ATLANTIS PRESS Journal of Statistical Theory and Applications, Vol. 16, No. 4 (December 2017) 427–447

The Marshall-Olkin-Kumaraswamy-*G* family of distributions

Laba Handique

Department of Statistics, Dibrugarh University, Dibrugarh-786004, India
handiquelaba@gmail.com

Subrata Chakraborty [*]

Department of Statistics, Dibrugarh University, Dibrugarh-786004, India
subrata_stats@dibru.ac.in

G.G. Hamedani

Department of MSCS, Marquette University, Milwaukee, WI 53201-1881, USA
g.hamedani@mu.edu

Received 5 October 2016

Accepted 6 March 2017

Abstract

A new family of continuous probability distributions is proposed by using Kumaraswamy-*G* distribution as the base line distribution in the Marshall-Olkin construction. A number of known distributions are derived as particular cases. Various properties of the proposed family like formulation of the pdf as different mixture of exponentiated baseline distributions, order statistics, moments, moment generating function, Rényi entropy, quantile function and random sample generation have been investigated. Asymptotes, shapes and stochastic ordering are also investigated. Characterizations of the proposed family based on truncated moments, hazard function and reverse hazard function are also presented. The parameter estimation by method of maximum likelihood, their large sample standard errors and confidence intervals and method of moment are also discussed. Two members of the proposed family are compared with different sub models and also with the corresponding members of Kumaraswamy-Marshall-Olkin-*G* family by fitting of two real life data sets.

Keywords: Exponentiated family; Power weighted moment; Characterizations; AIC; BIC; CAIC; HQIC and K-S test.

1. Introduction

One of the preferred area of research in the filed of the probability distribution is that of generating new distributions starting with a base line distribution by adding one or more additional parameters. Notable among them are the Azzalini's skewed family (Azzalini, 1985), Marshall-Olkin extended (*MOE*) family (Marshall and Olkin, 1997), exponentiated family of distributions (Gupta *et al*., 1998), or the composite methods of combining two or more known competing distributions through transformations like beta-generated family (Eugene *et al*., 2002; Jones 2004), gamma-generated familiy (Zografos and Balakrishnan 2009, Ristic and Balakrishnan, 2012), Kumaraswamy-*G* (*Kw-G*) family (Cordeiro and de Castro, 2011), McDonald-*G* family (Alexander *et al*., 2012), beta extended-*G* family (Cordeiro *et al*., 2012), Kumaraswamy-beta generalized family (Pescim *et al*., 2012), exponentiated transformed transformer family (Alzaghal *et al.* 2013), exponentiated generalized family (Cordeiro *et al*., 2013), geometric exponential-Poisson family

[*] Corresponding author.

(Nadarajah *et al*., 2013a), truncated-exponential skew symmetric family (Nadarajah *et al*., 2013b), logistic-generated family (Torabi and Montazari, 2014), Kumaraswamy Marshall-Olkin-*G* family (Alizadeh *et al*., 2015a), generalized odd log-logistic family (Cordeiro *et al*., 2016), generalized gamma-Weibull distribution (Meshkat *et al.* 2016) and generalized transmuted-*G* family (Nofal *et al*., 2017). While the additional parameter(s) bring in more flexibility at the same time they also complicates the mathematical form of the resulting distribution, often considerably enough to render it not amenable to further analytical and numerical manipulations. But with the advent of sophisticated powerful mathematical and statistical softwares unlike in past now a days more and more complex distributions are getting accepted as viable models of data analysis. Tahir and Nadarajah (2015) provided a detail review of how new families of univariate continuous distributions can be generated through introduction of additional parameter(s).

1.1 *Marshall-Olkin Extended (MOE) family of distributions*

Starting with a given baseline distribution with pdf $f(t)$, cdf $F(t)$, Marshall and Olkin (1997) proposed a new flexible semi parametric family of distributions and defined a new survival function (sf) $\overline{F}^{MO}(t)$ by introducing an additional parameter $\alpha > 0$. The sf $\overline{F}^{MO}(t)$ of the *MOE* family of distributions is defined by

$$\overline{F}^{MO}(t) = \alpha\,\overline{F}(t)\big/\{1-\overline{\alpha}\,\overline{F}(t)\}\,. \tag{1}$$

where $-\infty < t < \infty, \alpha > 0$ and $\overline{\alpha} = 1-\alpha$. The parameter $'\alpha'$ is known as the tilt parameter since the hrf of the new family is shifted below (above) the hrf of the base line distribution for $\alpha \geq 1 (0 < \alpha \leq 1)$ (Nanda and Das, 2012). That is for all $t \geq 0$, $h^{MO}(t) \leq h(t)$ when $\alpha \geq 1$, and $h^{MO}(t) \geq h(t)$ when $0 < \alpha \leq 1$, where $h^{MO}(t)$ and $h(t)$ are the hrf's of the *MOE* and baseline distributions respectively. Consequently,

$$F^{MO}(t) = F(t)\big/\{1-\overline{\alpha}\,\overline{F}(t)\}\,, \tag{2}$$

and

$$f^{MO}(t) = \alpha\, f(t)\big/[1-\overline{\alpha}\,\overline{F}(t)]^2\,. \tag{3}$$

If $\alpha = 1$, then we have $\overline{F}^{MO}(t) = \overline{F}(t)$. Other reliability measures like the hrf, rhrf and chrf associated with (1) are given by

$$h^{MO}(t) = h(t)\big/\{1-\overline{\alpha}\,\overline{F}(t)\}, \quad r^{MO}(t) = \alpha\, r(t)\big/\{1-\overline{\alpha}\,\overline{F}(t)\}$$

and

$$H^{MO}(t) = -\log\,[\alpha\,\overline{F}(t)\big/\{1-\overline{\alpha}\,\overline{F}(t)\}]$$

respectively. Where $h(t)$ and $r(t)$ is the hrf and rhrf of the baseline distribution.

It is obvious that many new families can be derived from Marshall-Olkin set up by considering different base line distribution *F* in the equation (1). These new families are usually termed as Marshall-Olkin extended *F* distribution. For detail see Tahir and Nadarajah (2015).

Jayakumar and Mathew (2008) generalized the Marshall-Olkin set up by exponentiating the Marshall-Olkin sf as

$$\overline{F}^{GMO}(t) = [\,\alpha\,\overline{F}(t)\big/\{1-\overline{\alpha}\,\overline{F}(t)\}\,]^{\theta}\,.$$

This method is called method of Lehmann alternative 1 due to Lehmann (1953).

Tahir and Nadarajah (2015) proposed another generalization through Lehmann alternative 2 due to Lehmann (1953) by exponentiating the Marshall-Olkin cdf as

$$F^{GMO}(t) = 1-[\{\alpha\,\overline{F}(t)\}/\{1-\overline{\alpha}\,\overline{F}(t)\}]^{\theta}\,.$$

[for more on Marshall-Olkin family see Barakat *et al.* (2009), Jose *et al.* (2011), Barreto-Souza *et al*. (2013), Cordeiro *et al*. (2014), Alizadeh *et al*., (2015a)].

1.2 *Kumaraswamy-G* (*Kw-G*) *family of distributions*

For a baseline cdf $G(t)$ with pdf $g(t)$, Cordeiro and de Castro (2011) defined *Kw-G* distribution with cdf and pdf

$$F^{KwG}(t) = 1 - [1 - G(t)^a]^b \tag{4}$$

and

$$f^{KwG}(t) = a\,b\,g(t)G(t)^{a-1}[1 - G(t)^a]^{b-1}, \tag{5}$$

where $t > 0,\ g(t) = G'(t)$ and $a > 0, b > 0$ are shape parameters in addition to those in the baseline distribution which partly govern skewness and variation in tail weights. For a lifetime random variable *T*, the sf, hrf, rhrf and chrf for distribution in (4) are given respectively by

$$\overline{F}^{KwG}(t) = [1 - G(t)^a]^b,$$

$$h^{KwG}(t) = a\,b\,g(t)G(t)^{a-1}[1 - G(t)^a]^{-1},$$

$$r^{KwG}(t) = a\,b\,g(t)G(t)^{a-1}[1 - G(t)^a]^{b-1}\left\{1 - [1 - G(t)^a]^b\right\}^{-1}$$

and

$$H^{KwG}(t) = -b\log[1 - G(t)^a].$$

Recently, Alizadeh *et al*., (2015a) proposed the Kumaraswamy Marshall-Olkin family of distributions by using the Marshall-Olkin cdf in that Kw-G family and studied its many properties.

The main motivation behind the present article is to propose a new family of continuous probability distributions that generalizes the Kw-*G* family as well as the Marshall-Olkin extended family by integrating the former as the base line distribution in the latter. We call this new family the Marshall-Olkin Kumaraswamy-*G* (*MOKw-G*) family of distributions which encompasses many known families of distributions and study some of its general properties, characterizations and parameter estimation. We also carry out two real life data modeling applications.

The rest of this article is organized in six more sections. In section 2 the new family is defined along with its physical basis and list of some important sub models. In section 3 we discuss few important general results of the proposed family. In the next section some characterization results are presented while different methods of estimation of parameters are presented in section 5. In section 6 we present two examples of comparative data fitting. The paper ends with concluding remarks in the final section.

2. *MOKw-G* family of distributions

We now propose a new extension of the *MO* family by considering the cdf and pdf of *Kw-G* distribution in (4) and (5) as the $f(t)$ and $F(t)$ respectively in the *MO* formulation in (3) and call it *MOKw-G* distribution with pdf given by

$$f^{MOKwG}(t) = \alpha\,a\,b\,g(t)G(t)^{a-1}[1 - G(t)^a]^{b-1}\big/[1 - \overline{\alpha}[1 - G(t)^a]^b]^2, -\infty < t < \infty, \alpha > 0, a > 0, b > 0. \tag{6}$$

Similarly using equation (4) in (2) the cdf, sf, hrf, rhrf and chrf of *MOKw-G* are respectively obtained as

$$F^{MOKwG}(t) = 1 - [1 - G(t)^a]^b \big/ 1 - \overline{\alpha}[1 - G(t)^a]^b, \tag{7}$$

$$\overline{F}^{MOKwG}(t) = \alpha[1 - G(t)^a]^b \big/ 1 - \overline{\alpha}[1 - G(t)^a]^b, \tag{8}$$

$$h^{MOKwG}(t) = \{a\,b\,g(t)G(t)^{a-1}[1 - G(t)^a]^{-1}\} \big/ \{1 - \overline{\alpha}[1 - G(t)^a]^b\}, \tag{9}$$

$$r^{MOKwG}(t) = \{\alpha\,a\,b\,g(t)G(t)^{a-1}[1 - G(t)^a]^{b-1}[1 - [1 - G(t)^a]^b]^{-1}\} \big/ \{1 - \overline{\alpha}[1 - G(t)^a]^b\},$$

and

$$H^{MOKwG}(t) = -\log\left\{\alpha[1-G(t)^a]^b \big/ 1-\overline{\alpha}[1-G(t)^a]^b\right\}.$$

The pdf in equation (6) for $\alpha = 1$, reduces to that of *Kw-G* in equation (5) and for $a = b = 1$, reduces to that of *MO* in equation (3).

2.1 *Some special members*

In this section we provide some special cases of the *MOKw-G* family of distributions and list their main distributional characteristics.

2.1.1 *The MOKw- exponential (MOKw-E) distribution*

Let the base line distribution be exponential with parameter $\lambda > 0$, $g(t) = \lambda e^{-\lambda t}$ and $G(t) = 1 - e^{-\lambda t}$, $t > 0$, then for the *MOKw-E* model we get the pdf, cdf and hrf respectively as

$$f^{MOKwE}(t) = \alpha\, a\, b\, \lambda\, e^{-\lambda t}(1-e^{-\lambda t})^{a-1}[1-(1-e^{-\lambda t})^a]^{b-1} \big/ [1-\overline{\alpha}[1-(1-e^{-\lambda t})^a]^b]^2,$$

$$F^{MOKwE}(t) = 1-[1-(1-e^{-\lambda t})^a]^b \big/ 1-\overline{\alpha}[1-(1-e^{-\lambda t})^a]^b,$$

$$h^{MOKwE}(t) = [a\, b\, \lambda\, e^{-\lambda t}(1-e^{-\lambda t})^{a-1}[1-(1-e^{-\lambda t})^a]^{-1}] \big/ [1-\overline{\alpha}[1-(1-e^{-\lambda t})^a]^b].$$

2.1.2 *The MOKw-Weibull (MOKw-W) distribution*

Taking the Weibull distribution (Ghitany *et al*., 2005, Zhang and Xie, 2007, and Caroni, 2010) with parameters $\lambda > 0$ and $\beta > 0$ having pdf and cdf $g(t) = \lambda\beta t^{\beta-1} e^{-\lambda t^\beta}$ and $G(t) = 1 - e^{-\lambda t^\beta}$, $t > 0$ respectively we get the pdf, cdf and hrf of *MOKw-W* distribution respectively as

$$f^{MOKwW}(t) = \alpha\, a\, b\, \lambda\, \beta t^{\beta-1} e^{-\lambda t^\beta}[1-e^{-\lambda t^\beta}]^{a-1}[1-[1-e^{-\lambda t^\beta}]^a]^{b-1} \big/ [1-\overline{\alpha}[1-[1-e^{-\lambda t^\beta}]^a]^b]^2,$$

$$F^{MOKwW}(t) = \{1-[1-[1-e^{-\lambda t^\beta}]^a]^b\} / \{1-\overline{\alpha}[1-[1-e^{-\lambda t^\beta}]^a]^b\},$$

$$h^{MOKwW}(t) = \{a\, b\, \lambda\, \beta t^{\beta-1} e^{-\lambda t^\beta}[1-e^{-\lambda t^\beta}]^{a-1}[1-[1-e^{-\lambda t^\beta}]^a]^{-1}\} / \{1-\overline{\alpha}[1-[1-e^{-\lambda t^\beta}]^a]^b\}.$$

2.1.3 *The MOKw-Lomax (MOKw-L) distribution*

Considering the Lomax distribution (Ghitany *et al*., 2007) with pdf and cdf given by $g(t) = (\beta/\delta)[1+(t/\delta)]^{-(\beta+1)}$ and $G(t) = 1-[1+(t/\delta)]^{-\beta}$, $t > 0$, $\beta > 0$ and $\delta > 0$ the pdf, cdf and hrf of *MOKw-L* distribution are respectively given by

$$f^{MOKwL}(t) = \frac{\alpha\, a\, b(\beta/\delta)[1+(t/\delta)]^{-(\beta+1)}[1-[1+(t/\delta)]^{-\beta}]^{a-1}[1-[1-[1+(t/\delta)]^{-\beta}]^a]^{b-1}}{[1-\overline{\alpha}[1-[1-[1+(t/\delta)]^{-\beta}]^a]^b]^2},$$

$$F^{MOKwL}(t) = [1-[1-\{1-[1+(t/\delta)]^{-\beta}\}^a]^b]/[1-\overline{\alpha}[1-\{1-[1+(t/\delta)]^{-\beta}\}^a]^b],$$

$$h^{MOKwL}(t) = \frac{a\, b(\beta/\delta)[1+(t/\delta)]^{-(\beta+1)}[1-[1+(t/\delta)]^{-\beta}]^{a-1}[1-\{1-[1+(t/\delta)]^{-\beta}\}^a]^{-1}}{1-\overline{\alpha}[1-\{1-[1+(t/\delta)]^{-\beta}\}^a]^b}.$$

2.2 *Genesis of the distribution*

Let T_i $(i = 1,2,...,N)$ be a sequence of *i.i.d.* random variables with survival function $[1-G(t)^a]^b$, then

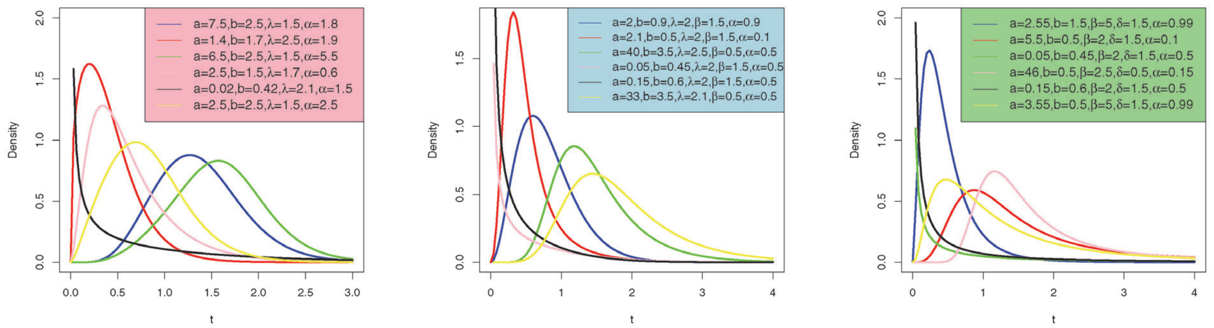
Journal of Statistical Theory and Applications, Vol. 16, No. 4 (December 2017) 427–447

i. if N has a geometric distribution with parameter $\alpha\,(0<\alpha\le 1)$ independent of T_i's, then $\min(T_1, T_2 ..., T_N)$ is distributed as $MOKw-G(\alpha, a, b)$ or

ii. if N has a geometric distribution with parameter $1/\alpha\,(\alpha>1)$ independent of T_i's, then $\max(T_1, T_2 ..., T_N)$ is distributed as $MOKw-G(\alpha^{-1}, a, b)$.

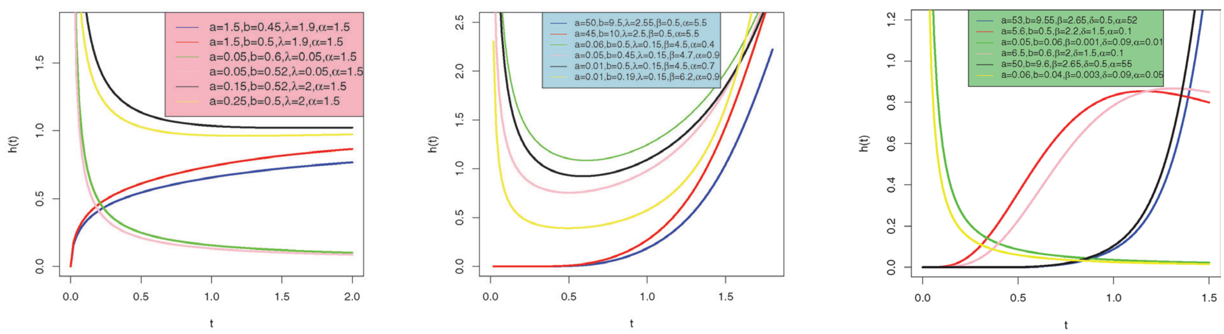
2.3 *Shape of the density and hazard functions*

Here we have plotted the pdf and hrf of the *MOKw-G* by choosing exponential, Weibull and Lomax distributions for *G* with given parameter values to study the variety of shapes assumed by the family.



(a) (b) (c)

Fig. 1. Density plots of (a) *MOKw-E*, (b) *MOKw-W* and (c) *MOKw-L* distributions.



(a) (b) (c)

Fig. 2. Hazard plots of (a) *MOKw-E*, (b) *MOKw-W* and (c) *MOKw-L* distributions.

From the plots in figures 1 and 2, it can be seen that the family is very flexible and can offer many different types of shapes of density and hazard rate functions including the bath tub shaped for hazard.

3. General results

In this section we derive certain general results for the proposed *MOKw-G* family following the methods described in Barreto-Souza *et al*. (2013), Cordeiro *et al*. (2014) and Alizadeh *et al*. (2015a).

3.1 *Series expansions*

Consider the series representation

$$(1-z)^{-k} = \sum_{j=0}^{\infty} \frac{\Gamma(k+j)}{\Gamma(k)\,j!} z^j \,. \tag{10}$$

This is valid for $|z|<1$ and $k>0$, where $\Gamma(.)$ is the gamma function. If $\alpha \in (0,1)$ and using equation (10) in equation (6), we obtain

$$f^{MOKwG}(t) = f^{KwG}(t;a,b) \sum_{j=0}^{\infty} \kappa_j [\overline{F}^{KwG}(t;a,b)]^j \tag{11}$$

$$= -\sum_{j=0}^{\infty} \kappa_j^{/} \frac{d}{dt} [\overline{F}^{KwG}(t;a,b)]^{j+1} = \sum_{j=0}^{\infty} \kappa_j^{/} f^{KwG}(t;a,b(j+1)) \,,$$

where $\kappa_j^{/} = \kappa_j^{/}(\alpha) = \alpha(1-\alpha)^j$; $\kappa_j = \kappa_j(\alpha) = (j+1)\kappa_j^{/}$ and $\overline{F}^{KwG}(t) = [1-G(t)^a]^b$ is the sf of $Kw-G$ distribution. Again,

$$f^{MOKwG}(t) = f^{KwG}(t;a,b) \sum_{j=0}^{\infty} \sum_{k=0}^{j} \varphi_{j,k} [F^{KwG}(t;a,b)]^{j-k} \tag{12}$$

$$= \sum_{j=0}^{\infty} \sum_{k=0}^{j} \varphi_{j,k}^{/} \frac{d}{dt} [F^{KwG}(t;a,b)]^{j-k+1} \,,$$

Where

$$\varphi_{j,k}^{/} = \varphi_{j,k}^{/}(\alpha) = \alpha(1-\alpha)^j (-1)^{j-k} \binom{j+1}{k}, \quad \varphi_{j,k} = \varphi_{j,k}(\alpha) = (j-k+1)\,\varphi_{j,k}^{/}$$

and $F^{KwG}(t) = 1-[1-G(t)^a]^b$ is the cdf of *Kw-G* distribution in equation (4). Similar expansion for the survival function of *MOKw-G* [for $\alpha \in (0,1)$] can be derived as

$$\overline{F}^{MOKwG}(t) = \alpha\, \overline{F}^{KwG}(t;a,b) \sum_{j=0}^{\infty} (1-\alpha)^j \{\overline{F}^{KwG}(t;a,b)\}^j = \sum_{j=0}^{\infty} \kappa_j^{/}\, \overline{F}^{KwG}(t;a,b(j+1)) \,,$$

The density function (6) can also be expressed as

$$f^{MOKwG}(t) = \{a\,b\,g(t)G(t)^{a-1}[1-G(t)^a]^{b-1}\} \big/ \{\alpha\,[1-\{(\alpha-1)[1-[1-G(t)^a]^b]\}/\alpha]^2\} \,.$$

Hence for $\alpha > 1$, using equation (10), we get

$$f^{MOKwG}(t) = f^{KwG}(t;a,b) \sum_{j=0}^{\infty} \eta_j \{F^{KwG}(t;a,b)\}^j \tag{13}$$

$$= \sum_{j=0}^{\infty} \eta_j^{/} \frac{d}{dt} \{F^{KwG}(t;a,b)\}^{j+1} \,, \tag{14}$$

where $\eta_j^{/} = \eta_j^{/}(\alpha) = (\alpha-1)^j \big/ \alpha^{j+1}$ and $\eta_j = \eta_j(\alpha) = (j+1)\eta_j^{/}$.

Similarly for $\alpha > 1$, the survival function of *MOKw-G* can be expressed as

$$\overline{F}^{MOKwG}(t) = [1-G(t)^a]^b \big/ [1-\{(\alpha-1)[1-[1-G(t)^a]^b]\}/\alpha] \,.$$

On using equation (10) we get

$$\overline{F}^{MOKwG}(t) = \overline{F}^{KwG}(t;a,b) \sum_{j=0}^{\infty} C_j' [F^{KwG}(t;a,b)]^j$$

where $C_j' = (\alpha/j+1)\,\eta_j$.

3.2 *Order statistics*

Suppose $T_1, T_2, ..., T_n$ is a random sample from *MOKw-G* distribution. Let $T_{i:n}$ denote the i^{th} order statistic. The pdf of $T_{i:n}$ can be expressed as

$$f_{i:n}(t) = \{n!/(i-1)!\,(n-i)!\}\, f^{MOKwG}(t)[1-\overline{F}^{MOKwG}(t)]^{i-1}\,\overline{F}^{MOKwG}(t)^{n-i}$$
$$= \{n!/(i-1)!\,(n-i)!\}\, f^{MOKwG}(t)\sum_{p=0}^{i-1}(-1)^p\binom{i-1}{p}\overline{F}^{MOKwG}(t)^{n+p-i}\,.$$

Now using the general expansion of the pdf and sf of *MOKw-G* distribution from section 3.1, we get the pdf of the i^{th} order statistic for of the *MOKw-G* for $\alpha \in (0,1)$ as

$$f_{i:n}(t) = f^{KwG}(t;a,b)\sum_{j,q=0}^{\infty} X_{j,q}\,[\overline{F}^{KwG}(t;a,b)]^{j+q+n+p-i} \tag{15}$$
$$= \sum_{j,q=0}^{\infty} X'_{j,q} f^{KwG}(t;a,b(j+q+n+p-i+1)),$$

where

$$X_{j,q} = n\kappa_j\,\kappa'_q\binom{n-1}{i-1}\sum_{p=0}^{i-1}\binom{i-1}{p}(-1)^{p+1}\;, X'_{j,q} = X_{j,q}/(j+q+n+p-i+1),$$

κ_j and κ'_q are defined earlier.

Again using the general expansion of the pdf and sf of *MOKw-G* distribution from section 3.1, we get the pdf of the i^{th} order statistic for of the *MOKw-G* for $\alpha > 1$ as

$$f_{i:n}(t) = f^{KwG}(t;a,b)\;[\overline{F}^{KwG}(t;a,b)]^{(n+p-i)}\sum_{j,k=0}^{\infty}\lambda_{j,k}\{F^{KwG}(t;a,b)\}^{j+k}, \tag{16}$$

where

$$\lambda_{j,k} = n\,\eta_j\,d_{n+p-i,k}\binom{n-1}{i-1}\sum_{p=0}^{i-1}(-1)^p\binom{i-1}{p}\quad,$$

$$d_{n+p-i,k} = \frac{1}{k\,C'_0}\sum_{h=1}^{k}[h(n+p-i-1)-k]C'_h\,d_{n+p-i,k-h}\,,\eta_j \quad \text{and} \quad C'_k \text{ defined in section 3.1}$$

Remark 1. Equations (11)-(16) reveal that the density functions of the *MOKw-G* distribution and that of its order statistics can be expressed as a product of the baseline density $f(t)$ with an infinite power series of $G(.)$ and also as a mixture of exponentiated-*G* distributions under Lehman alternatives. These results play important role and may be used to obtain explicit expressions for the moments and moment generating function (mgf) of the *MOKw-G* distribution and of its order statistics in a general framework and also for special models using the corresponding results of exponentiated-*G* distributions.

3.3 *Moments*

The probability weighted moments (PWMs), first proposed by Greenwood *et al.* (1979), are expectations of certain functions of a random variable whose mean exists. The $(p,q,r)^{th}$ PWM of *T* is having a base line cdf $F(t)$ defined by

$$\Gamma_{p,q,r}=\int_{-\infty}^{\infty} t^p [F(t)]^q [1-F(t)]^r f(t)\,dt\,.$$

The s^{th} moment of T from equations (11), (12) for $\alpha\in(0,1)$ and from equation (13) for $\alpha>1$, can be written as

$$E(T^s)=\int_{-\infty}^{+\infty} t^s\, a\,b\,g(t)G(t)^{a-1}[1-G(t)^a]^{b-1}\sum_{j=0}^{\infty}\kappa_j[1-G(t)^a]^{bj}dt=\sum_{j=0}^{\infty}\kappa_j\,\Gamma_{s,\,0,\,j}\,,$$

$$E(T^s)=\sum_{j=0}^{\infty}\sum_{k=0}^{j}\varphi_{j,k}\;\Gamma_{s,\,j-k,\,0}$$

and as

$$E(T^s)=\sum_{j=0}^{\infty}\eta_j\,\Gamma_{s,\,j,\,0}\,,$$

respectively, where

$$\Gamma_{p,q,r}=\int_{-\infty}^{\infty} t^p\;\{1-[1-G(t)^a]^b\}^q\{[1-G(t)^a]^b\}^r[a\,b\,g(t)\;G(t)^{a-1}[1-G(t)^a]^{b-1}]dt$$

is the PWM of *Kw-G* the for the baseline distribution *G*.

The PWM can generally be used for estimating parameters quantiles of generalized distributions. These moments have low variances and do not possess severe biases and they compare favourably with estimators obtained by maximum likelihood (Alizadeh *et al*., 2015b).

Proceeding as above we can also derive s^{th} moment of the i^{th} order statistic $T_{i:n}$, in a random sample of size *n* from *MOKw-G* for $\alpha\in(0,1)$ and $\alpha>1$, on using equations (15) and (16), as

$$E(T^s{}_{i,n})=\sum_{j,q=0}^{\infty}X_{j,q}\,\Gamma_{s,0,\;j+q+n+p-i}\;\text{ and }\;E(T^s{}_{i,n})=\sum_{j,k=0}^{\infty}\lambda_{j,k}\,\Gamma_{s,\;j+k,\;n+p-i}$$

respectively, where the quantities $\varphi_{j,k},\kappa_j,\,\eta_j,\,X_{j,q}$ and $\lambda_{j,k}$ are defined in sections 3.1 and 3.2.

3.4 *Moment generating function*

The mgf of *MOKw-G* family can be easily expressed in terms of those of the exponentiated *Kw-G* distribution using the results of section 3.1. For example, using equation (14), it can be seen that

$$M_T(s)=E[e^{st}]=\int_{-\infty}^{\infty}e^{st}f(t)\,dt=\int_{-\infty}^{\infty}e^{st}\sum_{j=0}^{\infty}\eta_j^{/}\frac{d}{dt}\left\{F^{KwG}(t;a,b)\right\}^{j+1}dt$$

$$=\sum_{j=0}^{\infty}\eta_j^{/}\int_{-\infty}^{\infty}e^{st}\frac{d}{dt}\left\{F^{KwG}(t;a,b)\right\}^{j+1}dt=\sum_{j=0}^{\infty}\eta_j M_X(s)\,,$$

where η_j is define in section 3.1 and *X* follows exponentiated *Kw-G* distribution.

3.5 *Entropy*

In this subsection we present results on Renyi entropy and relative entropy.

3.5.1 *Rényi entropy*

The entropy of a random variable is a measure of uncertainty. The Rényi entropy is defined as

$$I_R(\delta) = (1-\delta)^{-1} \log\left(\int_{-\infty}^{\infty} f(t)^{\delta}\, dt \right),$$

where $\delta > 0$ and $\delta \neq 1$. For furthers details, see Song (2001). Using expansion in equation (10) in equation (6) we can write for $\alpha \in (0,1)$

$$f(t)^{\delta} = \frac{\alpha^{\delta}[\,a\,b\,g(t)\,G(t)^{a-1}[1-G(t)^{a}]^{b-1}\,]^{\,\delta}}{\Gamma(2\delta)} \sum_{j=0}^{\infty} (1-\alpha)^{j}\,\Gamma(2\delta+j) \frac{[\{1-G(t)^{a}\}^{b}]^{j}}{j!}.$$

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

$$I_R(\delta) = (1-\delta)^{-1} \log\left(\sum_{j=0}^{\infty} q_j \int_{-\infty}^{\infty} [\,a\,b\,g(t)\,G(t)^{\,a-1}[1-G(t)^{\,a}]^{b-1}\,]^{\,\delta} [\{1-G(t)^{a}\}^{\,b}\,]^{\,j}\, dt \right),$$

where

$$q_j = q_j(\alpha) = \{\alpha^{\,\delta}(1-\alpha)^{\,j}\,\Gamma(2\delta+j)\} / (\Gamma(2\delta)\, j!)\ ,$$

while for $\alpha > 1$, we can write

$$f(t)^{\delta} = \frac{[\,a\,b\,g(t)\,G(t)^{\,a-1}[1-G(t)^{\,a}\,]^{\,b-1}]^{\delta}}{\alpha^{\delta+j}\ \Gamma(2\delta)} \sum_{j=0}^{\infty} (\alpha-1)^{\,j}\,\Gamma(2\delta+j)\ \frac{\{1-[1-G(t)^{\,a}]^{b}\}^{\,j}}{j!}.$$

Hence for $\alpha > 1$, the Rényi entropy of T can be obtained as

$$I_R(\delta) = (1-\delta)^{-1} \log\left(\sum_{j=0}^{\infty} r_j \int_{-\infty}^{\infty} [\,a\,b\,g(t)\,G(t)^{\,a-1}[1-G(t)^{\,a}\,]^{\,b-1}\,]^{\,\delta} [1-[1-G(t)^{\,a}\,]^{\,b}\,]^{\,j}\, dt \right),$$

where

$$r_j = r_j(\alpha) = \{(\alpha-1)^{\,j}\,\Gamma(2\delta+j)\} / (\alpha^{\delta+j}\,\Gamma(2\delta)\, j!).$$

3.5.2 *Relative entropy*

The relative entropy (R.E.) between two distributions with pdfs $f_1(t)$ and $f_2(t)$ is define as

$$R.E\,(f_1, f_2) = E_{f_1}\big(\log\{f_1(t)/f_2(t)\}\big) = \int \big(\log\{f_1(t)/f_2(t)\}\big)\, f_1(t)\ dt.$$

This is also known as Kullback-Leibler divergence (distance) (Kullback, 1959). It is a measure of distance between $f_1(t)$ and $f_2(t)$ and can be seen as inefficiency of assuming f_2 for modeling when f_1 is the true distribution. We have presented the estimated $R.E.(MOKw-G,\ f_2)$ where f_2 is a competing distribution for all the data fitting examples in Tables 1 and 2 of section 6.

3.6 *Quantile function and random sample generation*

The p^{th} Quantile t_p for *MOKw-G* can be easily obtained by solving the equation $F^{MOKwG}(t) = p$ as $t_p = G^{-1}[1-[1-\{\alpha\, p/(1-\overline{\alpha}\, p)\}]^{\,1/b}\,]^{1/a}$. In general to generate a random number $'t'$ from *MOKw-G* from a uniform random number 'u' we can use the formula

$$t = G^{-1}[1-[1-\{\alpha\, u/(1-\overline{\alpha}\, u)\}]^{\,1/b}\,]^{1/a}.$$

For example, let G be exponential distribution with parameter $\lambda > 0$. Then, the p^{th} quantile is obtained as $-(1/\lambda)\log[1-p]$. Therefore, the p^{th} quantile t_p, of *MOKw-E* is given by

$$t_p = (-1/\lambda)\ \log[1-[1-[1-\{\alpha\, p/(1-\overline{\alpha}\, p)\}]^{\,1/b}\,]^{1/a}\,]\ .$$

3.7 *Asymptotes and shapes*

Here we investigate the asymptotic shapes of the *MOKw-G* family.

Proposition 1. *The asymptotes of equations* (6), (7) *and* (9) *as* $t \to 0$ *are given by*

$$f(t) \sim \{a\,b\,g(t)\,G(t)^{\,a-1}\}/\alpha \qquad as\ t \to 0$$

$$F(t) \sim 0 \qquad as\ t \to 0$$

$$h(t) \sim \{a\,b\,g(t)\,G(t)^{\,a-1}\}/\alpha \qquad as\ t \to 0.$$

Proposition 2. *The asymptotes of equations* (6), (7) *and* (9) *as* $t \to \infty$ *are given by*

$$f(t) \sim \alpha a\,b\,g(t)[1-G(t)^a]^{\,b-1} \quad as\ t \to \infty$$

$$F(t) \sim 1-[1-G(t)^a]^{\,b} \qquad as\ t \to \infty$$

$$h(t) \sim a\,b\,g(t)[1-G(t)^a]^{-1} \qquad as\ t \to \infty.$$

The shapes of the density and hazard rate functions of the *MOKw-G* can be described analytically by looking at their critical points which are the roots of the equations

$$(d/d\,t)\log[f(t)] = g'(t)/g(t) + (a-1)\{g(t)/G(t)\} + \{a\,(1-b)\,g(t)G(t)^{a-1}\}/1-G(t)^a$$
$$-2\ \{\overline{\alpha}\,a\,b\,g(t)\,G(t)^{\,a-1}\,[1-G(t)^{\,a}\,]^{\,b-1}\}/1-\overline{\alpha}\,[1-G(t)^a]^{\,b} = 0 \tag{17}$$

and

$$\frac{d}{d\,t}\log[h(t)] = \frac{g'(t)}{g(t)} + (a-1)\frac{g(t)}{G(t)} + \frac{a\,g(t)\,G(t)^{a-1}}{1-G(t)^a} - \frac{\overline{\alpha}\,a\,b\,g(t)\,G(t)^{a-1}\,[1-G(t)^{\,a}\,]^{\,b-1}}{1-\overline{\alpha}\,[1-G(t)^a]^{\,b}} = 0. \tag{18}$$

The number of roots of (17) may be more than one. In particular, if $t = t_0$ is a root of (17) then it corresponds to a local maximum, a local minimum or a point of inflexion depending on whether $\psi(t_0) < 0, \psi(t_0) > 0$ *or* $\psi(t_0) = 0$ and similarly for (18) $\omega(t_0) < 0,\ \omega(t_0) > 0$ *or* $\omega(t_0) = 0$ where $\psi(t) = (d^2/dt^2)\log[f(t)]$ and $\omega(t) = (d^2/dt^2)\log[h(t)]$.

3.8 *Stochastic orderings*

In this section we study the stochastic ordering of the *MOKw-G* distribution. Stochastic ordering properties have applications in diverse fields such as economics, reliability, survival analysis, insurance, finance, actuarial and management sciences (Shaked and Shanthikumar, 2007).
Let X and Y be two random variables with cfd's F and G, respectively, survival functions $\overline{F} = 1-F$ and $\overline{G} = 1-G$, and corresponding pdf's f and g. Then X is said to be smaller than Y in the likelihood ratio order ($X \le_{lr} Y$) if $f(t)/g(t)$ is decreasing in $t \ge 0$; stochastic order ($X \le_{st} Y$) if $\overline{F}(t) \le \overline{G}(t)$ for all $t \ge 0$; hazard rate

order ($X \le_{hr} Y$) if $\overline{F}(t)/\overline{G}(t)$ is decreasing in $t \ge 0$; reversed hazard rate order ($X \le_{rhr} Y$) if $F(t)/G(t)$ is decreasing in $t \ge 0$. These four stochastic orders are related to each other, as

$$X \le_{rhr} Y \Leftarrow X \le_{lr} Y \Rightarrow X \le_{hr} Y \Rightarrow X \le_{st} Y$$

Theorem 1. *Let $X \sim MOKwG(\alpha_1, a, b)$ and $Y \sim MOKwG(\alpha_2, a, b)$. If $\alpha_1 < \alpha_2$, then $X \le_{lr} Y$. As a consequence it follows that ($X \le_{hr} Y$, $X \le_{rhr} Y$, $X \le_{st} Y$).*

Proof. Let $X \sim MOKw\text{-}G(\alpha_1, a, b)$ and $Y \sim MOKw\text{-}G(\alpha_2, a, b)$. If $\alpha_1 < \alpha_2$, then

$$\frac{f(t)}{g(t)} = \frac{\alpha_1\, a\, b\, g(t)\, G(t)^{a-1}[1-G(t)^a]^{b-1} / [1-\overline{\alpha}_1[1-G(t)^a]^b]^2}{\alpha_2\, a\, b\, g(t)\, G(t)^{a-1}[1-G(t)^a]^{b-1} / [1-\overline{\alpha}_2[1-G(t)^a]^b]^2}$$

$$= (\alpha_1/\alpha_2)[\{1-\overline{\alpha}_2[1-G(t)^a]^b\} / \{1-\overline{\alpha}_1[1-G(t)^a]^b\}]^2.$$

Now,

$$d/dt(f(t)/g(t)) = [2\alpha_1(\alpha_1-\alpha_2)\{1-\overline{\alpha}_2[1-G(t)^a]^b\}\, a\, b[1-G(t)^a]^{b-1}\, G(t)^{a-1}\, g(t)] / [\alpha_2\{1-\overline{\alpha}_1[1-G(t)^a]^b\}^3].$$

This is always less than 0 since $\alpha_1 < \alpha_2$. Hence, $f(t)/g(t)$ is decreasing in t. That is $X \le_{lr} Y$.

4. Characterizations

Characterization of distributions is an important research area which has recently attracted the attention of many researchers. This section deals with various characterizations of *MOKw-G* of distribution. These characterizations are based on (i) a simple relationship between two truncated moments; (ii) the hazard function; and (iii) It should be mentioned that for characterization (i) the cdf is not required to have a closed form.

4.1 *Characterization based on two truncated moments*

Here we present a characterization of *MOKw-G* distribution in terms of a simple relationship between two truncated moments. The first characterization result employs a theorem due to Glanzel (1987), stated below. Note that the result holds also when the interval *H* is not closed. Moreover, as mentioned above, it could be also applied when the cdf *F* does not have a closed form. As shown in Glanzel (1990), this characterization is stable in the sense of weak convergence.

Theorem 2 (By Glanzel, 1987). *Let (Ω, F, P) be a given probability space and let $H = [d, e]$ be an interval for some $d < e$ ($d = -\infty$, $e = \infty$ might as well be allowed). Let $T : \Omega \to H$ be a continuous random variable with the distribution function F and let q_1 and q_2 be two real function defined on H such that*

$$E[q_2(T) / T \ge t] = E[q_1(T) / T \ge t]\, \xi(t), \qquad t \in H,$$

is defined with some real function ξ. Assume that $q_1, q_2 \in C^1(H)$, $\xi \in C^2(H)$ and F is twice continuously differentiable and strictly monotone function on the set H. Finally, assume that the equation $\xi q_1 = q_2$ has no real solution in the interior of H. Then F is uniquely determined by the functions q_1, q_2 and ξ particularly

$$F(t) = \int_a^t C \left| \xi'(u) / [\xi(u)\, q_1(u) - q_2(u)] \right| \exp(-s(u))\, du\ ,$$

where the function s *is a solution of the differential equation* $s' = \xi' q_1 / (\xi q_1 - q_2)$ *and C is the normalization constant, such that* $\int_H dF = 1$.

Proposition 3. *Let* $T : \Omega \to \Re$ *be a continuous random variable and let*

$$q_1(t) = \{1 - \bar{\alpha}[1 - G(t)^a]^b\}^2 \text{ and } q_2(t) = q_1(t)[1 - G(t)^a]$$

for $t \in \Re$. *The random variable T has pdf in equation* (6) *if and only if the function* ξ *defined in Theorem* 2 *has the form*

$$\xi(t) = (b/(b+1))[1 - G(t)^a], \qquad t \in \Re.$$

Proof. Let T be a random variable with pdf in equation (6), then

$$(1 - F^{MOKwG}(t))E[q_1(T) / T \ge t] = \alpha[1 - G(t)^a], \qquad t \in \Re,$$

and

$$(1 - F^{MOKwG}(t))E[q_2(T) / T \ge t] = (\alpha b/(b+1))[1 - G(t)^a]^{b+1}, \qquad t \in \Re,$$

and finally

$$\xi(t) q_1(t) - q_2(t) = -q_1(t)\, (1/(b+1))[1 - G(t)^a] < 0 \quad \text{for} \quad t \in \Re.$$

Conversely, if ξ is given as above, then

$$s'(t) = \xi'(t) q_1(t) / (\xi(t) q_1(t) - q_2(t)) = [a b g(t) G(t)^{a-1}] / [1 - G(t)^a], \; t \in \Re,$$

and hence

$$s(t) = -b \log[1 - G(t)^a], \qquad t \in \Re.$$

Now, in view of Theorem 2, T has the density in equation (6).

Corollary 1. *Let* $T : \Omega \to \Re$ *be a continuous random variable and let* $q_1(t)$ *be as in Proposition* 3. *The pdf of T is given by equation* (6) *if and only if there exist functions* Q_2 *and* ξ *defined in Theorem* 2 *satisfying the differential equation*

$$\xi'(t) q_1(t) / \{\xi(t) q_1(t) - q_2(t)\} = a b g(t) G(t)^{a-1} / \{1 - G(t)^a\}, \qquad t \in \Re.$$

The general solution of the differential equation in Corollary 1 is

$$\xi(t) = [1 - G(t)^a]^{-1} [-\int a b g(t) G(t)^{a-1} (q_1(t))^{-1} q_2(t) + D],$$

where D is a constant. Note that a set of function satisfying the above differential equation is given in Proposition 3 with $D = 0$. However, it should be also noted that there are other triplets (q_1, q_2, ξ) satisfying the conditions of Theorem 2.

4.2 *Characterization based on hazard function*

It is known that the hazard function, h_F, of a twice differentiable distribution function, *F,* satisfies the first order differential equation

$$f'(t)/f(t) = [h'_F(t)/h_F(t)] - h_F(t).$$

For many univariate continuous distributions, this is the only characterization available in terms of the hazard function. The following non-trivial characterizations of *MOKw-G* distribution in terms of the hazard function, which is not of the above trivial form.

Proposition 4. *Let $T:\Omega \to \Re$ be a continuous random variable. The pdf of T is given by equation (6) if and only if its hazard function $h^{MOKwG}(t)$ satisfies differential equation*

$$h^{MOKwG}(t) - \{g'(t)/g(t)\}h^{MOKwG}(t)$$
$$= \frac{a\,b\,g(t)^2[(a-1)+G(t)^a - \bar{\alpha}\{(a-1)+(1+a\,b)G(t)^a\}]\{1-G(t)^a\}^b}{\{1-G(t)^a\}^2[1-\bar{\alpha}\{1-G(t)^a\}^b]}, \qquad t \in \Re,$$

with the initial condition $\lim_{t\to-\infty} h^{MOKwG}(t) = 0$ for $a > 1$.

Proof. If T has pdf in equation (6), then clearly the above differential equation holds. Now, if the differential equation holds, then

$$\frac{d}{dt}[(g(t))^{-1}h^{MOKwG}(t)] = a\,b\frac{d}{dt}\left\{\frac{G(t)^{a-1}[1-G(t)^a]^{-1}}{1-\bar{\alpha}\{1-G(t)^a\}^b}\right\},$$

or

$$h^{MOKwG}(t) = [a\,b\,g(t)\,G(t)^{a-1}[1-G(t)^a]^{-1}]/[1-\bar{\alpha}\{1-G(t)^a\}^b],$$

which is the hazard function of the *MOKw-G* distribution.

4.3 *Characterization in terms of the reverse (or reversed) hazard function*

The reverse hazard function, r_F, of a twice differentiable distribution function, *F*, is defined as

$$r_F = f(t)/F(t), \quad t \in \text{support of } F.$$

Proposition 5. *Let $T:\Omega \to \Re$ be a continuous random variable. The pdf of T is given by equation (6) if and only if its reverse hazard function $r^{MOKwG}(t)$ satisfies differential equation*

$$r^{MOKwG}(t) - \{g'(t)/g(t)\}r^{MOKwG}(t) = \alpha\,a\,b\,g(t)\frac{d}{dt}\left\{\frac{G(t)^{a-1}[1-G(t)^a]^{-1}}{[1-\{1-G(t)^a\}^b][1-\bar{\alpha}\{1-G(t)^a\}^b]}\right\}, \quad t\in\Re.$$

Proof. The proof is similar to that of Proposition 4.

5. Estimation

5.1 *Maximum likelihood estimation*

The model parameters of the *MOKw-G* distribution can be estimated by maximum likelihood. Let $t = (t_1, t_2, ..., t_n)'$ be a random sample of size n from *MOKw-G* with parameter vector $\boldsymbol{\theta} = (\alpha, a, b, \boldsymbol{\beta}^T)'$, where $\boldsymbol{\beta} = (\beta_1, \beta_2, ..., \beta_q)'$ corresponds to the parameter(s) of the baseline distribution *G*. Then the log-likelihood function is given by

$$\ell = \ell(\boldsymbol{\theta}) = n\log\alpha + n\log(a\,b) + \sum_{i=1}^{n}\log[g(t_i,\boldsymbol{\beta})] + (a-1)\sum_{i=1}^{n}\log[G(t_i,\boldsymbol{\beta})]$$
$$+(b-1)\sum_{i=1}^{n}\log[1-G(t_i,\boldsymbol{\beta})^{a}] - 2\sum_{i=1}^{n}\log[1-\bar{\alpha}[1-G(t_i,\boldsymbol{\beta})^{a}]^{b}].$$

This log-likelihood function can not be solved analytically because of its complex form but it can be maximized numerically by employing global optimization methods available with softwares like R, Mathematica.
By taking the partial derivatives of the log-likelihood function with respect to α, a, b and $\boldsymbol{\beta}$ components of the score vector $U_{\boldsymbol{\theta}} = (U_\alpha, U_a, U_b, U_{\beta^T})^T$ can be obtained as follows:

$$U_\alpha = \frac{\partial\ell}{\partial\alpha} = \frac{n}{\alpha} - 2\sum_{i=1}^{n}\frac{[1-G(t_i,\boldsymbol{\beta})^{a}]^{b}}{1-\bar{\alpha}\,[1-G(t_i,\boldsymbol{\beta})^{a}]^{b}},$$

$$U_a = \frac{\partial}{\partial a}\ell = \frac{n}{a} + \sum_{i=1}^{n}\log[G(t_i,\boldsymbol{\beta})] + (1-b)\sum_{i=1}^{n}\frac{G(t_i,\boldsymbol{\beta})^{a}\log[G(t_i,\boldsymbol{\beta})]}{1-G(t_i,\boldsymbol{\beta})^{a}}$$
$$-2\sum_{i=1}^{n}\frac{b\bar{\alpha}[1-G(t_i,\boldsymbol{\beta})^{a}]^{b-1}G(t_i,\boldsymbol{\beta})^{a}\log[G(t_i,\boldsymbol{\beta})]}{1-\bar{\alpha}\,[1-G(t_i,\boldsymbol{\beta})^{a}]^{b}},$$

$$U_b = \frac{\partial}{\partial b}\ell = \frac{n}{b} + \sum_{i=1}^{n}\log[1-G(t_i,\boldsymbol{\beta})^{a}] + 2\sum_{i=1}^{n}\frac{\bar{\alpha}[1-G(t_i,\boldsymbol{\beta})^{a}]^{b}\log[1-G(t_i,\boldsymbol{\beta})^{a}]}{1-\bar{\alpha}[1-G(t_i,\boldsymbol{\beta})^{a}]^{b}} \quad \text{and}$$

$$U_\beta = \frac{\partial}{\partial\boldsymbol{\beta}}\ell = \sum_{i=1}^{n}\frac{g^{(\boldsymbol{\beta})}(t_i,\boldsymbol{\beta})}{g(t_i,\boldsymbol{\beta})} + (a-1)\sum_{i=1}^{n}\frac{G^{(\boldsymbol{\beta})}(t_i,\boldsymbol{\beta})}{G(t_i,\boldsymbol{\beta})} + (1-b)\sum_{i=1}^{n}\frac{a\,G(t_i,\boldsymbol{\beta})^{a-1}G^{(\boldsymbol{\beta})}(t_i,\boldsymbol{\beta})}{1-G(t_i,\boldsymbol{\beta})^{a}}$$
$$-2\sum_{i=1}^{n}\frac{b\bar{\alpha}[1-G(t_i,\boldsymbol{\beta})^{a}]^{b-1}a\,G(t_i,\boldsymbol{\beta})^{a-1}G^{(\boldsymbol{\beta})}(t_i,\boldsymbol{\beta})}{1-\bar{\alpha}\,[1-G(t_i,\boldsymbol{\beta})^{a}]^{b}}.$$

Simultaneous solution of the equations $U_\theta = (U_\alpha, U_a, U_b, U_{\beta^T})' = 0$ gives the maximum likelihood estimate (MLE) $\hat{\theta} = (\hat{\alpha}, \hat{a}, \hat{b}, \hat{\boldsymbol{\beta}}^T)'$ of $\theta = (\alpha, a, b, \boldsymbol{\beta}^T)'$.

5.1.1 *Asymptotic standard error and confidence interval for the MLEs*

The asymptotic variance-covariance matrix of the MLEs of parameters can obtained by inverting the Fisher information matrix $\mathrm{I}(\boldsymbol{\theta})$ which can be derived using the second partial derivatives of the log-likelihood function with respect to each parameter. The $i\,j^{th}$ elements of $\mathrm{I}_n(\boldsymbol{\theta})$ are given by

$$\mathrm{I}_{ij} = -E[\partial^2 l(\boldsymbol{\theta}) / \partial\theta_i\,\partial\theta_j], \qquad i, j = 1, 2, \ldots, 3+q.$$

The exact evaluation of the above expectations may be cumbersome. In practice one can estimate $\mathrm{I}_n(\boldsymbol{\theta})$ by the observed Fisher's information matrix $\hat{\mathrm{I}}_n(\hat{\boldsymbol{\theta}}) = (\hat{\mathrm{I}}_{ij})$ is defined as:

$$\hat{\mathrm{I}}_{ij} \approx \left(-\partial^2 l(\boldsymbol{\theta}) / \partial\theta_i\,\partial\theta_j\right)_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}, \qquad i, j = 1, 2, \ldots, 3+q.$$

Using the general theory of MLEs under some regularity conditions on the parameters as $n \to \infty$ the asymptotic distribution of $\sqrt{n}\,(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})$ is $N_k(0, V_n)$ where $V_n = (v_{jj}) = \mathrm{I}_n^{-1}(\boldsymbol{\theta})$. The asymptotic behaviour remains valid if V_n is replaced by $\hat{V}_n = \hat{\mathrm{I}}^{-1}(\hat{\boldsymbol{\theta}})$. This result can also be used to provide large sample standard errors and also construct confidence intervals for the model parameters. Thus an approximate standard error and $(1-\gamma/2)100\%$ confidence interval for the MLE of j^{th} parameter θ_j are respectively given by $\sqrt{\hat{v}_{jj}}$ and $\hat{\theta}_j \pm Z_{\gamma/2}\sqrt{\hat{v}_{jj}}$, where $Z_{\gamma/2}$ is the $\gamma/2$ point of standard normal distribution.

5.2 *Estimation by method of moments*

Here an alternative method to estimation of the parameters is discussed. Since the moments are not in closed form, the estimation by the usual method of moments may not be tractable. Therefore following Barreto-Souzai *et al*. (2013) we get

$$E[[1-\bar{\alpha}\{1-\bar{\alpha}[1-G(t)^a]^b\}]^v] = \begin{cases} -\alpha\log(\alpha)/1-\alpha, & v=1 \\ \alpha(1-\alpha^{v-1})/\bar{\alpha}\,(v-1), & v=2,3,...,q+1. \end{cases} \tag{19}$$

One can use (19) to give a new method of estimation i.e. for a random sample $t_1, t_2, ..., t_n$ from a population with survival function in equation (8) the model parameters can be estimated from the equations

$$\frac{1}{n}\sum_{i=1}^{n}[[1-\bar{\alpha}\left\{1-\bar{\alpha}[1-G(t_i)^a]^b\right\}]^v] = \begin{cases} -\alpha\log(\alpha)/1-\alpha, & v=1 \\ \alpha(1-\alpha^{v-1})/\bar{\alpha}\,(v-1), & v=2,3,...,q+1, \end{cases}$$

where q is the number of parameter(s) of the base line distribution.

6. Real life applications

In this subsection, we consider fitting two real data sets to show that the distributions from the proposed *MOKw-G* family can provide better model than the corresponding distributions from *KwMO-G* (Alizadeh *et al*., 2015a) by considering the Frechet and the Exponential distribution as our *G*. Further we compare it with its sub models namely *MO-G, Kw-G* to show its superiority. Here the parameters are estimated by numerical maximization of log-likelihood function and their asymptotic standard errors are computed using large sample approach.

In order to compare the distributions, we have considered known criteria like AIC (Akaike Information Criterion), BIC (Bayesian Information Criterion), CAIC (Consistent Akaike Information Criterion) and HQIC (Hannan-Quinn Information Criterion). It may be noted that $AIC = 2k - 2l$; $BIC = k\log(n) - 2l$; $CAIC = AIC + (2k(k+1))/(n-k-1)$; and $HQIC = 2k\log[\log(n)] - 2l$, where k is the number of parameters in the statistical model, n the sample size and l is the maximized value of the log-likelihood function under the considered model. The Kolmogorov-Smirnov (K-S) statistics for goodness of fit is also carried out to compare the fitted models. We have further provided the estimated relative entropy (R.E.) values, which can be interpreted as the cost of encoding *MOKw-G* through *Kw-G, MO-G* and *KwMO-G*.

Example I. Here we work with the following data about 346 nicotine measurements made from several brands of cigarettes in 1998. The data have been collected by the Federal Trade Commission which is an independent agency of the US government, whose main mission is the promotion of consumer protection. [http: //www.ftc.gov/ reports/tobacco or http: // pw1.netcom.com/ rdavis2/ smoke. html.]

In Table 1, the MLEs and standard errors (SEs) (in parentheses) of the parameters from all the fitted distributions along with AIC, BIC, CAIC, HQIC and R.E. values are presented. According to the lowest values of the AIC, BIC, CAIC and HQIC the *MOKw-Fr* (*MOKw-E*) distribution could be chosen as a better model than *Kw-Fr, MO-Fr* and *KwMO-Fr* (*Kw-E, MO-E, KwMO-E*) distribution.

It can be seen that the R. E. values between *MOKw-G* and its competing distributions in table 1 are high except in the case of *KwMO-E*, indicating the fact that the distance between all the competing distributions (except *KwMO-E*) and are high.

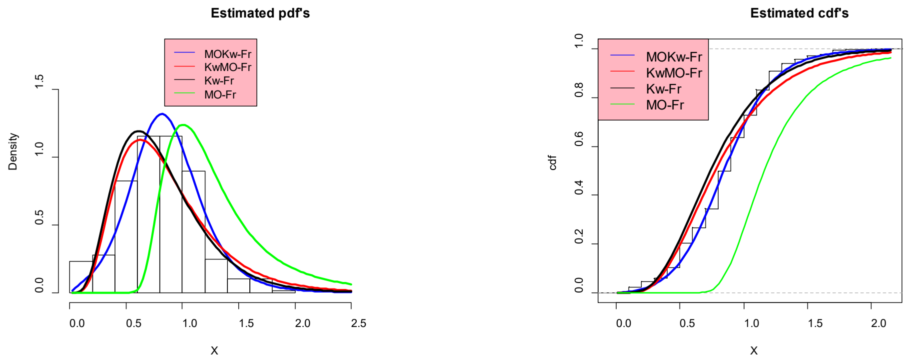
More information is provided for a visual comparison in the form of histogram and Ogive of the observed data with the fitted densities and fitted cdf's displayed in Figure 3 and Figure 4. These plots indicate that the proposed distributions provide the closest fit to this data.

Example II. This data set consists of 100 observations of breaking stress of carbon fibres (in Gba) given by Nichols and Padgett (2006).

In Table 2, the MLEs and standard errors (SEs) (in parentheses) of the parameters from all the fitted distributions along with AIC, BIC, CAIC, HQIC and R.E. values are presented. According to the lowest values of the AIC, BIC, CAIC and HQIC the *MOKw-Fr* (*MOKw-E*) distribution could be chosen as a better model than *Kw-Fr, MO-Fr* and *KwMO-Fr* (*Kw-E, MO-E, KwMO-E*) distribution.

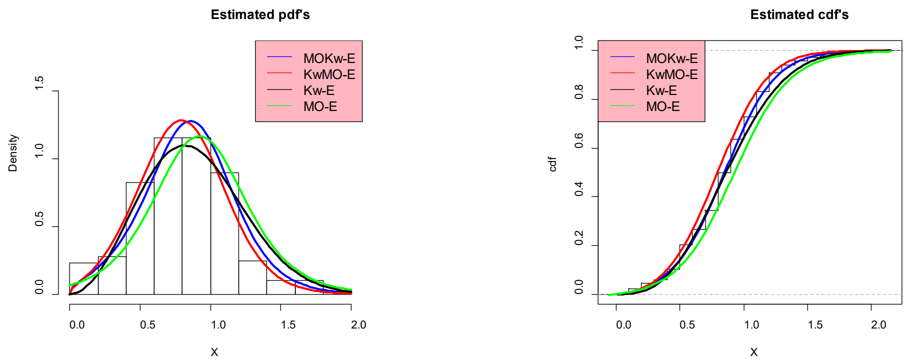
It can be seen that compared to the R.E. values in Table 1, here the values are small indicating the fact that the distance between the competing distributions and *MOKw-G* are small.

Like in the last example here also we have presented the histogram and Ogive of the observed data with the fitted densities and fitted cdf's in Figures 5 and Figures 6 for visual inspection of the closeness of the fittings. These plots indicate that the distributions from the proposed family provide the closest fit to this data.



(a) (b)

Fig. 3. Plots of the (a) observed histogram and estimated pdf's and (b) observed Ogives and estimated cdf's for the *MOKw-Fr*, *Kw-Fr, MO-Fr* and *KwMO-Fr* distributions for example I.



(a) (b)

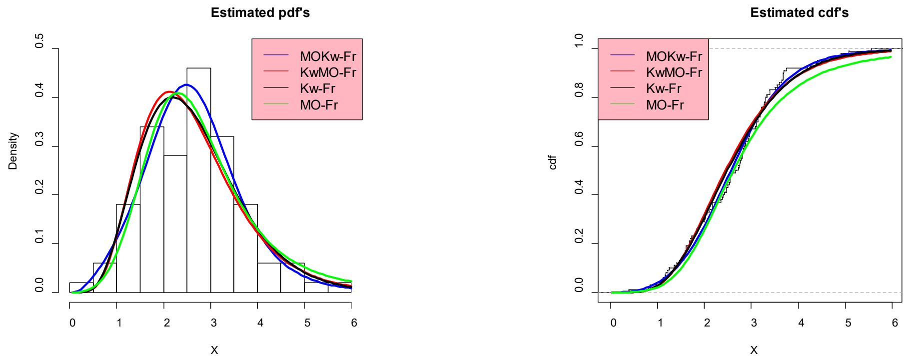
Fig. 4. Plots of the (a) observed histogram and estimated pdf's and (b) observed Ogives and estimated cdf's for the *MOKw-E, Kw-E, MO-E* and *KwMO-E* distributions for example I.

Table 1. MLEs, standard errors (in parentheses) and the AIC, BIC, CAIC, HQIC, K-S with p-values and relative entropy (R.E.) for the nicotine measurements data.

Parameters	$Kw-Fr$	$MO-Fr$	$MOKw-Fr$	$KwMO-Fr$	$Kw-E$	$MO-E$	$MOKw-E$	$KwMO-E$
$\hat{a}$	57.191 (5.806)	---	1.176 (0.497)	10.462 (2.669)	3.020 (0.163)	---	1.511 (0.515)	1.434 (0.236)
$\hat{b}$	81.249 (7.937)	---	126.659 (100.668)	75.085 (19.512)	105.575 (38.348)	---	11.063 (35.759)	1.959 (0.779)
$\hat{\lambda}$	0.497 (0.423)	3.914 (1.121)	0.303 (0.063)	0.251 (0.014)	0.252 (0.045)	4.612 (0.254)	0.586 (1.572)	3.557 (0.677)
$\hat{\delta}$	0.005 (0.059)	0.999 (0.213)	29.261 (23.497)	16.270 (4.834)	---	---	---	---
$\hat{\alpha}$	---	1.521 (0.927)	60.442 (85.106)	0.075 (0.013)	---	70.241 (18.043)	20.97 (19.978)	21.214 (10.377)
log-likelihood	-147.12	-143.10	**-110.27**	-155.49	-115.46	-112.81	**-106.61**	-107.89
AIC	302.24	292.20	**230.54**	320.98	236.92	229.62	**221.22**	223.78
BIC	317.81	303.74	**249.77**	340.21	248.46	237.32	**236.61**	239.17
CAIC	302.35	292.27	**230.72**	321.16	236.99	229.65	**221.34**	223.89
HQIC	308.37	300.33	**238.19**	328.64	241.51	232.68	**227.35**	229.91
K-S	0.277	0.280	0.236	0.274	0.287	0.299	0.282	0.296
p-value	0.086	0.082	**0.202**	0.093	0.068	0.062	**0.078**	0.056
R.E.	47.59	58.07	**---**	52.49	14.29	21.56	**---**	1.27

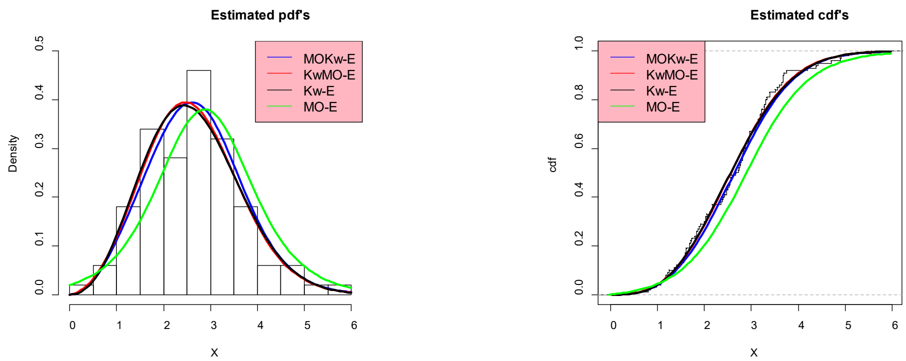
Table 2. MLEs, standard errors (in parentheses) and the AIC, BIC, CAIC, HQIC, K-S with *p*-values and relative entropy (R.E.) for the breaking stress of carbon fibres data

Parameters	$Kw-Fr$	$MO-Fr$	$MOKw-Fr$	$KwMO-Fr$	$Kw-E$	$MO-E$	$MOKw-E$	$KwMO-E$
$\hat{a}$	55.362 (1.429)	---	6.777 (4.936)	9.243 (2.958)	3.439 (0.552)	---	3.226 (1.409)	2.647 (2.066)
$\hat{b}$	140.774 (34.243)	---	38.279 (67.729)	19.203 (14.763)	48.150 (14.213)	---	9.065 (42.865)	4.571 (14.461)
$\hat{\lambda}$	0.518 (0.213)	4.118 (0.342)	0.522 (0.259)	0.953 (0.172)	0.129 (0.103)	1.535 (0.153)	0.295 (0.834)	0.591 (1.103)
$\hat{\delta}$	0.027 (0.008)	0.328 (0.163)	0.416 (0.500)	0.051 (0.026)	---	---	---	---
$\hat{\alpha}$	---	21.243 (9.235)	16.978 (31.897)	17.222 (7.881)	---	75.663 (33.464)	2.566 (5.709)	3.899 (10.290)
log-likelihood	-144.69	-146.28	**-141.63**	-144.83	-145.32	-147.52	**-141.09**	-141.25
AIC	297.38	298.56	**293.26**	299.66	296.64	299.04	**290.18**	290.50
BIC	307.80	306.37	**306.29**	312.69	304.45	304.25	**300.60**	300.92
CAIC	297.80	298.81	**293.89**	300.29	296.89	299.16	**290.60**	290.92
HQIC	301.59	301.72	**298.54**	304.94	299.80	301.06	**294.40**	294.73
K-S	0.089	0.095	0.068	0.102	0.074	0.141	0.064	0.108
p-value	0.516	0.321	**0.825**	0.344	0.732	0.074	**0.869**	0.281
R.E.	1.02	1.24	**---**	1.26	0.63	1.91	**---**	0.14



(a) (b)

Fig. 5. Plots of the (a) observed histogram and estimated pdf's and (b) observed Ogives and estimated cdf's for the *MOKw-Fr*, *Kw-Fr, MO-Fr* and *KwMO-Fr* distributions for example II.



(a) (b)

Fig. 6. Plots of the (a) observed histogram and estimated pdf's and (b) observed Ogives and estimated cdf's for the the *MOKw-E, Kw-E, MO-E* and *KwMO-E* distributions for example II.

7. Conclusion

In this paper, the Marshall-Olkin extended Kumaraswamy generalized distributions is introduced and some of its important distributional and mathematical properties including characterizations are investigated. The maximum likelihood and moment method for estimating the parameters are discussed. Comparative data modeling considering two important members of the proposed *MOKw-G* family with sub models *MO-G, Kw-G* and the corresponding members of *KwMO-G* family (Alizadeh *et al*., 2015a) revealed formers superiority in both examples considered here.

Acknowledgements

The authors would like to thank the three anonymous referees for their valuable suggestions, which lead to substantial improvement of the paper.

References

[1] Alexander, C., Cordeiro, G.M., Ortega, E.M.M. and Sarabia, J.M. (2012). Generalized beta-generated distributions. *Comput Stat Data Anal*, 56 (6), 1880-1897.

[2] Alizadeh, M., Tahir, M.H., Cordeiro, G.M., Zubair, M. and Hamedani, G.G. (2015a). The Kumaraswamy Marshal-Olkin family of distributions. *Journal of the Egyptian Mathematical Society,* 23 (3), 546-557.

[3] Alizadeh, M., Tahir, M.H., Cordeiro, G.M., Zubair, M. and Hamedani, G.G. (2015b). The beta Marshall-Olkin family of distributions. *Journal of Statistical Distributions and Applications,* 2:4 doi:10.1186/s40488-015-0027-7.

[4] Alzaghal, A., Lee, C. and Famoye, F. (2013). Exponentiated T-X family of distributions with some applications. *Int J Statist Probab*, 2 (3), 31-49.

[5] Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scand J Statist*, 12 (2), 171-178.

[6] Barakat, H.M., Ghitany, M.E. and AL-Hussaini, E.K. (2009). Asymptotic distributions of order statistics and record values under the Marshall-Olkin parameterization operation.*Commun Statist Theor Meth*, 38 (13), 2267-2273.

[7] Barreto-Souza, W., Lemonte, A.J. and Cordeiro, G.M. (2013). General results for Marshall and Olkin's family of distributions. *An Acad Bras Cienc*, 85 (1), 3-21.

[8] Cordeiro, G.M., Alizadeh, M., Tahir, M.H. Mansoor, M., Bourguignon, M. and Hamedani G.G. (2017). The generalized odd log logistic family of distributions: properties, regression models and applications, *Journal of Statistical Computation and Simulation*, 87 (5), 908-932.

[9] Cordeiro, G.M. and De Castro, M. (2011). A new family of generalized distributions. *J Stat Comput Simul*, 81 (7), 883-893

[10] Cordeiro, G.M., Lemonte, A.J. and Ortega, E.E.M. (2014). The Marshall-Olkin family of distributions: Mathematical properties and new models. *J Stat Theor Pract*, 8 (2), 343-366.

[11] Cordeiro, G.M., Ortega, E.M.M. and DA Cunha, D.C.C. (2013). The exponentiated generalized class of distributions. *J Data Sci*, 11, 1-27.

[12] Cordeiro, G.M., Silva, G.O. and Ortega, E.M.M. (2012). The beta extended Weibull distribution. *J Probab Stat Sci,* 10, 15-40.

[13] Eugene, N., Lee, C. and Famoye, F. (2002). Beta-normal distribution and its applications. *Commun Statist Theor Meth*, 31 (4), 497-512.

[14] Ghitany, M.E., Al-Hussaini, E.K. and Al-Jarallah, R.A. (2005). Marshall-Olkin extended Weibull distribution and its application to censored data. *J Appl Stat*, 32 (10), 1025-1034.

[15] Ghitany, M.E., Al-Awadhi, F.A. and Alkhalfan, L.A. (2007). Marshall-Olkin extended Lomax distribution and its application to censored data. *Commun Statist Theor Meth*, 36 (10), 1855-1866.

[16] Glanzel, W. (1987). A Characterization theorem based on truncated moments and its application to some distribution families, *Mathematical Statistics and Probability Theory*, Vol. B, Statistical Inference and Methods Proceedings of the 6th Pannonian Symposium on Mathematical Statistics, Bad Tatzmannsdorf, Austria, September 14–20, 1986, 75-84.

[17] Glanzel, W. (1990). Some consequences of a characterization theorem based on truncated moments, *Statistics*: *A Journal of Theoretical and Applied Statistics,* 21 (4), 613-618.

[18] Greenwood, J.A., Landwehr, J.M., Matalas, N.C. and Wallis, J.R. (1979). Probability weighted moments: definition and relation to parameters of several distributions expressible in inverse form. *Water Resour Res,* 15 (5), 1049-1054.

[19] Gupta, R.C., Gupta, P.I. and Gupta, R.D. (1998). Modeling failure time data by Lehmann alternatives. *Commun Statist Theor Meth***,** 27(4), 887-904.

[20] Jayakumar, K. and Mathew, T. (2008). On a generalization to Marshall-Olkin scheme and its application to Burr type XII distribution. *Stat Pap,* 49, 421-439.

[21] Jones, M.C. (2004). Families of distributions arising from the distributions of order statistics. *Test,* 13 (1), 1-43.

[22] Jose, K.K. and Krishna, E. (2011). Marshall-Olkin extended uniform distribution. *Prob Stat Forum***,** 4, 78-88.

[23] Kullback, S. (1959). *Information Theory and Statistics.* John Wiley & Sons, New York.

[24] Lehmann, E.L. (1953). The power of rank tests. *Ann Math Statist*, 24 (1), 23-43.

[25] Marshall, A. and Olkin, I. (1997). A new method for adding a parameter to a family of distributions with applications to the exponential and Weibull families. *Biometrika*, 84 (3), 641-652.

[26] Meshkat, R.S., Torabi, H. and Hamedani, G.G. (2016). A generalized gamma -Weibull distribution: model, properties and applications, *Pakistan Journal of Statistics and Operation Research* 12 (2), 201-212.

[27] Nadarajah, S., Cancho, V.G. and Ortega, E.M.M. (2013a). The geometric exponential Poisson distribution. *Stat Methods Appl*, 22 (3), 355-380.

[28] Nadarajah, S., Jayakumar, K. and Risti, c M.M. (2013b). A new family of lifetime models. *J Stat Comput Simul*, 83 (8), 1389-1404.

[29] Nanda, A.K. and Das, S. (2012). Stochastic orders of the Marshall-Olkin extended distribution. *Statist Probab Lett,* 82 (2), 295-302.

[30] Nichols, M.D. and Padgett, W.J.A. (2006). A bootstrap control chart for Weibull percentiles. *Quality and Reliability Engineering International,* 22 (2), 141-151.

[31] Nofal, Z. M., Afify, A. Z., Yousof, H. M. and Cordeiro, G. M. (2017). The generalized transmuted-G family of distributions. *Commun Statist Theor Meth*, 46 (8), 4119-4136.

[32] Pescim, R.R., Cordeiro, G.M., Demetrio, C.G.B., Ortega, E.M.M. and Nadarajah, S. (2012). The new class of Kummer beta generalized distributions. *SORT,* 36 (2), 153-180.

[33] Ristic, M.M. and Balakrishnan (2012). The gamma-exponentiated exponential distribution. *J Stat Comput Simul*, 82 (8), 1191-1206.

[34] Shaked, M. and Shanthikumar, J.G. (2007). *Stochastic Orders*, Springer, New York.
[35] Song, K.S. (2001). Rényi information, log likelihood and an intrinsic distribution measure. *J Stat Plan Infer,* 93(1-2), 51-69.
[36] Tahir, M.H. and Nadarajah, S. (2015). Parameter induction in continuous univariate distributions: Well-established *G* families. *Anais da Academia Brasileira de Ciências* (*Annals of the Brazilian Academy of Sciences*)*,* 87(2), 539-568.
[37] Torabi, H. and Montazari, N.H. (2014). The logistic-uniform distribution and its application. *Commun Statist Simul Comput*, 43 (10), 2551-2569.
[38] Zhang, T. and Xie, M. (2007). Failure data analysis with extended Weibull distribution. *Commun Statist Simul Comput*, 36 (3), 579–592.
[39] Zografos, K. and Balakrishnan (2009). On families of beta- and generalized gamma generated distributions and associated inference. *Stat Methodol***,** 6 (4), 344-362.