1-1-2017

Some Extended Classes of Distributions: Characterizations and Properties

Gholamhossein G. Hamedani
Marquette University, gholamhoss.hamedani@marquette.edu

Gauss M. Cordeiro
Universidade Federal de Pernambuco

M. C.S. Lima
Universidade Federal de Pernambuco

A. D.C. Nascimento
Universidade Federal de Pernambuco

Some Extended Classes of Distributions: Characterizations and Properties

G. G. Hamedani
Department of Mathematics, Statistics and Computer Science
Marquette University, Milwaukee, USA
gholamhoss.hamedani@marquette.edu

G. M. Cordeiro
Department of Statistics, Federal University of Pernambuco, Recife, Brazil
Gauss@de.ufpe.br

M. C. S. Lima
Department of Statistics, Federal University of Pernambuco, Recife, Brazil
maria@de.ufpe.br

A.D. C. Nascimento
Department of Statistics, Federal University of Pernambuco, Recife, Brazil
abraao@de.ufpe.br

Abstract

Based on a simple relationship between two truncated moments and certain functions of the nth order statistic, we characterize some extended classes of distributions recently proposed in the statistical literature, videlicet Beta-G, Gamma-G, Kumaraswamy-G and McDonald-G. Several properties of these extended classes and some special cases are discussed. We compare these classes in terms of goodness-of-fit criteria using some baseline distributions by means of two real data sets.

1. Introduction

The recent literature has suggested several ways of extending well-known distributions. One of the earliest is the class of distributions generated by a standard beta distribution pionnered by Eugene *et al.* (2002). The more recent ones are: the class of distributions generated by Kumaraswamy (1980)'s distribution defined by Cordeiro and de Castro (2011), and the class of distributions generated by McDonald (1984)'s generalized beta distribution introduced by Alexander *et al.* (2012). Generalized distributions usually provide flexible framework for modeling a wide range of data sets, that is, these models are very useful for fitting a wide spectrum of real world lifetime data in biology, medicine, engineering, economics and other fields.

Alexander *et al.* (2012) proposed the *generalized beta-generated* (GBG) family of distributions (also called McDonald generalized, McG) with the probability density function (pdf) given by

$$f_{McG-K}(x; a, b, c, \lambda) = \frac{c}{B(a,b)} k(x) K(x)^{ac-1} [1 - K(x)^c]^{b-1}, x \geq 0, \qquad (1)$$

and cumulative distribution function (cdf) in the form

$$F_{McG-K}(x; a, b, c, \lambda) = \frac{1}{B(a,b)} \int_0^{K(x)^c} w^{a-1} (1-w)^{b-1} dw, \quad x > 0, \qquad (2)$$

where $a > 0$, $b > 0$ and $c > 0$ are shape parameters, $B(a,b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$ is the beta function, $\Gamma(\cdot)$ is the gamma function and $K(x)$ is a cdf with support in any subinterval of $\mathbb{R}$ and corresponding pdf $k(x) = dK(x)/dx$, which depends on a parameter vector λ. Hereafter, we shall refer to model (2) as the *McDonald generalized-K* (denoted by the prefix "McG-K" for short) family since the McDonald density function is a basic exemplar when $K(x) = x$ for $x \in (0,1)$. The family of distributions (2) includes two important special classes: the *beta generalized* (BG) (Eugene *et al.*, 2002) for $c = 1$, and the *Kumaraswamy generalized* (KwG) (Cordeiro and de Castro, 2011) for $a = 1$. It follows from (2) that the McG-K family with baseline cdf $K(x)$ is the BG distribution with baseline cdf $K(x)^c$. This simple transformation may facilitate the derivation of some of its structural properties.

For example, the pdf and cdf of the *McDonald Normal* (McN) distribution are given by

$$f(x; a, b, c, \mu, \sigma) = \frac{c}{\sigma B(a,b)} \phi\left(\frac{x-\mu}{\sigma}\right) \left[\Phi\left(\frac{x-\mu}{\sigma}\right)\right]^{ac-1} \left\{1 - \left[\Phi\left(\frac{x-\mu}{\sigma}\right)\right]^c\right\}^{b-1}, x \in \mathbb{R},$$

and

$$F(x; a, b, c, \mu, \sigma) = \frac{1}{B(a,b)} \int_0^{\left[\Phi\left(\frac{x-\mu}{\sigma}\right)\right]^c} t^{a-1} (1-t)^{b-1} dt, \quad x \in \mathbb{R}, \qquad (3)$$

respectively, where $a > 0$, $b > 0$, $c > 0$, $\sigma > 0$ and $\mu \in \mathbb{R}$ are parameters and $\phi(x)$ and $\Phi(x)$ are the pdf and cdf of the normal $N(0,1)$ distribution.

For example, the pdf and cdf of the Kumaraswamy-inverse Weibull (Kw-IW) distribution are given by

$$f(x; a, b, \alpha, \beta) = \frac{a \, b \, \alpha \, \beta}{x^{\beta+1}} \exp\left(-\frac{a\alpha}{x^\beta}\right) \left[1 - \exp\left(-\frac{a\alpha}{x^\beta}\right)\right]^{b-1}, \quad x > 0, \qquad (4)$$

and

$$F(x; a, b, \alpha, \beta) = 1 - \left[1 - \exp\left(-\frac{a\alpha}{x^\beta}\right)\right]^b, \quad x \geq 0, \qquad (5)$$

respectively, where $a > 0$, $b > 0$, $\alpha > 0$ and $\beta > 0$ are parameters.

Cordeiro *et al.* (2012) proposed the *beta extended Weibull* (BEW) family of distributions on the basis of the extended class of Weibull distributions studied by Nadarajah and Kotz (2005). The pdf of the BEW family takes the form

$$f_{BEW}(x; a, b, \alpha, \tau) = \frac{\alpha}{B(a,b)} u(x) e^{-\alpha b U(x)} [1 - e^{-\alpha U(x)}]^{a-1}, x > 0 \qquad (6)$$

The corresponding cdf is given by

$$F_{BEW}(x; a, b, \alpha, \tau) = \frac{1}{B(a,b)} \int_0^{1-e^{-\alpha U(x)}} w^{a-1} (1-w)^{b-1} dw, x \geq 0 \qquad (7)$$

where $a > 0$ and $b > 0$ are shape parameters, $\alpha > 0$ is a scale parameter and τ denotes the vector of unknown parameters in $U(x)$. We assume that $U(x)$ is a monotonically

increasing function of x with $U(x) \geq 0$, $\lim_{x \to 0^+} U(x) = 0$ and the derivative $u(x) = dU(x)/dx$ belongs to the interval $(0, \infty)$. A characterization of the BEW family is that its hazard rate function (hrf) can be bathtub shaped, monotonically increasing or decreasing and upside-down bathtub depending basically on the parameter values. This family contains as special models several well-known distributions. Some useful distributions in this family are presented in Cordeiro *et al.* (2012).

The generator proposed by Zografos and Balakrishnan (2009) and Ristic' and Balakrishnan (2012), called the *gamma-G* ("GG" for short) family defined from any baseline cdf $G(x; \ \tau)$, $x \in \mathbb{R}$, considers an extra shape parameter $a > 0$. They defined the GG family by the pdf and cdf

$$f_{GG}(x; \ \tau, \delta) = \frac{g(x; \ \tau)}{\Gamma(\delta)} \{-\log[1 - G(x; \ \tau)]\}^{\delta-1} \tag{8}$$

and $F_{GG}(x; \ \tau, a) = \frac{1}{\Gamma(\delta)} \int_0^{-\log[1-G(x; \ \tau)]} t^{\delta-1} e^{-t} dt = \gamma_1(\delta, -\log[1 - G(x; \ \tau)])$,

respectively, where $g(x; \ \tau) = dG(x; \ \tau)/dx$, $\gamma(\delta, z) = \int_0^z t^{\delta-1} e^{-t} dt$ and $\gamma_1(\delta, z) = \gamma(\delta, z)/\Gamma(\delta)$ are the incomplete gamma function and the incomplete gamma function ratio, respectively. Each new GG distribution can be generated from a specified G distribution.

Nascimento *et al.* (2014) introduced a new class of distributions called the *gamma extended Weibull* (GEW) family based on the work of Zografos and Balakrishnan (2009). The pdf and cdf of this family are defined by

$$f_{GEW}(x; \delta, \alpha, \xi) = \frac{\alpha^\delta}{\Gamma(\delta)} v(x; \xi) V(x; \xi)^{\delta-1} e^{-\alpha V(x;\xi)}, x > 0 \tag{9}$$

and

$$F_{GEW}(x; \delta, \alpha, \xi) = \frac{1}{\Gamma(\delta)} \int_0^{\alpha V(x;\xi)} t^{\delta-1} e^{-t} dt, x \geq 0 \tag{10}$$

respectively, where $\delta > 0$ is a shape parameter, $\alpha > 0$ is a scale parameter and ξ is a vector of unknown parameters in $V(x)$. We assume that $V(x) \geq 0$ is monotonically increasing in x with $\lim_{x \to 0^+} V(x) = 0$, $\lim_{x \to \infty} V(x) = \infty$ and the derivative $v(x) = dV(x)/dx$ is defined in $(0, \infty)$. The proposed family includes several well-known models as special cases such as the exponential, Pareto, Gomertz, Weibull and modified Weibull distributions, among others.

The distribution proposed by Mead (2014), and called the generalized beta extended Pareto (GBEP) distribution, has pdf and cdf given by (for $x > d$)

$$f(x) = f(x; a, b, \lambda, d, k, c) = \frac{c\lambda k d^k x^{-(k+1)}}{B(a,b)} [1 - (d/x)^k]^{\lambda ac-1} \times \{1 -$$

$[1 - (d/x)^k]^{\lambda c}\}^{b-1}$ and $F(x) = F(x; a, b, \lambda, d, k, c) = B(a,b)^{-1} \int_0^{[1-(d/x)^k]^{\lambda c}} w^{a-1}(1 - w)^{b-1} dw$, respectively, where a, b, λ, d, k, c and c are all positive parameters.

Proposition 1.1 *Let* $X: \Omega \to (d, \infty)$ *be a continuous random variable and let*

$$h(x) = \{1 - [1 - (d/x)^k]^{\lambda c}\}^{1-b} \text{ and } g(x) = h(x)[1 - (d/x)^k]^{\lambda ac}$$

for $x \in (d, \infty)$. *The* pdf *of* X *is that of GBEP if and only if the function* η *defined in Theorem 2.1 (of Section 2) has the form* $\eta(x) = \frac{1}{2}\{1 + [1 - (d/x)^k]^{\lambda ac}\}$, $x > d$.

Hashimoto *et al.* (2014) introduced a distribution called the Poisson Birnbaun-Saunders (PBS) model with long-term survivors and pdf and cdf (for $x > 0$) given by

$$f(x) = f(x; \alpha, \lambda, \varphi) = \frac{\varphi \, x^{-3/2} \, (x + \lambda)}{2 \, \alpha \, \sqrt{2\pi\lambda} \, (1 - e^{-\varphi})}$$

$$\times \exp\left\{ -\frac{1}{2\alpha^2}\left(\frac{x}{\lambda} + \frac{\lambda}{x} - 2\right) - \varphi \, \Phi\left[\frac{1}{\alpha}\left(\sqrt{\frac{x}{\lambda}} - \sqrt{\frac{\lambda}{x}}\right)\right]\right\},$$

and

$$F(x) = F(x; \alpha, \lambda, \varphi) = 1 - 2 \, \alpha \, \sqrt{2\pi \, \lambda} \left\{\exp\left(-\varphi \, \Phi\left[\frac{1}{\alpha}\left(\sqrt{\frac{x}{\lambda}} - \sqrt{\frac{\lambda}{x}}\right)\right]\right) - e^{-\varphi}\right\},$$

respectively, where α, λ and φ are all positive parameters and Φ is the standard normal cdf.

Proposition 1.2 *Let $X: \Omega \to (0, \infty)$ be a continuous random variable and let*

$$h(x) = x^{-1/2}(x - \lambda)\exp\left\{\varphi \, \Phi\left[\frac{1}{\alpha}\left(\sqrt{\frac{x}{\lambda}} - \sqrt{\frac{\lambda}{x}}\right)\right]\right\} \text{ and } g(x) = h(x)\left(1 - \frac{\lambda}{x}\right)(x^2 + \lambda^3)$$

for $x \in (0, \infty)$. The pdf of X is that of PBS if and only if the function η defined in Theorem 2.1 has the form $\eta(x) = 2\alpha^2\lambda + x + \lambda^2 x^{-1}, \quad x > 0$.

The goal of this paper is to provide characterizations of the McG-K, BEW, GEW, McN and Kw-IW families described above. These characterizations are based on: (i) a simple relationship between two truncated moments, (ii) certain functions of the nth order statistic, (iii) certain functions of the first order statistic. It is widely known that the problem of characterizing a distribution is an important issue, which has attracted the attention of many researchers. Thus, various characterizations have been established in many different directions. For example, we can refer to Galambos and Kotz (1978), Glänzel (1987), Hamedani (1993, 2002, 2006), Glänzel and Hamedani (2001), Bairamov *et al.* (2005), Ahsanullah and Hamedani (2007), Tavangar and Asadi (2007), Beg and Ahsanullah (2007), Bieniek (2007), Baratpour *et al.* (2008), Nevzotov *et al.* (2003), Su *et al.* (2008), Ahmadi and Fashandi (2009), Haque *et al.* (2009), Akhundov and Nevzorov (2010), Khan *et al.* (2010), Hamedani and Ahsanullah (2011), Yanev and Ahsanullah (2012), among others.

Although in many applications an increase in the number of parameters provides a more suitable model, in characterization problems a lower number of parameters (without seriously affecting the suitability of the model) is mathematically more appealing (see Glänzel and Hamedani, 2001). In the applications, where the underlying distribution is assumed to be McG-K, BEW, GEW, McN or Kw-IW distribution, the investigator needs to verify that the underlying distribution is in fact the McG-K or BEW or GEW or McN or Kw-IW distribution. To this end, the investigator has to rely on the characterizations of these distributions and determine if the corresponding conditions are satisfied. Thus, the problem of characterizing these families of distributions become essential. As mentioned before, our objective is to present characterizations of the McG-K, BEW, GEW, McN and Kw-IW families.

These classes of distributions provide tools to obtain new parametric distributions from existing ones and have applications in many fields, in particular in lifetime modeling.

The paper is organized as follows. In Section 2, we consider a characterization based on two truncated moments. In Section 3, we discuss about characterizations based on truncated moment of the nth order statistic. In Section 4, we provide characterizations based on truncated moment of the first order statistic. In Section 5, we derive expansions for the pdfs of those families as linear combinations of exponentiated - G (Exp-G) families, where G is the baseline model. Some mathematical properties are addressed (Section 6) and two applications are explored to prove the efficiency of the new generators (Section 7). Some concluding remarks are provided in Section 8.

2. Characterization based on two truncated moments

In this section, we present characterizations of the McG-K, BEW, GEW, McN and Kw-IW families in terms of a simple relationship between two truncated moments. The characterizations derived here employ an interesting result due to Glänzel (1987), which is given by the following theorem.

Theorem 2.1 *Let $(\Omega, \mathcal{F}, \boldsymbol{P})$ be a given probability space and let $H = [a, b]$ be an interval for some $a < b$ ($a = -\infty, b = \infty$ mightaswellbeallowed). Let $X : \Omega \to H$ be a continuous random variable with distribution function F and let g and h be two real functions defined on H such that*
$$\mathbf{E}[g(X)|X \geq x] = \mathbf{E}[h(X)|X \geq x]\eta(x), \quad x \in H,$$
is defined for some real function η. Consider that $g, h \in C^1(H)$, $\eta \in C^2(H)$ and F are twice continuously differentiable and strictly monotone function on the set H. Finally, assume that the equation $h\eta = g$ has no real solution in the interior of H. Then, F is uniquely determined by the functions g, h and η, particularly
$$F(x) = \int_a^x C \left| \frac{\eta'(u)}{\eta(u)h(u) - g(u)} \right| \exp[-s(u)] \, du,$$

where the function s is a solution of the differential equation $s' = \frac{\eta' h}{\eta h - g}$ and C is a constant chosen to make $\int_H dF = 1$.

Remarks 2.1 (a) *In Theorem 2.1, the interval H need not be closed.*

(b) The goal is to have the function η as simple as possible.

Proposition 2.1 *Let $X : \Omega \to (0, \infty)$ be a continuous random variable and let $h(x) = K(x)^{c\,(1-a)}$ and $g(x) = K(x)^{c\,(1-a)}[1 - K(x)^c]$ for $x \in (0, \infty)$. The pdf of X is (1) if and only if the function η defined in Theorem 2.1 has the form*
$$\eta(x) = \frac{b}{b+1}[1 - K(x)^c], \quad x > 0.$$

Proof. Let X have pdf (1). Then, for $x > 0$,
$$[1 - F(x)]\mathbf{E}[h(X)|X \geq x] = \frac{1}{b\,B(a, b)}[1 - K(x)^c]^{\,b}$$

and

$$[1 - F(x)]\mathbf{E}[g(X)|X \geq x] = \frac{1}{(b+1)B(a,b)}[1 - K(x)^c]^{\,b+1}.$$

Observe that,

$$\eta(x)h(x) - g(x) = -\frac{1}{b+1}K(x)^{c\,(1-a)}[1 - K(x)^c] < 0.$$

Conversely, if η is given as above, then

$$s'(x) = \frac{\eta'(x)\,h(x)}{\eta(x)\,h(x) - g(x)} = \frac{c\,b\,k(x)\,K(x)^{c-1}}{1 - K(x)^c}$$

and hence

$$s(x) = -\log[1 - K(x)^c]^{\,b} + C_1,$$

where C_1 is a constant. Now, in view of Theorem 2.1, X has pdf (1) and cdf (2).

Corollary 2.1 *Let $X: \Omega \to (0, \infty)$ be a continuous random variable and let $h(x)$ be as in Proposition 2.1. The pdf of X is (1) if and only if there exist functions g and η defined in Theorem 2.1 satisfying the differential equation*

$$\frac{\eta'(x)h(x)}{\eta(x)h(x) - g(x)} = \frac{cbk(x)K(x)^{c-1}}{1 - K(x)^c}, \quad x > 0.$$

Remarks 2.2 (a) *The general solution of the differential equation in Corollary 2.1 is*

$$\eta(x) = [1 - K(x)^c]^{-b}\left[-\int g(x)c\,b\,k\,(x)K(x)^{c\,a-1}[1 - K(x)^c]^{b-1}\,\mathrm{d}x + D\right],$$

for $x > 0$, where D is a constant. One set of appropriate functions satisfying the above equation is given in Proposition 1.2 with $D = 0$.

(b) Clearly, there are other triplets of functions (h, g, η) satisfying the conditions of Theorem 2.1.

Proposition 2.2 *Let $X: \Omega \to \mathbb{R}$ be a continuous random variable and let*

$$h(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)^{c(1-a)} \quad \text{and} \quad g(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)^{c(1-a)}\left[1 - \Phi\left(\frac{x-\mu}{\sigma}\right)^c\right]$$

for $x \in \mathbb{R}$. The cdf of X is (3) if and only if the function η defined in Theorem 2.1 has the form $\eta(x) = \frac{b}{b+1}\left\{1 - \left[\Phi\left(\frac{x-\mu}{\sigma}\right)\right]^c\right\}$, $x \in \mathbb{R}$.

Proposition 2.3 *Let $X: \Omega \to (0, \infty)$ be a continuous random variable and let*

$$h(x) = e^{\alpha(b-1)U(x)}\left[1 - e^{-\alpha U(x)}\right]^{1-a} \quad \text{and} \quad g(x) = e^{-\alpha U(x)}\left[1 - e^{-\alpha U(x)}\right]^{1-a}$$

for $x \in (0, \infty)$. The cdf of X is (7) if and only if the function η defined in Theorem 2.1 has the form $\eta(x) = (b+1)^{-1}\,e^{-\alpha\,b\,U(x)}$, $x > 0$.

Proposition 2.4 *Let $X: \Omega \to (0, \infty)$ be a continuous random variable and let*

$$h(x) = V(x;\xi)^{1-\delta} \quad \text{and} \quad g(x) = e^{-\alpha V(x;\xi)}\,V(x;\xi)^{1-\delta}$$

for $x \in (0, \infty)$. The cdf of X is (10) if and only if the function η defined in Theorem 2.1 has the form $\eta(x) = \frac{1}{2}e^{-\alpha V(x;\xi)}$, $x > 0$.

Remarks 2.3 (*a*) *Letting* $a = 1$ *and then calling* c *as* a, *the pdf (1) with* $K(x) = exp\left(-\frac{\alpha}{x^\beta}\right)$ *and* $k(x) = \frac{\alpha\beta}{x^{\beta+1}} exp\left(-\frac{\alpha}{x^\beta}\right)$ *reduces to the pdf (4). So, the Kw-IW model is a special case of the McG-K distribution.* (*b*) *A corollary and a remark similar to Corollary 2.1 and Remark 2.2(a) can be stated for the BEW, GEW and McN distributions in the same way. For example, for the BEW distribution, the general solution of the differential equation is*

$$\eta(x) = e^{\alpha U(x)}\left[-\int a\,u(x)\,g(x)\,e^{-\alpha b\,U(x)}\left[1 - e^{-\alpha U(x)}\right]^{a-1} dx + D\right],$$

and for $g(x)$ and $\eta(x)$ given in Proposition 2.1, the constant $D = 0$.

Proof. We have $h(x) = e^{\alpha(b-1)U(x)}\left[1 - e^{-\alpha U(x)}\right]^{1-a}$, $g(x) = e^{-\alpha U(x)}\left[1 - e^{-\alpha U(x)}\right]^{1-a}$, $\eta(x) = \frac{1}{b+1}\,e^{-\alpha b\,U(x)}$ and $\eta'(x) = \frac{-a\,b\,u(x)}{b+1}e^{-\alpha b\,U(x)}$. Thus,

$$\frac{\eta'(x)h(x)}{\eta(x)h(x) - g(x)} = \frac{\frac{-abu(x)}{b+1}e^{-\alpha b\,U(x)}e^{\alpha(b-1)U(x)}\left[1 - e^{-\alpha U(x)}\right]^{1-a}}{\frac{1}{b+1}e^{-\alpha bU(x)}e^{\alpha(b-1)U(x)}[1 - e^{-\alpha U(x)}]^{1-a} - e^{-\alpha U(x)}[1 - e^{-\alpha U(x)}]^{1-a}}$$

or

$$\frac{\eta'(x)h(x)}{\eta(x)h(x) - g(x)} = \frac{-abu(x)e^{-\alpha bU(x)}e^{\alpha(b-1)U(x)}}{e^{-\alpha bU(x)}e^{\alpha(b-1)U(x)} - (b+1)e^{-\alpha U(x)}}$$

or

$$\frac{\eta'(x)h(x)}{\eta(x)h(x) - g(x)} = \frac{a\,u(x)e^{-\alpha b\,U(x)}e^{\alpha(b-1)U(x)}}{e^{-\alpha U(x)}} = a\,u(x)$$

or

$$\eta'(x)h(x) - a\,u(x)\eta(x)h(x) = -a\,u(x)g(x)$$

or

$$\eta'(x) - a\,u(x)\eta(x) = -a\,u(x)g(x)e^{-\alpha(b-1)U(x)}\left[1 - e^{-\alpha U(x)}\right]^{a-1}$$

or

$$e^{-\alpha U(x)}\eta'(x) - a\,u(x)e^{-\alpha U(x)}\eta(x) = -a\,u(x)g(x)e^{-\alpha bU(x)}\left[1 - e^{-\alpha U(x)}\right]^{a-1}$$

or

$$\eta(x) = e^{\alpha U(x)}\left[-\int au(x)g(x)e^{-\alpha bU(x)}\left[1 - e^{-\alpha U(x)}\right]^{a-1}dx + D\right]$$

or

$$\eta(x) = e^{\alpha U(x)}\left[-\int a\,u(x)e^{-\alpha U(x)}\left[1 - e^{-\alpha U(x)}\right]^{1-a}e^{-\alpha bU(x)}\left[1 - e^{-\alpha U(x)}\right]^{a-1}dx + D\right]$$

or

$$\eta(x) = e^{\alpha U(x)}\left[-\int a\,u(x)e^{-\alpha U(x)}e^{-\alpha bU(x)}dx + D\right]$$

or

$$\eta(x) = e^{\alpha U(x)}\left[-\int a\,u(x)e^{-\alpha(b+1)U(x)}dx + D\right] = e^{\alpha U(x)}\left[\frac{1}{b+1}e^{-\alpha(b+1)U(x)} + D\right]$$

$$= \frac{1}{b+1}e^{-\alpha bU(x)},$$

where $D = 0$.

3. Truncated moment of the nth order statistic

Let $X_{1:n} \leq X_{2:n} \leq \ldots \leq X_{n:n}$ be the corresponding order statistics from a random sample of size n from a continuous cdf F. We briefly discuss here characterization results based on functions of the nth order statistic. We have the following proposition.

Proposition 3.1 *Let $X:\Omega \to (0,\infty)$ be a continuous random variable with cdf F. Let ψ and q be two differentiable functions in $(0,\infty)$ such that $\lim_{x\to 0}\psi(x)F(x)^n = 0$, $\int_0^\infty \frac{q'(t)}{[\psi(t)-q(t)]}dt = \infty$.*

Then,
$$E[\psi(X_{n:n})|X_{n:n} < t] = q(t), \ t > 0, \tag{11}$$
implies
$$F(x) = \exp\left\{-\int_x^\infty \frac{q'(t)}{n[\psi(t)-q(t)]}dt\right\}, \quad x \geq 0. \tag{12}$$

Proof. If (11) holds, then using integration by parts on the left hand side of (11) and the condition $\lim_{x\to 0}\psi(x)F(x)^n = 0$, we have $\int_0^t \psi'(x)\big(F(x)\big)^n dx = [\psi(t) - q(t)]F(t)^n$.

Differentiating both sides of the above equation with respect to t, we obtain
$$\frac{f(t)}{F(t)} = \frac{q'(t)}{n[\psi(t) - q(t)]}, \quad t > 0.$$

Now, integrating the last equation from x to ∞ , we have, in view of $\int_0^\infty \frac{q'(t)}{[\psi(t)-q(t)]}dt = \infty$, that the cdf F is given by (12).

Remarks 3.1. (a) Taking, for instance, $\psi(x) = \left[1 - e^{-\alpha U(x)}\right]^{na}$ and $q(x) = \frac{1}{2}\psi(x)$ in Proposition 3.1 the above equation reduces to $f(x)F(x)^{-1} = a\,\alpha\,u(x)\,e^{-\alpha U(x)}[1 - e^{-\alpha U(x)}]^{-1}$, from which, in view of (12), we have $F(x) = \left[1 - e^{-\alpha U(x)}\right]^a$, which is the cdf (7) with $b = 1$.

(b) Taking, for instance, $\psi(x) = \left[\Phi\left(\frac{x-\mu}{\sigma}\right)\right]^{nac}$ and $q(x) = \frac{1}{2}\psi(x)$ in Proposition 3.1 the last above equation becomes $f(x)F(x)^{-1} = ac\frac{d}{dx}\left\{\left(\Phi\left(\frac{x-\mu}{\sigma}\right)\right)^{c(1-a)}\right\}\left(\Phi\left(\frac{x-\mu}{\sigma}\right)\right)^{-1}$, from which, in view of (12), we have $F(x) = \left[\Phi\left(\frac{x-\mu}{\sigma}\right)\right]^{ac}$, which is the cdf (3) with $b = 1$.

4. Characterizations based on the truncated moment of the first order statistic

We state here two characterizations based on certain functions of the first order statistic. We like to mention that the proof of Proposition 4.1 below is a straightforward extension of Theorem 2.2 of Hamedani (2010). We give a short proof of it for the sake of completeness.

900

Pak.j.stat.oper.res. Vol.XIII No.4 2017 pp893-908

Proposition 4.1 *Let* $X: \Omega \to (0, \infty)$ *be a continuous random variable with cdf* F. *Let* $\psi(x)$ *and* $q(x)$ *be two differentiable functions on* $(0, \infty)$ *such that* $\lim_{x\to\infty} \psi(x)[1 - F(x)]^n = 0$, $\int_0^\infty \frac{q'(t)}{[q(t)-\psi(t)]} dt = \infty$. *Then,*

$$E[\psi(X_{1:n})|X_{1:n} > t] = q(t), \quad t > 0, \tag{13}$$

implies $F(x) = 1 - \exp\left\{-\int_0^x \frac{q'(t)}{n[q(t)-\psi(t)]} dt\right\}, \quad x \geq 0.$

Proof. If (13) holds, then using integration by parts on the left hand side of (13) and the assumption $\lim_{x\to\infty} \psi(x)[1 - F(x)]^n = 0$, we have $\int_t^\infty \psi'(x)\left(1 - F(x)\right)^n dx = [q(t) - \psi(t)]\left(1 - F(t)\right)^n.$

Differentiating both sides of the above equation with respect to t, we obtain

$$\frac{f(t)}{1-F(t)} = \frac{q'(t)}{n[q(t)-\psi(t)]}, \quad t > 0. \tag{14}$$

Now, integrating both sides of (14) from 0 to x, we have, in view of $\int_0^\infty \frac{q'(t)}{[q(t)-\psi(t)]} dt = \infty$, the cdf F given in Proposition 4.1.

Remarks 4.1. (a) Taking, for instance, $\psi(x) = e^{-naV(x;\xi)}$ and $q(x) = 1/2 \, \psi(x)$ in Proposition 4.1, we obtain (10) for $b = 1$. (b) Taking, for instance, $\psi(x) = \left[1 - e^{-\frac{a\alpha}{x^\beta}}\right]^{nb}$ and $q(x) = \frac{1}{2}\psi(x)$ in Proposition 4.1, we obtain (5).

5. Useful representation

Theorem 5.1 *Let* X *be a random variable having any of the five families of distributions discussed so far and the function* $m_k(a, c) = c(a + k)$, *where* $k = 1, 2, \dots$ *and* $a, c \in \mathbb{R}^+$. *The pdf of* X *can be expressed as the linear combination*

$$f(x) = \sum_{k=0}^\infty b_k \, h_{c(a+k)}(x), \tag{15}$$

where $h_{c(a+k)}(x)$ *denotes the* $Exp - G\left(c(a + k)\right)$ *density function.*

Proof. First, consider the GG family. From equation (8) and based on an expansion due to Nadarajah *et al.* (2015), we can write (for $a > 0$)

$$f(x) = \sum_{k=0}^\infty b_k \, h_{a+k}(x) = \sum_{k=0}^\infty b_k \, h_{m_k(a,1)}(x), \tag{16}$$

where $b_k = \frac{1}{(a+k)\Gamma(a-1)} k + 1 - a_k \sum_{j=0}^k \frac{(-1)^{j+k} \, p_{j,k}}{(a-1-j)} k_j$, and $h_{m_k(a,1)}(x) = h_{(a+k)}(x)$ denotes the Exp-$G(c(a + k))$ density function with $c = 1$.

Second, we consider $X \sim$McG-K(a, b, c, τ). Expanding the binomial in (1) yields:

$$f(x) = c \, B(a, b+1)^{-1} f(x) \sum_{k=0}^\infty (-1)^k b_k F(x)^{c(a+k)-1} =$$

$$\sum_{k=0}^\infty b_k \, h_{m_k(a,c)}(x), \tag{17}$$

where $h_{m_k(a,c)}(x) = h_{c(a+k)}(x)$ denotes the density of Exp-G($c(a + k)$) and $b_k = (-1)^k (a + k)^{-1} b_k B(a, b+1)^{-1}.$

Third, we consider $X \sim \mathrm{BG}(a, b, \tau)$. This distribution is a special case of the McG-K distribution with $c = 1$ and the same b_k. Now, consider $X \sim \mathrm{Kw\text{-}G}(a, b, \tau)$. This distribution is a special case of the McG-K distribution too, but with $a = 1$, changing $c = a$, and $b_k = (-1)^k (k+1)^{-1} b - 1_k b$. Thus, we prove equation (15) in five parts (for each five families), as shown in equations (16) and (17). Besides that, each family has specific weights.

6. Mathematical properties

In this section, we derive moments, moment generating function (mgf) and quantile function (qf) of those distributions.

6.1 Moments

We derive several representations for the moment $\mu_s' = E(X^s)$ of X having all of five families discussed in this paper. Note that other kinds of moments related to the L-moments of Hosking (1990) may also be obtained in closed-form, but we confine ourselves here to μ_s' for brevity.

Henceforth, we assume that $Y_{c(a+k)} \sim \mathrm{Exp\text{-}G}(c(a + k))$. The importance of moments in Statistics especially in applications is obvious. A first formula for the nth moment of X can be obtained from (15) and the monotone convergence theorem as $\mu_n' = E(X^n) = \sum_{k=0}^{\infty} b_k E(Y_{c(a+k)}^n)$. A second formula for $E(X^n)$ follows from the last identity in terms of the baseline qf $Q_G(u) = G^{-1}(u)$ as $\mu_n' = \sum_{k=0}^{\infty} c(a + k) b_k \tau(n, k)$, where $\tau(n, k) = \int_{-\infty}^{\infty} x^n G(x)^k g(x) dx = \int_0^1 Q_G(u)^n u^k du$.

6.2 Moment generating function

The mgf provides the basis of an alternative route to analytical results compared with working directly with the pdf and cdf and it is widely used in the characterization of distributions and the application of the skew-normal test (Meintanis, 2010) and other goodness of fit tests (Ghosh, 2013).

Here, we provide two formulae for the mgf $M(t) = E[\exp(t X)]$ of X. A first formula for $M(t)$ comes from (15) and the monotone convergence theorem as $M(t) = \sum_{k=0}^{\infty} b_k M_{c(a+k)}(t)$, where $M_{c(a+k)}(t)$ is the mgf of $Y_{c(a+k)}$. Hence, $M(t)$ can be determined from the generating function of the Exp-G distribution. An alternative formula for $M(t)$ can be derived from the last identity as $M(t) = \sum_{i=0}^{\infty} c(a + k) b_k \rho(t, k)$, where $\rho(t, k) = \int_{-\infty}^{\infty} e^{tx} G(x)^k g(x) dx = \int_0^1 \exp\{t Q_G(u)\} u^k du$.

6.3 Quantile function

The GBG qf is obtained by inverting the parent cdf $K(x)$. We have $Q_{GBG}(u; \tau, a, b, c) = K^{-1}\left(\left[Q_{\beta(a,b)}(u)\right]^{1/c}\right)$, where $Q_{\beta(a,b)}(u) = I^{-1}(u; a, b)$ is the ordinary beta qf. It is possible to obtain some expansions for the beta qf with positive parameters a and b. One of them can be found on the Wolfram website

(http://functions.wolfram.com/06.23.06.0004.01) as $z = Q_{\beta(a,b)}(u) = a_1 v + a_2 v^2 + a_3 v^3 + a_4 v^4 + O(v^{5/a})$, where $v = [a\, B(a,b)\, u]^{1/a}$ for $a > 0$ and $a_0 = 0$, $a_1 = 1$, $a_2 = (b-1)/(a+1)$, $a_3 = (b-1)[a^2 + (3b-1)a + 5b - 4]/[2(a+1)^2(a+2)]$, $a_4 = (b-1)[a^4 + (6b-1)a^3 + (b+2)(8b-5)a^2 + (33b^2 - 30b + 4)a + b(31b - 47) + 18]/[3(a+1)^3(a+2)(a+3)]$, ... The coefficients a_i (for $i \geq 2$) can be derived from a cubic recursion of the form $a_i = [i^2 + (a-2)i - (a-1)]^{-1}\{(1 - \delta_{i,2})\sum_{r=2}^{i-1} a_r\, a_{i+1-r}\, [r(1-a)(i-r) - r(r-1)] + \sum_{r=1}^{i-1}\sum_{s=1}^{i-r} a_r\, a_s\, a_{i+1-r-s}\, [r(r - a) + s(a + b - 2)(i + 1 - r - s)]\}$, where $\delta_{i,2} = 1$ if $i = 2$ and $\delta_{i,2} = 0$ if $i \neq 2$. In the last equation, we note that the quadratic term only contributes for $i \geq 3$.

7. Applications

In this section, we compare the fits of the BG, GG, KwG and McG with the baselines Gamma (Γ), Weibull (W) and Inverse Weibull (IW) to two real data sets from Murthy *et al.* (2004).

7.1 Application 1: Stress data

These data refer to accelerated life testing of ($n = 40$) items with change in stress from 100 to 150 at $t = 15$. The data are:

4.79, 7.17, 7.31, 7.43, 7.84, 8.49, 8.94, 9.40, 9.61, 9.84, 10.58, 11.18, 11.84, 13.28, 14.47, 14.79, 15.54, 16.90, 17.25, 17.37, 18.69, 18.78, 19.88, 20.06, 20.10, 20.95, 21.72, 23.87.

Table 1 provides a summary of these data. The stress data have positive skewness and negative kurtosis.

Table 1: Descriptive statistics. [a]There are various modes.

Data	Mean	Median	Mode	Std. Dev.	Variance	Skewness	Kurtosis	Min.	Max.
Stress	10.45	9.51	1.3[a]	6.99	48.86	0.23	-1.19	0.13	23.87

Table 2 lists the values of the following statistics for some models: Akaike Information Criterion (AIC), Consistent Akaike Information Criterion (AICc) and Bayesian Information Criterion (BIC). The figures involving the Γ and IW baselines in Table 2 indicate that the Kw-G model has the smallest values of these statistics among all fitted models. So, it could be chosen as the more suitable model in these cases (when we use gamma and IW as the parent distributions). However, note that when the baseline is Weibull, the GG family presents better performance than the others. Thus, we can say that is important to propose new generators in order to provide better fits to real data sets.

Table 2: Relative goodness-of-fit for the selected generators

Models	Measures		
	AIC	AICc	BIC
(Baseline: Gamma)			
BΓ	262.5893	263.7322	269.3448
GΓ	267.1444	267.8111	272.2111
KwΓ	**261.6627**	**262.8056**	**268.4182**
McΓ	263.6835	265.4482	272.1279
(Baseline: Weibull)			
BW	282.0882	283.2311	288.8437
GW	**261.4682**	**262.1349**	**266.5349**
KwW	271.4083	272.5511	278.1638
McW	289.3399	291.1046	297.7843
(Baseline: Inverse Weibull)			
BIW	288.9176	290.0604	295.6731
GIW	295.3439	296.0105	300.4105
KwIW	**278.4004**	**279.5432**	**285.1559**
McIW	279.9951	281.7598	288.4395

Besides that, note that when we compare the models KwΓ, GW and KwIW (models that field better adjustments), the best of them was the second, showing, in this study, that the gamma generator provides the best performance among the others generators. Moreover, we also provide a visual comparison of the histogram of the data with the fitted density functions. The plots of the fitted densities for the baselines Γ, W and IW are displayed in Figures 1(a), 1(b) and 1(c), respectively, for the data set. We only reinforce what has been said above.
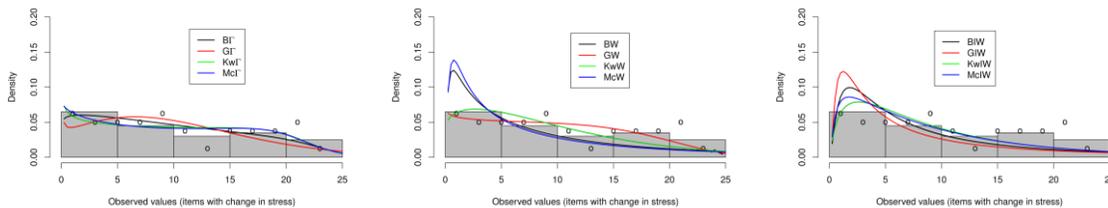

Figure 1: Estimated densities of the selected generators for stress data.

7.2 Application 2: Repairable data

The following data refer to the time between failures for repairable itens ($n = 30$):
1.43, 0.11, 0.71, 0.77, 2.63, 1.49, 3.46, 2.46, 0.59, 0.74, 1.23, 0.94, 4.36, 0.40, 1.74, 4.73, 2.23, 0.45, 0.70, 1.06, 1.46, 0.30, 1.82, 2.37, 0.63, 1.23, 1.24,1.97, 1.86, 1.17.

Table 3 provides a summary of these data. The repairable data has positive skewness and kurtosis, and has less variability.

Table 3: Descriptive statistics

904

Pak.j.stat.oper.res. Vol.XIII No.4 2017 pp893-908

Data	Mean	Median	Mode	Std. Dev.	Variance	Skewness	Kurtosis	Min.	Max.
Repairable	1.54	1.24	1.23	1.13	1.27	1.37	1.8	0.11	4.73

Table 4 lists the values of the following statistics for some models: AIC, AICc and BIC. The figures involving Γ and W baselines in Table 4 indicate that the GG model has the smallest values of these statistics among all fitted models. So, it could be chosen as the more suitable model in this case (when we take gamma and Weibull as the baselines). However, note that when the baseline is Weibull, the GG generator presents better performance than the others, as in the first application. Besides that, note too that when we compare the GΓ, GW and KwIW models (those that yield better adjustments), the best of them is the second, showing, in this study, that the GG generator provides the best performance among the other current models. These results are exhibited in Figure 2.

Table 4:   Relative goodness-of-fit for the selected generators

| Models	Measures		
	AIC	AICc	BIC
(Baseline: Gamma)			
BΓ	87.22407	88.82407	92.82886
GΓ	**85.25093**	**86.17401**	**89.45453**
KwΓ	87.2274	88.8274	92.83219
McΓ	89.23693	91.73693	96.24291
(Baseline: Weibull)			
BW	87.19683	88.79683	92.80162
GW	**85.23609**	**86.15917**	**89.43968**
KwW	87.24511	88.84511	92.8499
McW	90.57175	93.07175	97.57774
(Baseline: Inverse Weibull)			
BIW	87.70437	89.30437	93.30916
GIW	92.60976	93.53284	96.81335
KwIW	**87.48851**	**89.08851**	**93.0933**
McIW	102.6596	105.1596	109.6656
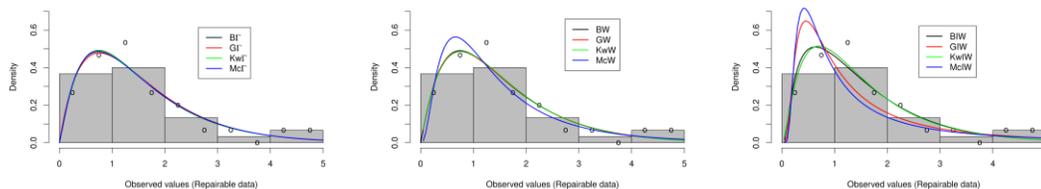
Figure 2: Estimated densities for the selected generators for repairable data.

8.   Concluding remarks

We study the characterizations of some important classes of generalized distributions such as the beta-G, Gamma-G, Kumaraswamy-G and McDonald-G, in three different directions. We believe that our characterizations will be the only ones for some classes due to the complexity of their cumulative distribution functions. Further, we discuss certain properties of these distributions, which would be valuable to researchers in applications.

References

1. Ahmadi J., Fashandi M. (2009). Some characterization and ordering results based on entropies of current records. *Statistics Probability Letters*, 79, 2053-2059. ISSN 0167-7152.

2. Ahsanullah M., Hamedani G.G. (2007). Certain characterizations of power function and beta distributions based on order statistics. *Journal of Statistical Theory and Applications*, 6, 220-226.

3. Akhundov I., Nevzorov V.B. (2010). A simple characterization of Student's t3 distribution. *Statistics Probability Letters*, 80, 293295. ISSN 0167-7152.

4. Alexander C., Cordeiro G.M., Ortega E.M.M., Sarabia J.M. (2012). Generalized beta-generated distributions. *Computational Statistics Data Analysis*, 56, 1880-1897. ISSN 0167-9473.

5. Bairamov I., Ahsanullah M., Pakes A. (2005). A characterization of continuous distributions via regression on pairs of record values. *Australian New Zealand Journal of Statistics*, 47, 543-547. ISSN 1369-1473.

6. Baratpour S., Ahmadi J., Arghami N.R. (2008). Characterizations based on Renyi entropy of order statistics and record values. *Journal of Statistical Planning and Inference*, 138, 2544-2551. ISSN 0378-3758.

7. Beg M.I., Ahsanullah M. (2007). On characterizing distributions by conditional expectations of functions of generalized order statistics. *Journal of Applied Statistical Science*, 15, 229-244.

8. Bieniek M. (2007). On characterizations of distributions by regression of adjacent generalized order statistics. *Metrika*, 66, 233-242. ISSN 0026-1335.

9. Cordeiro G.M., de Castro M. (2011). A new family of generalized distributions. *Journal of Statistical Computation and Simulation*, 81, 883-898. http://dx.doi.org/10.1080/ 00949650903530745.

10. Cordeiro G.M., Silva G.O., Ortega E.M.M. (2012). The beta extended Weibull family. *Journal of Probability and Statistical Science*, 10, 15-40.

11. Eugene N., Lee C., Famoye F. (2002). Beta-normal Distribution and Its Applications. *Com- munications in Statistics - Theory and Methods*, 31, 497-512.

12. Galambos J., Kotz S. (1978). Characterizations of probability distributions. A unifed approach with an emphasis on exponential and related models. Number 675 in Lecture Notes in Mathematics. Springer, Berlin.

13. Ghosh S. (2013). Normality testing for a long-memory sequence using the empirical moment generating function. *Journal of Statistical Planning and Inference*, 143, 944-954. ISSN 0378-3758.

14. Glanzel W. (1987). A characterization theorem based on truncated moments and its application to some distribution families, pp. 75-84. *Mathematical Statistics and Probability Theory (Bad Tatzmannsdorf, 1986)*. Dordrecht: Reidel.

15. Gl• anzel W., Hamedani G.G. (2001). Characterizations of univariate continuous distributions. *Studia Scientiarum Mathematicarum Hungarica*, 37, 83-118.

16. Hamedani G.G. (1993). Characterizations of Cauchy, normal and uniform distributions. *Studia Scientiarum Mathematicarum Hungarica*, 28, 243-247.

17. Hamedani G.G. (2002). Characterizations of univariate continuous distributions. II. *Studia Scientiarum Mathematicarum Hungarica*, 39, 407-424.

18. Hamedani G.G. (2006). Characterizations of univariate continuous distributions. III. *Studia Scientiarum Mathematicarum Hungarica*, 43, 361-385.

19. Hamedani G.G. (2010). Characterizations of continuous univariate distributions based on the truncated moments of functions of order statistics. *Studia Scientiarum Mathematicarum Hungarica*, 47, 462-484.

20. Hamedani G.G., Ahsanullah M. (2011). Characterizations of Weibull Geometric Distribution. *Journal of Statistical Theory and Applications*, 10, 581-590.

21. Haque Z., Athar H., Khan R.U. (2009). Characterization of probablity distributions through expectation of function of generalized order statistics. *Journal of Statistical Theory and Applications*, 8, 416-426.

22. Hashimoto E.M, Ortega E.M.M., Cordeiro GM, Cancho VG (2014). The Poisson Birnbaum- Saunders model with long-term survivors. *Statistics*, 48, 1394-1413. http://dx.doi.org/ 10.1080/02331888.2013.822503.

23. Hosking J.R.M. (1990). Ł-Moments: Analysis and Estimation of Distributions Using Linear Combinations of Order Statistics. *Journal of the Royal Statistical Society*. Series B (Methodological), 52, 105-124. ISSN 00359246.

24. Khan A.H., Anwar Z., Chishti S. (2010). Characterization of continuous distributions through conditional expectation of functions of dual generalized order statistics. *Pakistan Journal of Statistical*, 26, 615-628.

25. Kumaraswamy P. (1980). A generalized probability density function for double-bounded random processes. *Journal of Hydrology*, 46, 79-88. ISSN 0022-1694.

26. McDonald J.B. (2008). Some Generalized Functions for the Size Distribution of Income. *In D Chotikapanich (ed.), Modeling Income Distributions and Lorenz Curves, volume 5 of Economic Studies in Equality, Social Exclusion and Well-Being*, pp. 37-55. Springer New York. ISBN 978-0-387-72756-1.

27. Mead M. (2014). An Extended Pareto Distribution. *Pakistan Journal of Statistics and Operation Research*, 10, 1816-2711.

28. Meintanis S.G. (2010). Testing skew normality via the moment generating function. *Mathe- matical Methods of Statistics*, 19, 64-72. ISSN 1066-5307.

29.   Murthy D.N.P., Xie M., Jiang R. (2004). Weibull Models. Wiley Series in Probability and Statistics. Wiley. ISBN 9780471473275.

30.   Nadarajah S., Cordeiro G.M., Ortega E.M.M. (2015). The gamma-G family of distributions: Mathematical properties and applications. *To appear in Communications in Statistics Theory and Methods.*

31.   Nadarajah S., Kotz S. (2005). On some recent modifications of Weibull distribution. *IEEE Transaction on Reliability*, 54, 561-562.

32.   Nascimento A. D. C., Pereira M.B, Milena Zea L, Santo-Neto M, Bernardo R, Cordeiro GM (2014). The gamma extended Weibull family of distributions. *Journal of Statistical Theory and Applications*, 13, 1-16.

33.   Nevzorov V.B., Balakrishnan N., Ahsanullah M. (2003). characterizations of Student's t2-distribution. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 52, 395-400. ISSN 1467-9884.

34.   Ristic M.M., Balakrishnan N. (2012). The gamma-exponentiated exponential distribution. *Journal of Statistical Computation and Simulation*, 82, 1191-1206.

35.   Su J.C., Su N.C., Huang W.J. (2008). Characterizations based on record values and order statistics. *Journal of Statistical Planning and Inference*, 138, 1358-1367. ISSN 0378-3758.

36.   Tavangar M., Asadi M. (2007). Generalized Pareto Distributions Characterized by Generalized Order Statistics. *Communications in Statistics - Theory and Methods*, 36, 1333-1341.

37.   Yanev G.P., Ahsanullah M. (2012). Characterizations of Student's t3-distribution via regressions of order statistics. *Statistics*, 46, 429-435.

38.   Zografos K., Balakrishnan N. (2009). On families of beta and generalized gamma-generated distributions and associated inference. *Statistical Methodology*, 6, 344-362. ISSN 1572- 3127.