance level. Based on this result, we fit the $\mathcal{W}$, exponential power ($\mathcal{EP}$), $\mathcal{MOG}$ and Marshall-Olkin flexible Weibull extension ($\mathcal{MOFWE}$) models to the current data (see Table 1). These models are compared with two other three-parameter models, namely: the modified Weibull ($\mathcal{MW}$)

Table 4 Descriptive statistics

Mean	Mode	Median	Std. Desv.	Skewness	Kurtosis	Min	Max
19.30	20.40	19.25	8.23	0.19	2.62	3.00	47.50

Table 5 Comparison of fitted models using the LR test

Null hypothesis	Models	LR statistic	*p*-value
$\delta = 1$	$\mathcal{G} \times \mathcal{MOG}$	11.2963	0.0008
	$\mathcal{W} \times \mathcal{MOW}$	0.7638	0.3822
	$\mathcal{EP} \times \mathcal{MOEP}$	2.1959	0.1384
	$\mathcal{FWE} \times \mathcal{MOFWE}$	12.3659	0.0004

and generalized Birnbaum-Saunders ($\mathcal{GBS}$) (Owen 2006) distributions. The $\mathcal{GBS}$ density is given by

$$f(x;\phi,\eta,\kappa) = \frac{1}{\phi\sqrt{2\pi\eta}x^{\kappa}}\left(1-\kappa+\frac{\eta\kappa}{x}\right)\exp\left[-\frac{1}{2\phi^2}\frac{(x-\eta)^2}{\eta x^{2\kappa}}\right], \quad x>0.$$

In Table 6, we present the MLEs (standard errors in parentheses) of the parameters of the fitted $\mathcal{MOFWE}$, $\mathcal{MOG}$, $\mathcal{EP}$, $\mathcal{W}$, $\mathcal{MW}$ and $\mathcal{GBS}$ distributions. Also, we provide the goodness-of-fit measures (p-values in parentheses). Thus, these values indicate that the null models are strongly rejected for the $\mathcal{MOFWE}$ and $\mathcal{MOG}$ distributions, since the associated p-values are much lower than 0.001.

Table 7 gives the values of the Akaike information criterion (AIC), Bayesian information criterion (BIC), consistent Akaike information criterion (CAIC) and Hannan-Quinn information criterion (HQIC). Since the values of the AIC, CAIC and HQIC are smaller for the $\mathcal{MOFWE}$ distribution compared to those values of the other fitted models. Thus, this new distribution seems to be a very competitive model to explain the current data.

Figures 8(a) and 8(b) display the estimated density and survival functions of the $\mathcal{MOFWE}$ distribution. The plots confirm the excellent fit of this distribution to the data. Figure 8(c) shows that the estimated $\mathcal{MOFWE}$ hrf is an increasing curve.

7.2.2 Censored data

Now, we consider a set of remission times from 137 cancer patients [Lee and Wang (2003), pag. 231]. Lee and Wang (2003) showed that the log-logistic ($\mathcal{LL}$) model provides a good fit to the data. Ghitany *et al.* (2005) compared the fits of the $\mathcal{MOW}$ and $\mathcal{W}$ models to these data. Now, we present a more detailed study by comparing the fitted $\mathcal{W}$, $\mathcal{LL}$, $\mathcal{EP}$, $\mathcal{MOW}$, Marshall-Olkin log-logistic ($\mathcal{MOLL}$), $\mathcal{MOEP}$ and $\mathcal{GBS}$ models to these data.

Table 6 MLEs and goodness-of-fit statistics

Model	Estimates (standard errors)				Goodness-of-fit (*p*-value)	
	$\widehat{\alpha}$(or $\widehat{\phi}$)	$\widehat{\delta}$(or $\widehat{\eta}$)	$\widehat{\gamma}$ (or $\widehat{\kappa}$)	$\widehat{\beta}$ (or $\widehat{\lambda}$)	AD	CM
$\mathcal{MOFWE}$	1	2.9136	0.0552	14.3666	**0.1082**	**0.0115**
	×	(1.1321)	(0.0022)	(3.7615)	**(0.9939)**	**(0.9987)**
$\mathcal{MOG}$	0.1289	18.8183	–	0.0183	0.6825	0.0938
	(0.0151)	(2.9308)	×	(0.0063)	(0.0739)	(0.1361)
$\mathcal{EP}$	1	1	0.0359	1.7778	0.2537	0.0273
	×	×	(0.0008)	(0.0870)	(0.7301)	(0.8800)
$\mathcal{W}$	0.0004	1	2.5373	–	0.4344	0.0667
	(0.0002)	×	(0.1434)	×	(0.2985)	(0.3079)
$\mathcal{MW}$	0.0007	1	2.2292	0.0149	0.2761	0.0384
	(0.0007)	×	(0.4384)	(0.0191)	(0.6546)	(0.7094)
$\mathcal{GBS}$	1.3189	18.7623	0.1328	×	0.5672	0.0876
	(0.1847)	(0.5784)	(0.0513)	×	(0.1404)	(0.1642)

Table 7 Statistics AIC, BIC, CAIC and HQIC

Models	AIC	BIC	CAIC	HQIC
$\mathcal{MOFWE}$	**1753.989**	1764.553	**1754.087**	**1758.241**
$\mathcal{MOG}$	1767.305	1777.870	1767.403	1771.557
$\mathcal{EP}$	1764.136	1771.178	1764.184	1766.970
$\mathcal{W}$	1756.843	**1763.886**	1756.892	1759.678
$\mathcal{MW}$	1757.997	1768.561	1758.094	1762.248
$\mathcal{GBS}$	1761.136	1771.701	1761.234	1765.388

The functions $H(x;\gamma,c) = \log(1+\gamma\, x^c)$ and $h(x;\gamma,c) = \gamma\, c\, x^{c-1}/(1+\gamma\, x^c)$ are associated with the $\mathcal{LL}$ model.

The hypothesis that the underlying distribution is $\mathcal{W}$ (or $\mathcal{EP}$) versus the alternative hypothesis that the distribution is the $\mathcal{MOW}$ (or $\mathcal{MOEP}$) is rejected with p-value = 0.0055 (or p-value = <0.0001). Further, the hypothesis test that the underlying distribution is $\mathcal{LL}$ versus the $\mathcal{MOLL}$ distribution yields the p-value $= 1.0000$. Thus, we compare the $\mathcal{MOW}$, $\mathcal{MOEP}$, $\mathcal{LL}$ and $\mathcal{GBS}$ models to determine which model gives the best fit to the current data.

Table 8 lists the MLEs (and corresponding standard errors in parentheses) of the parameters and the values of the AD and CM statistics (their p-values in parentheses). The figures in this table, specially the p-values, suggest that the $\mathcal{MOW}$ distribution yields a better fit to these data than the other three distributions.

Table 9 lists the values of the AIC, BIC, CAIC and HQIC statistics. The figures in this table indicate that there is a competitiveness among the $\mathcal{MOW}$, $\mathcal{MOEP}$ and $\mathcal{LL}$ models. However, if we observe the Figures 9(a), 9(b) and 9(c), we note that the $\mathcal{MOW}$ and $\mathcal{MOEP}$ models present better fits to the current data.

Figure 9(d) really shows that the $\mathcal{MOW}$ and $\mathcal{MOEP}$ distributions present good fits to the current data. We can conclude that the $\mathcal{MOW}$ and $\mathcal{MOEP}$ distributions are excellent alternatives to explain this data set.

8 Conclusion

In this paper, the Marshall-Olkin extended Weibull family of distributions is proposed and some of its mathematical properties are studied. The maximum likelihood procedure is

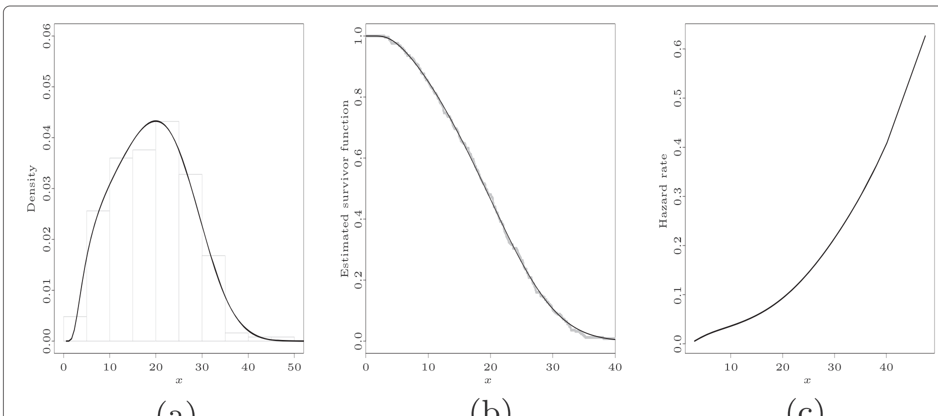


Figure 8 Plots of the estimated. (a) density, **(b)** survivor function **(c)** and hrf of the $\mathcal{MOFWE}$ model.

Table 8 MLEs and goodness-of-fit statistics

Model	Estimates (standard errors)				Goodness-of-fit (*p*-value)	
	$\widehat{\alpha}$ (or $\widehat{\phi}$)	$\widehat{\delta}$ (or $\widehat{\eta}$)	$\widehat{\gamma}$ (or $\widehat{\kappa}$ or $\widehat{\beta}$)	$\widehat{c}$ (or $\widehat{\lambda}$)	AD	CM
$\mathcal{MOW}$	0.0037	0.0736	1.5719	–	**0.1889**	**0.0264**
	(0.0043)	(0.0727)	(0.1616)	×	**(0.8994)**	**(0.8908)**
$\mathcal{MOEP}$	–	0.0233	0.0144	1.6012	0.2057	0.0279
	×	(0.0165)	(0.1423)	(0.0042)	(0.8686)	(0.8733)
$\mathcal{GBS}$	1.6313	7.1422	0.3356	–	1.2753	0.2116
	(0.1226)	(0.7374)	(0.0314)	×	(0.0025)	(0.0038)
$\mathcal{LL}$	–	–	0.0427	1.6900	0.2891	0.0380
	×	×	(0.0118)	(0.1249)	(0.6101)	(0.7164)

used for estimating the model parameters. Two special models in the family are described with some details. In order to assess the performance of the maximum likelihood estimates, a simulation study is performed by means of Monte Carlo experiments. Special models of the proposed family are compared (through goodness-of-fit measures) with other well-known lifetime models by means of two real data sets. The proposed model outperforms classical lifetime models to these data.

Appendix: An expansion for $f(x;\delta,\alpha,\xi)F(x;\delta,\alpha,\xi)^c$

Here, we obtain an expansion for the quantity $f(x;\delta,\alpha,\xi)F(x;\delta,\alpha,\xi)^c$. First, we consider an expansion for $F(x;\delta,\alpha,\xi)^c$. Based on (5), the power of the cdf can be expressed as

$$F(x;\delta,\alpha,\xi)^c = \underbrace{\{1-\exp[\,-\alpha H(x;\boldsymbol{\xi})]\,\}^c}_{\equiv A}\ \underbrace{\{1-\overline{\delta}\exp[\,-\alpha H(x;\boldsymbol{\xi})]\,\}^{-c}}_{\equiv B}.$$

Applying expansion (9), we have

$$A = \sum_{k=0}^{\infty}(-1)^k \binom{c}{k}\ \exp[\,-k\alpha H(x;\boldsymbol{\xi})]\,.$$

Now, we expand the quantity B. Equation (9) under the restriction $\delta < 1$ (implying that $\overline{\delta}\exp[\,-\alpha H(x;\boldsymbol{\xi})] < 1$) yields

$$B = \sum_{j=0}^{\infty}\frac{(c)_j}{j!}\ \overline{\delta}^j\ \exp[\,-j\alpha H(x;\boldsymbol{\xi})]\,.$$

Moreover, it is clear that $\delta = 1$ implies $B = 1$. Finally, for $\delta > 1$ (i.e., $\{1-\overline{\delta}\exp[\,-\alpha H(x;\boldsymbol{\xi})]\,\} > 1$), the quantity B can be rewritten as

$$B = \left\{1-[\,1-\{1-\overline{\delta}\exp[\,-\alpha H(x;\boldsymbol{\xi})]\,\}^{-1}]\right\}^c.$$

Table 9 Statistics AIC, BIC, CAIC and HQIC

Models	AIC	BIC	CAIC	HQIC
$\mathcal{MOW}$	**843.1171**	851.8770	**843.2975**	846.6769
$\mathcal{MOEP}$	843.1898	851.9498	843.3703	846.7497
$\mathcal{GBS}$	858.3686	867.1285	858.5490	861.9284
$\mathcal{LL}$	843.7586	**849.5986**	843.8481	**846.1318**

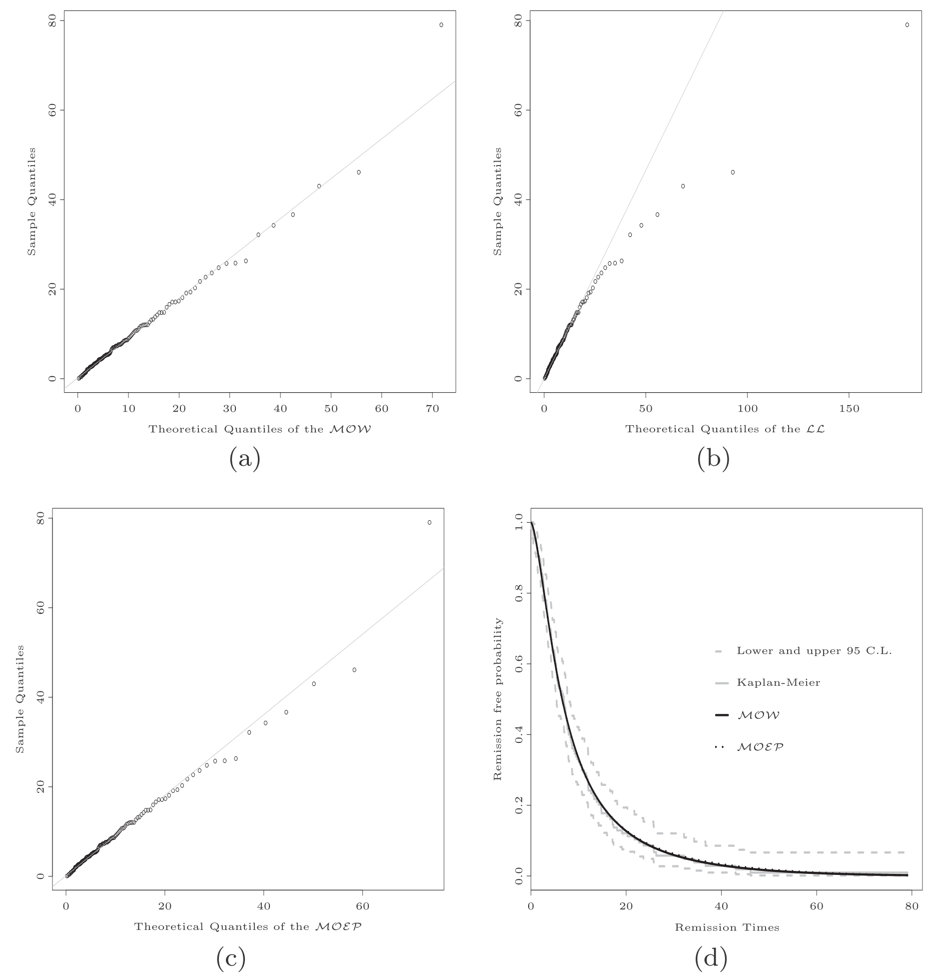


Figure 9 Plots of the estimated. (a) Q-Q plot of $\mathcal{MOW}$, **(b)** $\mathcal{LL}$, **(c)** $\mathcal{MOEP}$ distributions and **(d)** Kaplan-Meier curve estimated survival and upper and lower 95% confidence limits for the cancer patients.

Using the binomial expansion, we have

$$B = \sum_{j=0}^{\infty} (-1)^j \binom{c}{j} \left[1 - \left\{1 - \overline{\delta} \exp\left[-\alpha H(x; \boldsymbol{\xi})\right]\right\}^{-1}\right]^j.$$

Thus,

$$\begin{aligned} F(x;\delta,\alpha,\xi)^c = {} & \mathbb{I}_{(\delta<1)} \sum_{j,k=0}^{\infty} (-1)^k \frac{(c)_j}{j!} \binom{c}{k} \overline{\delta}^j \exp[-(j+k)\alpha H(x;\boldsymbol{\xi})] \\ & + \mathbb{I}_{(\delta=1)} \sum_{k=0}^{\infty} (-1)^k \binom{c}{k} \exp[-k\alpha H(x;\boldsymbol{\xi})] \\ & + \mathbb{I}_{(\delta>1)} \sum_{j,k=0}^{\infty} (-1)^{j+k} \binom{c}{k}\binom{c}{j} \exp[-k\alpha H(x;\boldsymbol{\xi})] \\ & \times [1 - \{1 - \overline{\delta} \exp[-\alpha H(x;\boldsymbol{\xi})]\}^{-1}]^j. \end{aligned}$$

Hence, based on Equation (13), the following expansion holds

$$f(x;\delta,\alpha,\xi)F(x;\delta,\alpha,\xi)^c = \left(\sum_{\nu=0}^{\infty} w_\nu\, g\,(x;(\nu+1)\alpha,\boldsymbol{\xi})\right) F(x;\delta,\alpha,\xi)^c = \mathbb{I}_{(\delta<1)} \sum_{j,k,\nu=0}^{\infty} (-1)^k$$

$$\times\, w_\nu\, \frac{(c)_j}{j!} \binom{c}{k} \overline{\delta}^{\,j}\ \exp[-(j+k)\alpha H(x;\boldsymbol{\xi})]\ g(x;(\nu+1)\alpha,\boldsymbol{\xi})$$

$$+\,\mathbb{I}_{(\delta=1)} \sum_{k,\nu=0}^{\infty} (-1)^k\, w_\nu \binom{c}{k}\ \exp[-k\alpha H(x;\boldsymbol{\xi})]\ g(x;(\nu+1)\alpha,\boldsymbol{\xi})$$

$$+\,\mathbb{I}_{(\delta>1)} \sum_{j,k,\nu=0}^{\infty} (-1)^{j+k}\ w_\nu \binom{c}{k}\binom{c}{j}\ \exp[-k\alpha H(x;\boldsymbol{\xi})]$$

$$\times\ \left[1-\{1-\overline{\delta}\exp[-\alpha H(x;\boldsymbol{\xi})]\}^{-1}\right]^j\ g(x;(\nu+1)\alpha,\boldsymbol{\xi}). \qquad (24)$$

Competing interests
The authors declare that they have no competing interests.

Authors' contributions
The authors MS-N, MB, LMZ, ADCN and GMC proposed a new class of models named the Marshall-Olkin extended Weibull distributions and investigated some of its structural properties including ordinary and incomplete moments, generating and quantile functions, mean deviations, information theory measures and some types of entropies. Two special models were discussed and the estimation of the family model parameters was performed by maximum likelihood. They provided a simulation study and two applications to real data. All authors read and approved the final manuscript.

Acknowledgements
The authors gratefully acknowledge financial support from CAPES and CNPq. The authors are also grateful to three referees and an associate editor for helpful comments and suggestions.

Author details
[1]Departamento de Estatística, Universidade Federal de Campina Grande, Bodocongó, 58429-970, Campina Grande, PB, Brazil. [2]Departamento de Estatística, Universidade Federal do Piauí, Ininga, 64049-550, Teresina, PI, Brazil. [3]Departamento de Estatística, Universidade Federal do Rio Grande do Norte, Campus Universitário Lagoa Nova, 59078-970, Natal, RN, Brazil. [4]Departamento de Estatística, Universidade Federal de Pernambuco, Cidade Universitária, 50740-540, Recife, PE, Brazil.

Received: 28 November 2013 Accepted: 22 April 2014
Published: 16 June 2014

References
Bain, LJ: Analysis for the linear failure-rate life-testing distribution. Technometrics. **16**, 551–559 (1974)
Barreto-Souza, W, Lemonte, AJ, Cordeiro, GM: General results for the Marshall and Olkin's family of distributions. An. Acad. Bras. Cienc. **85**, 3–21 (2013)
Bebbington, M, Lai, CD, Zitikis, R: A flexible Weibull extension. Reliability Eng. Syst. Saf. **92**, 719–726 (2007)
Carrasco, JMF, Ortega, EMM, Cordeiro, GM: A generalized modified Weibull distribution for lifetime modeling. Comput. Stat. Data Anal. **53**, 450–462 (2008)
Chen, Z: A new two-parameter lifetime distribution with bathtub shape or increasing failure rate function. Stat. Probability Lett. **49**, 155–161 (2000)
Chen, G, Balakrishnan, N: A general purpose approximate goodness-of-fit test. J. Qual. Technol. **27**, 154–161 (1995)
Cordeiro, GM, Lemonte, AJ: On the Marshall-Olkin extended Weibull distribution. Stat. Paper. **54**, 333–353 (2013)
Cover, TM, Thomas, JA: Elements of Information Theory. John Wiley & Sons, New York (1991)
Cox, DR, Hinkley, DV: Theoretical Statistics. Chapman and Hall, London (1974)
Doornik, J: Ox 5: object-oriented matrix programming language. 5th ed. Timberlake Consultants, London (2007)
EL-Bassiouny, AH, Abdo, NF: Reliability properties of extended makeham distributions. Comput. Methods Sci. Technol. **15**, 143–149 (2009)
Fisk, PR: The graduation of income distributions. Econometrica. **29**, 171–185 (1961)
Fréchet, M: Sur la loi de probabilité. Ann. Soc. Polon. Math. **6**, 93–93 (1927)
Garvan, F: The Maple Book. Chapman and Hall/CRC, London (2002)
Gompertz, B: On the nature of the function expressive of the law of human mortality and on the new model of determining the value of life contingencies. Philos. Trans. R. Soc. Lond. **115**, 513–585 (1825)
Guess, F, Proschan, F: Mean residual life: Theory and applications. In: Krishnaiah, PR, Rao, CR (eds.) *Handbook of Statistics, vol. 7*, pp. 215–224. Elsevier, (1988). http://dx.doi.org/10.1016/S0169-7161(88)07014-2
Ghitany, ME: Marshall-Olkin extended Pareto distribution and its application. Int. J. Appl. Math. **18**, 17–31 (2005)

Ghitany, ME, Kotz, S: Reliability properties of extended linear failure-rate distributions. Probability Eng. Informational Sci. **21**, 441–450 (2007)
Ghitany, ME, AL-Hussaini, EK, AL-Jarallah: Marshall-Olkin extended Weibull distribution and its application to Censored data. J. Appl. Stat. **32**, 1025–1034 (2005)
Ghitany, ME, AL-Awadhi, FA, Alkhalfan, LA: Marshall-Olkin extended Lomax distribution and its applications to censored data. Comm. Stat. Theor. Meth. **36**, 1855–1866 (2007)
Gurvich, M, DiBenedetto, A, Ranade, S: A new statistical distribution for characterizing the random strength of brittle materials. J. Mater. Sci. **32**, 2559–2564 (1997)
Hartley, RVLL: Transmission of information. Bell Syst. Techn. J. **7**, 535–563 (1928)
Jayakumar, K, Mathew, T: On a generalization to Marshall-Olkin scheme and its application to Burr type XII distribution. Stat. Paper. **49**, 421–439 (2008)
Johnson, NL, Kotz, S, Balakrishnan, N: Continuous Univariate Distributions, Vol. 1. Wiley, New York (1994)
Kies, JA: The Strength of Glass, NRL Report 5093. Naval Research Lab., Washington, DC (1958)
Krishna, E, Jose, KK, Ristić, M: Applications of Marshal-Olkin Fréchet distribution. Comm. Stat. Simulat. Comput. **42**, 76–89 (2013)
Kullback, S, Leibler, RA: On information and sufficiency. Ann. Math. Stat. **22**, 79–86 (1951)
Lai, CD, Xie, M, Murthy, DNP: A modified Weibull distribution. Trans. Reliab. **52**, 33–37 (2003)
Lawless, JF: Statistical Models and Methods for Lifetime Data. 2nd ed. Wiley, New York (2003)
Lee, ET, Wang, JW: Statistical Methods for Survival Data Analysis. 3rd ed. Wiley, New York (2003)
Lomax, KS: Business failures; another example of the analysis of failure data. J. Am. Stat. Assoc. **49**, 847–852 (1954)
Marshall, A, Olkin, I: A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families. Biometrika. **84**, 641–652 (1997)
McCool, JI: Using the Weibull Distribution: Reliability, Modeling and Inference. John Wiley & Sons, New Jersey (2012)
Nadarajah, S, Kotz, S: On some recent modifications of Weibull distribution. IEEE Trans. Reliab. **54**, 561–562 (2005)
Nascimento, ADC, Cintra, RJ, Frery, AC: Hypothesis testing in speckled data with stochastic distances. IEEE Trans. Geosci. Remote Sensing. **48**, 373–385 (2010)
Nascimento, ADC, Horta, MM, Frery, AC, Cintra, RJ: Comparing edge detection methods based on stochastic entropies and distances for PolSAR imagery. IEEE J. Selected Topics Appl. Earth Observations Remote Sensing. **7**, 648–663 (2014)
Nikulin, M, Haghighi, F: A chi-squared test for the generalized power Weibull family for the head-and-neck cancer censored data. J. Math. Sci. **133**, 1333–1341 (2006)
Owen, WJ: A new three-parameter extension to the Birnbaum-Saunders distribution. IEEE Trans. Reliab. **55**, 475–479 (2006)
Pham, H: A vtub-shaped hazard rate function with applications to system safety. Int. J. Reliab. Appl. **3**, 1–16 (2002)
Phani, KK: A new modified Weibull distribution function. Commun. Am. Ceramic Soc. **70**, 182–184 (1987)
R Development Core Team: R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna (2009)
Rayleigh, JWS: On the resultant of a large number of vibrations of the same pitch and of arbitrary phase. Phil. Mag. **10**, 73–78 (1880)
Ristić, MM, Jose, KK, Ancy, J: A Marshall-Olkin gamma distribution and minification process. STARS: Stress Anxiety Res. Soc. **11**, 107–117 (2007)
Rodriguez, N: A guide to the Burr type XII distributions. Biometrika. **64**, 129–134 (1977)
Sankaran, PG, Jayakumar, K: On proportional odds model. Stat. Paper. **49**, 779–789 (2008)
Seghouane, A-K, Amari, S-I: The AIC criterion and symmetrizing the Kullback–Leibler divergence. IEEE Trans. Neural Netw. **18**, 97–106 (2007)
Shannon, CE: A mathematical theory of communication. Bell Syst. Techn. J. **27**, 379–423 (1948)
Sigmon, K, Davis, TA: MATLAB Primer. 6th ed. Chapman and Hall/CRC, London (2002)
Smith, RM, Bain, LJ: An exponential power life testing distribution. Comm. Stat. Theor. Meth. **4**, 469–481 (1975)
Wolfram, S: The Mathematica Book. 5th ed. Wolfram Media, Cambridge (2003)
Xie, M, Lai, D: Reliability analysis using additive Weibull model with bathtub-shaped failure rate function. Reliab. Eng. Syst. Saf. **52**, 87–93 (1995)
Xie, M, Tang, Y, Goh, TN: A modified Weibull extension with bathtub-shaped failure rate function. Reliab. Eng. Syst. Saf. **76**, 279–285 (2002)
Zhang, T, Xie, M: Failure data analysis with extended Weibull distribution. Comm. Stat. Simulat. Comput. **36**, 579–592 (2007)

doi:10.1186/2195-5832-1-9
Cite this article as: Santos-Neto *et al.*: **The Marshall-Olkin extended Weibull family of distributions.** *Journal of Statistical Distributions and Applications* 2014 **1**:9.