

## The Burr X Generator of Distributions for Lifetime Data

Haitham M. Yousof and Ahmed Z. Afify

*Department of Statistics, Mathematics and Insurance  
Benha University, Egypt*

G. G. Hamedani

*Department of Mathematics, Statistics and Computer Science  
Marquette University, USA*

Gokarna Aryal

*Department of Mathematics, Statistics and Computer Science  
Purdue University Northwest, USA*

Received 2 May 2016

Accepted 20 July 2016

In this paper, we introduce a new class of distributions called the *Burr X* family. Some of its mathematical and structural properties are derived. The maximum likelihood is used for estimating the model parameters. The importance and flexibility of the new family are illustrated by means of an application to real data set.

*Keywords:* The Burr X Distribution, Order Statistics, Maximum Likelihood Estimation, Simulation

2000 Mathematics Subject Classification: 60E05, 62P99

### 1. Introduction

The statistical literature contains many new classes of distributions which have been constructed by extending common families of continuous distributions by means of adding one or more shape parameters. The inducted extra parameter(s) to the existing probability distribution have been shown to improve the flexibility and goodness of fits of the distribution against the intuition of model parsimony. Therefore, many methods of adding a parameter to distributions have been proposed by several researchers and these new families have been used for modeling data in many applied areas such as engineering, economics, biological studies, environmental sciences and many more. In fact the modern computing technology has made many of these techniques accessible if the analytical solutions are very complicated.

Gupta et al. [18] defined the exponentiated-G (exp-G) class, which consists of raising the cumulative distribution function (cdf) to a positive power parameter and proposed the exponentiated exponential (EE) distribution, defined by the cdf (for  $x > 0$ )  $F(x) = [1 - \exp(-\lambda x)]^\theta$ , where  $\lambda, \theta > 0$ . This equation is simply the  $\theta$ th power of the standard exponential cumulative distribution. Many

generalized distribution functions are constructed in a similar manner, for example the exponentiated gamma, exponentiated Fréchet and exponentiated Gumbel distributions [23], although the way they defined the cdfs of the last two distributions is slightly different.

Several other classes can be mentioned such as the Marshall-Olkin-G (MO-G) family by Marshall and Olkin [22], beta generalized-G (BG-G) family by Eugene et al. [14], Kumaraswamy-G (Kw-G) family by Cordeiro and de Castro [9] and exponentiated generalized-G (EG-G) family by Cordeiro et al. [8], the Lomax generator of distributions by Cordeiro et al. [12], beta odd log-logistic generalized (BOLL-G) by Cordeiro et al. [11], beta Marshall-Olkin (BMO-G) by Alizadeh et al. [4], Kumaraswamy odd log-logistic (KwOLL-G) by Alizadeh et al. [6], Kumaraswamy Marshall-Olkin (KwMO-G) by Alizadeh et al. [5], generalized transmuted-G (GT-G) by Nofal et al. [24], transmuted exponentiated generalized-G (TExG-G) by Yousof et al. [26], Kumaraswamy transmuted-G by Afify et al. [2] and transmuted geometric-G by Afify et al. [1].

In this paper, we introduce a new class of distributions using the Burr-X generator. Burr [7] introduced twelve different forms of cumulative distribution functions for modeling data. Among those twelve distribution functions, Burr-Type X and Burr-Type XII received the maximum attention. The two-parameter Burr-Type X distribution is related to well studied and popular distributions like gamma distribution, Weibull distribution so we will choose Burr-X distribution for our generator.

The rest of the paper is organized as follows. In Section 2, we define the Burr X-G (BX-G) family of distributions. In Section 3, we provide a useful mixture representation for its probability density function (pdf). In Section 4, we derive some of its general mathematical properties. In Section 5, we present some characterizations of BX-G family. Two special models of BX-G family are discussed in Section 6. Maximum likelihood estimation of the model parameters is addressed in Section 7. In Section 8, simulation results to assess the performance of the proposed maximum likelihood estimation procedure are discussed. In Section 9, we provide application to real data to illustrate the importance and flexibility of the new family. Finally, some concluding remarks are presented in Section 10.

## 2. The new family

Consider the cdf and the pdf of the Burr X distribution

$$F(x; \theta) = [1 - \exp(-x^2)]^\theta, \quad x > 0, \theta \geq 0 \tag{1}$$

and

$$f(x; \theta) = 2\theta x \exp(-x^2) [1 - \exp(-x^2)]^{\theta-1}, \tag{2}$$

respectively, where  $\theta$  is the shape parameter. Let  $g(x; \xi)$  and  $G(x; \xi)$  denote the density and cumulative functions of the baseline model with parameter vector  $\xi$  and consider the Burr type X cdf  $V(t) = [1 - \exp(-t^2)]^\theta$  (for  $t > 0$ ) with positive parameter  $\theta$ . We replace the argument  $t$  by  $G(x; \xi)/\overline{G}(x; \xi)$ , where  $\overline{G}(x; \xi) = 1 - G(x; \xi)$ . Hence, the cdf of the Burr X generator (BX-G) becomes

$$F(x; \theta, \xi) = 2\theta \int_0^{\frac{G(x; \xi)}{\overline{G}(x; \xi)}} t \exp(-t^2) [1 - \exp(-t^2)]^{\theta-1} dt = \left\{ 1 - \exp \left[ - \left( \frac{G(x; \xi)}{\overline{G}(x; \xi)} \right)^2 \right] \right\}^\theta. \tag{3}$$

The pdf of the BX-G is given by

$$f(x; \theta, \xi) = \frac{2\theta g(x; \xi)G(x; \xi)}{\bar{G}(x; \xi)^3} \exp \left[ - \left( \frac{G(x; \xi)}{\bar{G}(x; \xi)} \right)^2 \right] \left\{ 1 - \exp \left[ - \left( \frac{G(x; \xi)}{\bar{G}(x; \xi)} \right)^2 \right] \right\}^{\theta-1}. \quad (4)$$

The reliability function ( $R(x)$ ), hazard rate function ( $h(x)$ ), reversed hazard rate function ( $r(x)$ ) and cumulative hazard rate function ( $H(x)$ ) of  $X$  are given, respectively, by

$$R(x; \theta, \xi) = 1 - \left\{ 1 - \exp \left[ - \left( \frac{G(x; \xi)}{\bar{G}(x; \xi)} \right)^2 \right] \right\}^{\theta},$$

$$h(x; \theta, \xi) = \frac{2\theta g(x; \xi)G(x; \xi) \exp \left[ - \left( \frac{G(x; \xi)}{\bar{G}(x; \xi)} \right)^2 \right] \left\{ 1 - \exp \left[ - \left( \frac{G(x; \xi)}{\bar{G}(x; \xi)} \right)^2 \right] \right\}^{\theta-1}}{\bar{G}(x; \xi)^3 \left( 1 - \left\{ 1 - \exp \left[ - \left( \frac{G(x; \xi)}{\bar{G}(x; \xi)} \right)^2 \right] \right\}^{\theta} \right)},$$

$$r(x; \theta, \xi) = \frac{2\theta g(x; \xi)G(x; \xi)}{\bar{G}(x; \xi)^3} \exp \left[ - \left( \frac{G(x; \xi)}{\bar{G}(x; \xi)} \right)^2 \right] \left\{ 1 - \exp \left[ - \left( \frac{G(x; \xi)}{\bar{G}(x; \xi)} \right)^2 \right] \right\}^{-1}$$

and

$$H(x; \theta, \xi) = - \left[ \log \left( 1 - \left\{ 1 - \exp \left[ - \left( \frac{G(x; \xi)}{\bar{G}(x; \xi)} \right)^2 \right] \right\}^{\theta} \right) \right].$$

### 3. Linear representation

In this section, we provide a very useful linear representation for the BX-G density function. If  $|z| < 1$  and  $b > 0$  is a real non-integer, the power series holds

$$(1 - z)^{b-1} = \sum_{i=0}^{\infty} \frac{(-1)^i \Gamma(b)}{i! \Gamma(b-i)} z^i. \quad (5)$$

For simplicity, ignoring the dependence of  $G(x)$  and  $g(x)$  on  $\xi$  and applying (5) to (4) we have

$$f(x) = \frac{2\theta g(x)G(x)}{\bar{G}(x)^3} \sum_{i=0}^{\infty} \frac{(-1)^i \Gamma(\theta)}{i! \Gamma(\theta-i)} \exp \left[ - (i+1) \left( \frac{G(x; \xi)}{\bar{G}(x; \xi)} \right)^2 \right]. \quad (6)$$

Applying the power series to the term  $\exp \left[ - (i+1) \left( \frac{G(x; \xi)}{\bar{G}(x; \xi)} \right)^2 \right]$ , Equation (6) becomes

$$f(x) = 2\theta g(x) \sum_{i,j=0}^{\infty} \frac{(-1)^{i+j} (i+1)^j \Gamma(\theta)}{i! j! \Gamma(\theta-i)} \frac{G(x)^{2j+1}}{\bar{G}(x)^{2j+3}}. \quad (7)$$

Consider the series expansion

$$(1 - z)^{-b} = \sum_{k=0}^{\infty} \frac{\Gamma(b+k)}{k! \Gamma(b)} z^k, \quad |z| < 1, \quad b > 0. \quad (8)$$

Applying the expansion in (8) to (7) for the term  $\overline{G}(x; \xi)^{2j+3}$ , Equation (7) becomes

$$f(x) = 2\theta \sum_{i,j,k=0}^{\infty} \frac{(-1)^{i+j} (i+1)^j \Gamma(\theta) \Gamma(2j+k+3) [2j+k+2]}{i! j! k! \Gamma(\theta-i) \Gamma(2j+3) [2j+k+2]} g(x) G(x)^{2j+k+1}.$$

This can be written as

$$f(x) = \sum_{j,k=0}^{\infty} \Omega_{j,k} \pi_{2j+k+2}(x), \tag{9}$$

where

$$\Omega_{j,k} = \frac{2\theta (-1)^j \Gamma(\theta) \Gamma(2j+k+3)}{j! k! \Gamma(2j+3) (2j+k+2)} \sum_{i=0}^{\infty} \frac{(-1)^i (i+1)^j}{i! \Gamma(\theta-i)}$$

and  $\pi_{2j+k+2}(x) = (2j+k+2) g(x) G(x)^{2j+k+1}$ . Equation (9) reveals that the density of  $X$  can be expressed as a linear mixture of exp-G densities. So, several mathematical properties of the new family can be obtained by knowing those of the exp-G distribution. Similarly, the cdf of the BX-G family can also be expressed as a mixture of exp-G cdfs given by

$$F(x) = \sum_{j,k=0}^{\infty} \Omega_{j,k} \Pi_{2j+k+2}(x) \tag{10}$$

where  $\Pi_{2j+k+2}(x)$  is the cdf of the exp-G family with power parameter  $(2j+k+2)$ .

#### 4. Mathematical and statistical properties

In this section we will provide some mathematical properties of the BX-G distribution.

##### 4.1. Probability weighted moments

The probability weighted moment (PWM)s are expectations of certain functions of a random variable and they can be defined for any random variable whose ordinary moments exist. The PWM method can generally be used for estimating parameters of a distribution whose inverse form cannot be expressed explicitly.

The  $(s, r)$ th PWM of  $X$  following the Burr type X generator, say  $\rho_{s,r}$ , is formally defined by

$$\rho_{s,r} = E \{X^s F(X)^r\} = \int_{-\infty}^{\infty} x^s F(x)^r f(x) dx.$$

Using equations (3), (4), (9) and (10) we can write

$$f(x) F(x)^r = \sum_{j,k=0}^{\infty} \Delta_{j,k} \pi_{2j+k+2}(x),$$

where

$$\Delta_{j,k} = \frac{2\theta (-1)^j \Gamma(2j+k+3)}{j! k! \Gamma(2j+3) (2j+k+2)} \sum_{i=0}^{\infty} (-1)^i (i+1)^j \binom{\theta(r+1)-1}{i}.$$

Then, the  $(s, r)$ th PWM of  $X$  can be expressed as

$$\rho_{s,r} = \sum_{j,k=0}^{\infty} \Delta_{j,k} E(Y_{2j+k+2}^s) dx.$$

#### 4.2. Residual and reversed residual life

The  $n$ th moment of the residual life, say  $m_n(t) = E[(X-t)^n | X > t]$ ,  $n = 1, 2, \dots$ , uniquely determine  $F(x)$ . The  $n$ th moment of the residual life of  $X$  is given by

$$m_n(t) = \frac{1}{R(t)} \int_t^{\infty} (x-t)^n dF(x).$$

Therefore,

$$m_n(t) = \frac{1}{R(t)} \sum_{j,k=0}^{\infty} \Omega_{j,k} \sum_{r=0}^n \binom{n}{r} (-t)^{n-r} \int_t^{\infty} x^r \pi_{2j+k+2}(x).$$

Another interesting function is the mean residual life (MRL) function or the life expectation at age  $t$  defined by  $m_1(t) = E[(X-t) | X > t]$ , which represents the expected additional life length for a unit which is alive at age  $t$ . The MRL of  $X$  can be obtained by setting  $n = 1$  in the last equation.

The  $n$ th moment of the reversed residual life, say  $M_n(t) = E[(t-X)^n | X \leq t]$  for  $t > 0$  and  $n = 1, 2, \dots$  uniquely determines  $F(x)$ . We obtain

$$M_n(t) = \frac{1}{F(t)} \int_0^t (t-x)^n dF(x).$$

Then, the  $n$ th moment of the reversed residual life of  $X$  becomes

$$M_n(t) = \frac{1}{F(t)} \sum_{j,k=0}^{\infty} \Omega_{j,k} \sum_{r=0}^n (-1)^r \binom{n}{r} t^{n-r} \int_0^t x^r \pi_{2j+k+2}(x).$$

The mean inactivity time (MIT) or mean waiting time (MWT) also called the mean reversed residual life function, is given by  $M_1(t) = E[(t-X) | X \leq t]$ , and it represents the waiting time elapsed since the failure of an item on condition that this failure had occurred in  $(0, t)$ . The MIT of the Burr type X generator of distributions can be obtained easily by setting  $n = 1$  in the above equation.

#### 4.3. Stress-strength model

Stress-strength model is the most widely approach used for reliability estimation. This model is used in many applications of physics and engineering such as strength failure and system collapse. In stress-strength modeling,  $\mathbf{R} = \Pr(X_2 < X_1)$  is a measure of reliability of the system when it is subjected to random stress  $X_2$  and has strength  $X_1$ . The system fails if and only if the applied stress is greater than its strength and the component will function satisfactorily whenever  $X_1 > X_2$ .  $\mathbf{R}$  can be considered as a measure of system performance and naturally arise in electrical and electronic systems. Other interpretation can be given as the reliability  $\mathbf{R}$  of a system is the probability that the system is strong enough to overcome the stress imposed on it.

Let  $X_1$  and  $X_2$  be two independent random variables with  $BX-G(\theta_1, \xi)$  and  $BX-G(\theta_2, \xi)$  distributions, respectively. The pdf of  $X_1$  and the cdf of  $X_2$  can be written from Equations (10) and (9), respectively as

$$f_1(x; \theta_1, \xi) = f(x) = 2\theta_1 \sum_{i,j,k=0}^{\infty} \frac{(-1)^{i+j} (i+1)^j \Gamma(\theta_1) \Gamma(2j+k+3)}{i! j! k! \Gamma(\theta_1 - i) \Gamma(2j+3)} g(x) G(x)^{2j+k+1}$$

and

$$F_2(x; \theta_2, \xi) = 2\theta_2 \sum_{h,w,m=0}^{\infty} \frac{(-1)^{h+w} (h+1)^w \Gamma(\theta_2) \Gamma(2w+m+3)}{h! w! m! \Gamma(\theta_2 - h) \Gamma(2w+3) (2w+m+2)} G(x)^{2w+m+2}.$$

Then, the reliability is defined by

$$\mathbf{R} = \int_0^{\infty} f_1(x; \theta_1, \xi) F_2(x; \theta_2, \xi) dx.$$

We can write

$$\mathbf{R} = \sum_{j,k,w,m=0}^{\infty} \Psi_{j,k,w,m} \int_0^{\infty} \pi_{2j+2w+k+m+4}(x) dx,$$

where

$$\begin{aligned} \Psi_{j,k,w,m} &= 4\theta_1 \theta_2 \sum_{j,k,w,m=0}^{\infty} \frac{(-1)^{j+w} \Gamma(2j+k+3) \Gamma(2w+m+3)}{j! k! w! m! \Gamma(\theta_2 - h) \Gamma(2j+3) \Gamma(2w+3)} \\ &\sum_{i,h=0}^{\infty} \frac{(-1)^{i+h} (i+1)^j (h+1)^w \binom{\theta_1-1}{i} \binom{\theta_2-1}{h}}{(2w+m+2) (2j+k+2w+m+4)}. \end{aligned}$$

and  $\pi_{2j+2w+k+m+4}(x) = (2j+k+2w+m+4) g(x) G(x)^{2j+k+2w+m+3}$ . Thus, the reliability,  $\mathbf{R}$ , can be expressed as

$$\mathbf{R} = \sum_{j,k,w,m=0}^{\infty} \Psi_{j,k,w,m} E(Y_{2j+2w+k+m+4}),$$

#### 4.4. Order statistics

Order statistics make their appearance in many areas of statistical theory and practice. Let  $X_1, \dots, X_n$  be a random sample from the  $BX-G$  of distributions and let  $X_{(1)}, \dots, X_{(n)}$  be the corresponding order statistics. The pdf of  $i$ th order statistic, say  $X_{i:n}$ , can be written as

$$f_{i:n}(x) = \frac{f(x)}{B(i, n-i+1)} \sum_{j=0}^{n-i} (-1)^j \binom{n-i}{j} F^{j+i-1}(x), \tag{11}$$

where  $B(\cdot, \cdot)$  is the beta function.

Using (3), (4), (9) and (10) in equation (11) we get

$$f(x) F(x)^{j+i-1} = \sum_{w,k=0}^{\infty} \mathbf{b}_{w,k} \pi_{2w+k+2}(x),$$

where

$$\mathbf{b}_{w,k} = \frac{2\theta(-1)^w \Gamma(2w+k+3)}{w!k! \Gamma(2w+3) (2w+k+2)} \sum_{m=0}^{\infty} (-1)^m (m+1)^w \binom{\theta(j+i)-1}{m}.$$

The pdf of  $X_{i:n}$  can be expressed as

$$f_{i:n}(x) = \sum_{w,k=0}^{\infty} \sum_{j=0}^{n-i} \frac{(-1)^j \binom{n-i}{j} \mathbf{b}_{w,k}}{\mathbf{B}(i, n-i+1)} \pi_{2w+k+2}(x).$$

Then, the density function of the BX-G order statistics is a mixture of exp-G densities. Based on the last equation, we note that the properties of  $X_{i:n}$  follow from those properties of  $Y_{2w+k+2}$ . For example, the moments of  $X_{i:n}$  can be expressed as

$$E(X_{i:n}^q) = \sum_{w,k=0}^{\infty} \sum_{j=0}^{n-i} \frac{(-1)^j \binom{n-i}{j} \mathbf{b}_{w,k}}{\mathbf{B}(i, n-i+1)} E(Y_{2w+k+2}^q). \tag{12}$$

#### 4.5. Entropies

The Rényi entropy of a random variable  $X$  represents a measure of variation of the uncertainty. The Rényi entropy is defined by

$$I_{\delta}(X) = \frac{1}{1-\delta} \log \int_{-\infty}^{\infty} f(x)^{\delta} dx, \quad \delta > 0 \text{ and } \delta \neq 1.$$

Using the pdf (4), we can write

$$f(x)^{\delta} = \sum_{j,k=0}^{\infty} \mathbf{t}_{j,k} g(x; \xi)^{\delta} G(x; \xi)^{\delta+2j+k},$$

where

$$\mathbf{t}_{j,k} = \frac{(-1)^j \Gamma(3\delta+2j+k)}{j!k! \Gamma(3\delta+2j)} \sum_{i=0}^{\infty} \frac{2^{\delta} \theta^{\delta} (-1)^i (\theta\delta - \delta)_i}{i! (\delta+i)^{-j}}.$$

Then, the Rényi entropy of the BX-G is given by

$$I_{\delta}(X) = \frac{1}{1-\delta} \log \left\{ \sum_{j,k=0}^{\infty} \mathbf{t}_{j,k} \int_{-\infty}^{\infty} g(x; \xi)^{\delta} G(x; \xi)^{2j+k+\delta} dx \right\},$$

The  $\delta$ -entropy, say  $H_{\delta}(X)$ , can be obtained as

$$H_{\delta}(X) = \frac{1}{\delta - 1} \log \left\{ 1 - \left[ \sum_{j,k=0}^{\infty} \mathbf{t}_{j,k} \int_{-\infty}^{\infty} g(x; \xi)^q G(x; \xi)^{2j+k+q} dx \right] \right\}.$$

The Shannon entropy of a random variable  $X$ , say  $SI$ , is defined by

$$SI = E \{ -[\log f(X)] \},$$

follows by taking the limit of  $I_{\delta}(X)$  as  $\delta$  tends to 1.

## 5. Characterizations

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

### 5.1. Characterizations based on two truncated moments

In this subsection we present characterizations of BX-G distribution in terms of a simple relationship between two truncated moments. Our first characterization result employs a theorem due to Glänzel, see Theorem A.1 of Appendix A. 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 Glänzel [17], this characterization is stable in the sense of weak convergence.

**Proposition 5.1.** *Let  $X : \Omega \rightarrow (0, \infty)$  be a continuous random variable and let  $q_1(x) = \left\{ 1 - \exp \left[ - \left( \frac{G(x; \xi)}{\overline{G}(x)} \right)^2 \right] \right\}^{1-\theta}$  and  $q_2(x) = q_1(x) \exp \left[ - \left( \frac{G(x; \xi)}{\overline{G}(x)} \right)^2 \right]$  for  $x > 0$ . The random variable  $X$  belongs to BX-G family (4) if and only if the function  $\eta$  defined in Theorem A.1 has the form*

$$\eta(x) = \frac{1}{2} \exp \left[ - \left( \frac{G(x; \xi)}{\overline{G}(x)} \right)^2 \right], \quad x > 0. \tag{13}$$

*Proof.* Let  $X$  be a random variable with pdf (4), then

$$(1 - F(x)) E [q_1(x) | X \geq x] = \exp \left[ - \left( \frac{G(x; \xi)}{\overline{G}(x; \xi)} \right)^2 \right], \quad x > 0,$$

and

$$(1 - F(x)) E [q_2(x) | X \geq x] = \frac{1}{2} \exp \left[ - \left( \frac{G(x; \xi)}{\overline{G}(x; \xi)} \right)^2 \right], \quad x > 0,$$

and finally



$$\eta(x) q_1(x) - q_2(x) = -\frac{1}{2} q_1(x) \left\{ -\exp \left[ -\left( \frac{G(x; \xi)}{\bar{G}(x; \xi)} \right)^2 \right] \right\} < 0 \text{ for } x > 0.$$

Conversely, if  $\eta$  is given as above, then

$$s'(x) = \frac{\eta'(x) q_1(x)}{\eta(x) q_1(x) - q_2(x)} = \frac{2g(x; \xi) G(x; \xi)}{(\bar{G}(x; \xi))^3}, \quad x > 0,$$

and hence

$$s(x) = \left( \frac{G(x; \xi)}{\bar{G}(x; \xi)} \right)^2, \quad x > 0.$$

Now, in view of Theorem A.1,  $X$  has density (4). □

**Corollary 5.1.** *Let  $X : \Omega \rightarrow (0, \infty)$  be a continuous random variable and let  $q_1(x)$  be as in Proposition 5.1. The pdf of  $X$  is (4) if and only if there exist functions  $q_2$  and  $\eta$  defined in Theorem A.1 satisfying the differential equation*

$$\frac{\eta'(x) q_1(x)}{\eta(x) q_1(x) - q_2(x)} = \frac{2g(x; \xi) G(x; \xi)}{(\bar{G}(x; \xi))^3}, \quad x > 0. \tag{14}$$

The general solution of the differential equation in Corollary 5.1 is

$$\eta(x) = \exp \left[ -\left( \frac{G(x; \xi)}{\bar{G}(x; \xi)} \right)^2 \right] \left[ -\int \frac{2g(x; \xi) G(x; \xi)}{(\bar{G}(x; \xi))^3} \exp \left[ \left( \frac{G(x; \xi)}{\bar{G}(x; \xi)} \right)^2 \right] \times \right. \\ \left. (q_1(x))^{-1} q_2(x) dx + D, \right]$$

where  $D$  is a constant. Note that a set of functions satisfying the differential equation (14) is given in Proposition 5.1 with  $D = 0$ . However, it should be also noted that there are other triplets  $(q_1, q_2, \eta)$  satisfying the conditions of Theorem A.1.

### 5.2. 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

$$\frac{f'(x)}{f(x)} = \frac{h'_F(x)}{h_F(x)} - h_F(x). \tag{15}$$

For many univariate continuous distributions, this is the only characterization available in terms of the hazard function. The following characterization establishes a non-trivial characterization for BX-G distribution in terms of the hazard function when  $\theta = 1$ , which is not of the trivial form given in (15). Clearly, we assume that  $G(x; \xi)$  is twice differentiable.

**Proposition 5.2.** Let  $X : \Omega \rightarrow (0, \infty)$  be a continuous random variable. Then for  $\theta = 1$ , the pdf of  $X$  is (4) if and only if its hazard function  $h_F(x)$  satisfies the differential equation

$$h'_F(x) - \frac{g'(x)}{g(x)} h_F(x) = \frac{2g(x; \xi)(1 + 2G(x; \xi))}{(\bar{G}(x; \xi))^4}, \quad (16)$$

with the boundary condition  $h_F(0) = 0$ .

*Proof.* If  $X$  has pdf (4), then clearly (16) holds. Now, if (16) holds, then

$$\frac{d}{dx} \left\{ (g(x))^{-1} h_F(x) \right\} = 2 \frac{d}{dx} \left\{ \frac{G(x; \xi)}{(\bar{G}(x; \xi))^3} \right\},$$

or, equivalently,

$$h_F(x) = \frac{2g(x; \xi) G(x; \xi)}{(\bar{G}(x; \xi))^3},$$

which is the hazard function of the BX-G distribution. □

### 5.3. Characterization based on truncated moment of certain function of the random variable

The following propositions have already appeared in (Hamedani, Technical Report, [19]), so we will just state them here which can be used to characterize BX-G distribution.

**Proposition 5.3.** Let  $X : \Omega \rightarrow (a, b)$  be a continuous random variable with cdf  $F$ . Let  $\psi(x)$  be a differentiable function on  $(a, b)$  with  $\lim_{x \rightarrow a^+} \psi(x) = 1$ . Then for  $\delta \neq 1$ ,

$$E[\psi(X) | X \geq x] = \delta \psi(x), \quad x \in (a, b),$$

if and only if

$$\psi(x) = (1 - F(x))^{\frac{1}{\delta} - 1}, \quad x \in (a, b).$$

**Proposition 5.4.** Let  $X : \Omega \rightarrow (a, b)$  be a continuous random variable with cdf  $F$ . Let  $\psi_1(x)$  be a differentiable function on  $(a, b)$  with  $\lim_{x \rightarrow b^-} \psi_1(x) = 1$ . Then for  $\delta_1 \neq 1$ ,

$$E[\psi_1(X) | X \leq x] = \delta_1 \psi_1(x), \quad x \in (a, b),$$

if and only if

$$\psi_1(x) = (F(x))^{\frac{1}{\delta_1}-1}, \quad x \in (a, b).$$

**Remark.** It is easy to see that for certain functions  $\psi(x)$  and  $\psi_1(x)$  on  $(0, \infty)$ ; (a) Proposition 5.3 provides a characterization of BX-G distribution for  $\theta = 1$  and (b) Proposition 5.4 provides a characterization of BX-G distribution.

### 6. Special BX-G models

In this section, we provide two special cases of the BX-G family of distributions. The pdf (4) will be most tractable when  $G(x; \xi)$  and  $g(x; \xi)$  have simple analytic expressions. These special models generalize some well-known distributions in the literature.

#### 6.1. The Burr X-Weibull (BXW) distribution

Consider the cdf and pdf (for  $x > 0$ )  $G(x) = 1 - \exp[-(\alpha x)^\beta]$  and  $g(x) = \beta \alpha^\beta x^{\beta-1} \exp[-(\alpha x)^\beta]$ , respectively, of the Weibull distribution with positive parameters  $\alpha$  and  $\beta$ . Then, the pdf of the BXW model is given by

$$f(x) = 2\theta\beta\alpha^\beta x^{\beta-1} \exp\left\{2(\alpha x)^\beta - \left[\exp(\alpha x)^\beta - 1\right]^2\right\} \times \left\{1 - \exp\left[-(\alpha x)^\beta\right]\right\} \left(1 - \exp\left\{-\left[\exp(\alpha x)^\beta - 1\right]^2\right\}\right)^{\theta-1}.$$

The BXW distribution includes the Burr X-Rayleigh (BXR) distribution when  $\beta = 2$ . For  $\beta = 1$ , we have the Burr X-exponential (BXE) distribution. The plots of the BXW density for some parameter values are displayed in Figure 1.

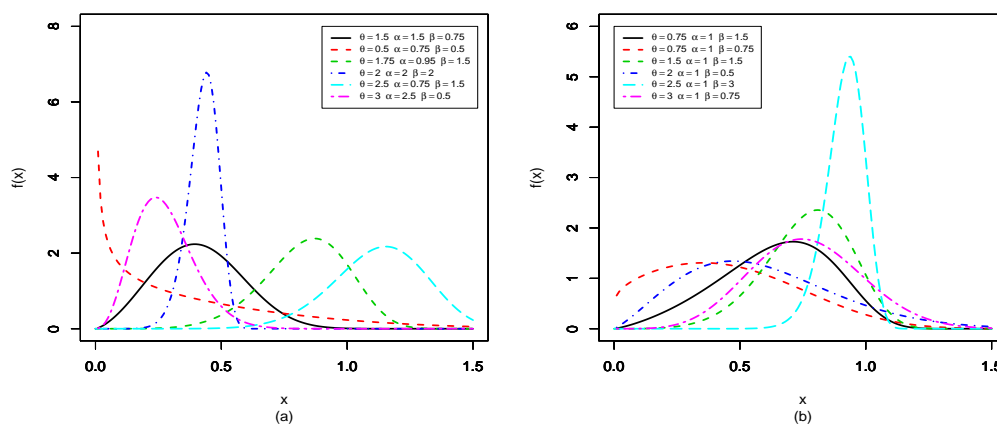


Fig. 1. pdf of BXW distribution for different values of parameters

**6.2. The Burr X-Lomax (BXL) distribution**

The cdf and pdf (for  $x > 0$ ) of the Lomax distribution with positive parameters  $\alpha$  and  $\beta$  are  $G(x) = 1 - [1 + (x/\beta)]^{-\alpha}$  and  $g(x) = (\alpha/\beta)[1 + (x/\beta)]^{-\alpha-1}$ , respectively. Then, the cdf and pdf of the BXL distribution becomes

$$F(x) = \left\{ 1 - \exp \left[ - \left( \left[ 1 + \frac{x}{\beta} \right]^\alpha - 1 \right)^2 \right] \right\}^\theta,$$

$$f(x) = \frac{2\theta\alpha}{\beta} \left[ 1 + \frac{x}{\beta} \right]^{\alpha-1} \left( \left[ 1 + \frac{x}{\beta} \right]^\alpha - 1 \right) \exp \left\{ - \left( \left[ 1 + \frac{x}{\beta} \right]^\alpha - 1 \right)^2 \right\} \left\{ 1 - \exp \left[ - \left( \left[ 1 + \frac{x}{\beta} \right]^\alpha - 1 \right)^2 \right] \right\}^{\theta-1}.$$

The plots of the BXL density are displayed in Figure 2 for some parameter values.

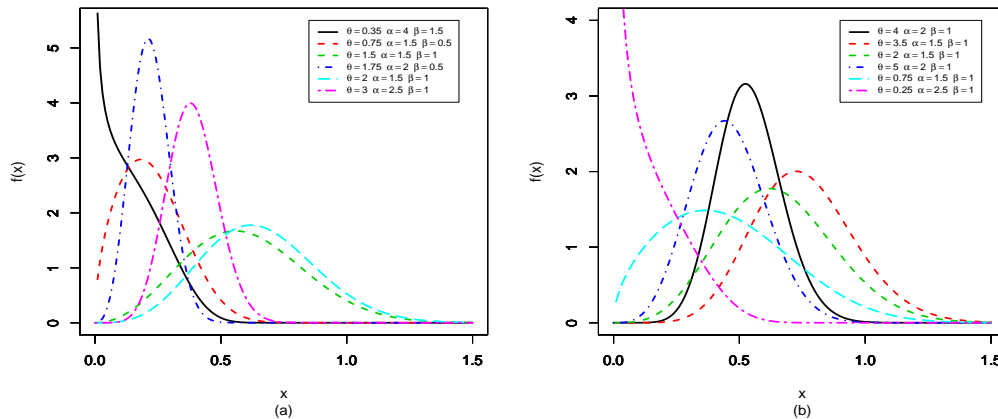


Fig. 2. pdf of BXL distribution for different values of parameters

**7. Parameter Estimation**

Several approaches for parameter estimation were proposed in the literature but the maximum likelihood method is the most commonly employed. So, we consider the estimation of the unknown parameters of this family from complete samples only by maximum likelihood.

Let  $x_1, \dots, x_n$  be a random sample from the BX-G family with parameters  $\theta$  and  $\xi$ . Let  $\Theta = (\theta, \xi^T)^T$  be the  $p \times 1$  parameter vector. For determining the MLE of  $\Theta$ , we have the log-likelihood function

$$\ell = \ell(\Theta) = n \log 2 + n \log \theta + \sum_{i=1}^n \log g(x_i; \xi) + \sum_{i=1}^n \log G(x_i; \xi) - 3 \sum_{i=1}^n \log \bar{G}(x_i; \xi) - \sum_{i=1}^n s_i^2 + (\theta - 1) \sum_{i=1}^n \log [1 - \exp(-s_i^2)],$$

where  $s_i = G(x_i; \xi) / \bar{G}(x_i; \xi)$ .

The components of the score vector,  $\mathbf{U}(\Theta) = \frac{\partial \ell}{\partial \Theta} = \left( \frac{\partial \ell}{\partial \theta}, \frac{\partial \ell}{\partial \xi} \right)^\top$ , are

$$U_\theta = \frac{n}{\theta} + \sum_{i=1}^n \log [1 - \exp(-s_i^2)]$$

and

$$U_\xi = \sum_{i=1}^n \frac{g'(x_i; \xi)}{g(x_i; \xi)} + \sum_{i=1}^n \frac{G'(x_i; \xi)}{G(x_i; \xi)} + 3 \sum_{i=1}^n \frac{G'(x_i; \xi)}{G(x_i; \xi)} - \sum_{i=1}^n 2s_i G'(x_i; \xi) / \overline{G}(x_i; \xi)^2 + 2(\theta - 1) \sum_{i=1}^n \frac{p_i s_i \exp(-s_i^2)}{1 - \exp(-s_i^2)},$$

where  $g'(x_i; \xi) = \partial g(x_i; \xi) / \partial \xi$ ,  $G'(x_i; \xi) = \partial G(x_i; \xi) / \partial \xi$  and  $p_i = \frac{\partial G(x_i; \xi) / \overline{G}(x_i; \xi)}{\partial \xi}$

Setting the nonlinear system of equations  $U_\theta = 0$  and  $U_\xi = \mathbf{0}$  and solving them simultaneously yields the MLE  $\widehat{\Theta} = (\widehat{\theta}, \widehat{\xi}^\top)^\top$ . To solve these equations, it is usually more convenient to use nonlinear optimization methods such as the quasi-Newton algorithm to numerically maximize  $\ell$ .

### 8. Simulation study

In this section, we present some simulations results for different sample sizes to assess the accuracy of the MLEs. For illustrative purposes, we will choose the BXL distribution. An ideal technique for simulating from the BXL distribution is the inversion method. We can simulate  $X$  by

$$X = \beta \left[ \left\{ 1 + \left[ -\log(1 - U^{1/\theta}) \right]^{1/2} \right\}^{1/\alpha} - 1 \right],$$

where  $U$  is a uniform random number in  $(0, 1)$ . For selected combinations of  $\alpha, \beta$  and  $\theta$ , we generate samples of different sizes from the BXL distribution. We repeat the simulations 1,000 times and evaluate the mean estimates and the root mean square errors (RMSEs). The required computations use a script *AdequacyModel* of the R-package written by Marinho *et al.* [21]. Empirical results obtained for selected values of triplets  $(\alpha, \beta, \theta)$  are given in Tables 1 and 2. One can observe that our estimates are pretty stable and as the sample size increases the mean square error decreases. Therefore, the maximum likelihood method works very well to estimate the model parameters of the BXL distribution.

### 9. Application to cancer patients data

In this section, we provide an application to real data to illustrate the flexibility of the BXL model presented in Section 6. The goodness-of-fit statistics for these models are compared with other competitive models and the MLEs of the model parameters are determined. The data set refers to the remission times (in months) of a random sample of 128 bladder cancer patients studied by Lee and Wang [20]. For these data, we compare the fit of the BXL distribution with those of the L, transmuted linear exponential (TLE), transmuted additive Weibull (TAW), BX and exponentiated transmuted generalized Rayleigh (ETGR) models ( $x > 0$  for all of them).

Table 1. Empirical means and the RMSEs of the BXL distribution for  $\alpha = 1, \beta = 2$  and  $\theta = 1.5$

$n$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\theta}$
10	1.16551 (0.42087)	2.34014 (1.19099)	1.97799 (1.10104)
20	1.14729 (0.36093)	2.38295 (1.01954)	1.66762 (0.533004)
50	1.04212 (0.23805)	2.12087 (0.75753)	1.57064 (0.31329)
100	1.06694 (0.17699)	2.20435 (0.56757)	1.51423 (0.20657)
200	0.98302 (0.08605)	1.93524 (0.26025)	1.54343 (0.17740)

Table 2. Empirical means and the RMSEs of the BXL distribution for  $\alpha = 1.5, \beta = 2$  and  $\theta = 3$

$n$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\theta}$
10	1.68605 (0.53095)	2.25143 (0.83961)	3.67227 (1.48749)
20	1.69259 (0.44873)	2.29826 (0.76963)	3.29538 (0.97611)
50	1.54496 (0.33149)	2.06713 (0.61679)	3.19786 (0.69367)
100	1.57582 (0.26203)	2.13539 (0.48819)	3.05440 (0.43556)
200	1.50372 (0.07566)	2.03331 (0.11298)	3.01899 (0.20671)

The TLE density (Tian et al., [25]) given by

$$f(x) = (\alpha + \gamma x) \left[ 1 - e^{-(\alpha x + \frac{\gamma}{2} x^2)} \right] \left\{ 1 - \lambda + 2\lambda e^{-(\alpha x + \frac{\gamma}{2} x^2)} \right\}.$$

The TAW density (Elbatal and Aryal, [13]) given by

$$f(x) = e^{-(\alpha x^\theta + \gamma x^\beta)} \left( \alpha \theta x^{\theta-1} + \gamma \beta x^{\beta-1} \right) \left\{ 1 - \lambda + 2\lambda e^{-(\alpha x^\theta + \gamma x^\beta)} \right\}.$$

The ETGR density (Afify et al., [3]) given by

$$f(x) = 2\alpha\delta\beta^2 x e^{-(\beta x)^2} \left( 1 - e^{-(\beta x)^2} \right)^{\alpha\delta-1} \times \left[ 1 + \lambda - 2\lambda \left( 1 - e^{-(\beta x)^2} \right)^\alpha \right] \left\{ 1 + \lambda - \lambda \left( 1 - e^{-(\beta x)^2} \right)^\alpha \right\}^{\delta-1}.$$

The parameters of the above densities are all positive real numbers except the parameter  $\lambda$  where  $|\lambda| \leq 1$ .

Table 3. The Goodness-of-fit criteria for the cancer data

Model	$-2\hat{\ell}$	<i>AIC</i>	<i>CAIC</i>	<i>HQIC</i>	<i>BIC</i>	$W^*$	$A^*$
BXL	822.277	828.277	828.47	831.753	836.833	0.05807	0.37468
L	827.644	831.644	831.74	833.962	837.348	0.05905	0.55862
TLE	826.971	832.971	833.165	836.448	841.528	0.06085	0.55402
TAW	828.478	838.478	838.97	844.272	852.739	0.11288	0.70326
BX	858.431	862.431	862.527	864.748	868.135	0.39647	2.35125
ETGR	858.35	866.35	866.675	870.985	877.758	0.39794	2.36077

Table 4. MLEs and their standard errors (in parentheses) for the cancer data

Model	Estimates				
TAW	$\hat{\alpha} = 0.1139$ (0.032)	$\hat{\beta} = 0.9722$ (0.125)	$\hat{\gamma} = 3.0936 \cdot 10^{-5}$ (0.006106)	$\hat{\theta} = 1.0065$ (0.035)	$\hat{\lambda} = -0.163$ (0.28)
ETGR	$\hat{\alpha} = 7.3762$ (5.389)	$\hat{\lambda} = 0.118$ (0.26)	$\hat{\beta} = 0.0473$ (0.003965)	$\hat{\delta} = 0.0494$ (0.036)	
BXL	$\hat{\alpha} = 0.2982$ (0.051)	$\hat{\beta} = 1.0194$ (0.664)	$\hat{\theta} = 0.9337$ (0.25)		
TLE	$\hat{\alpha} = 0.0612$ (0.01)	$\hat{\lambda} = 0.8568$ (0.203)	$\hat{\gamma} = 3.0877 \cdot 10^{-5}$ (0.0006819)		
L	$\hat{\alpha} = 13.9187$ (15.3472)	$\hat{\beta} = 120.8281$ (142.3413)			
BX	$\hat{\alpha} = 0.364$ (0.037)	$\hat{\beta} = 0.0476$ (0.00.907)			

In order to compare the fitted models, we consider some goodness-of-fit measures including the Akaike information criterion (*AIC*), consistent Akaike information criterion (*CAIC*), Hannan-Quinn information criterion (*HQIC*), Bayesian information criterion (*BIC*) and  $-2\hat{\ell}$ , where  $\hat{\ell}$  is the maximized log-likelihood. Further, we adopt the Anderson-Darling ( $A^*$ ) and Cramér-von Mises ( $W^*$ ) statistics in order to compare the fits of the two new models with other nested and non-nested models. The statistics are widely used to determine how closely a specific cdf fits the empirical distribution of a given data set. The smaller these statistics are, the better the fit.

Table 3 lists the values of  $-2\hat{\ell}$ , *AIC*, *CAIC*, *HQIC*, *BIC*,  $W^*$  and  $A^*$ , whereas the MLEs and their corresponding standard errors (in parentheses) of the model parameters are given in Table 4. In Table 3, we compare the fits of the BXL model with the L, TLE, TAW, BX and ETGR models. We note that the BXL model has the lowest values for the  $-2\hat{\ell}$ , *AIC*, *CAIC*, *HQIC*, *BIC*,  $W^*$  and  $A^*$  statistics (for the cancer data) among the fitted models. So, the BXL model could be chosen as the best model. Figures 3 and 4 display the fitted pdfs and Q-Q plots for the BXL distribution and other distributions. They reveal that the new distribution can be a better model for the data than other competitive lifetime models.

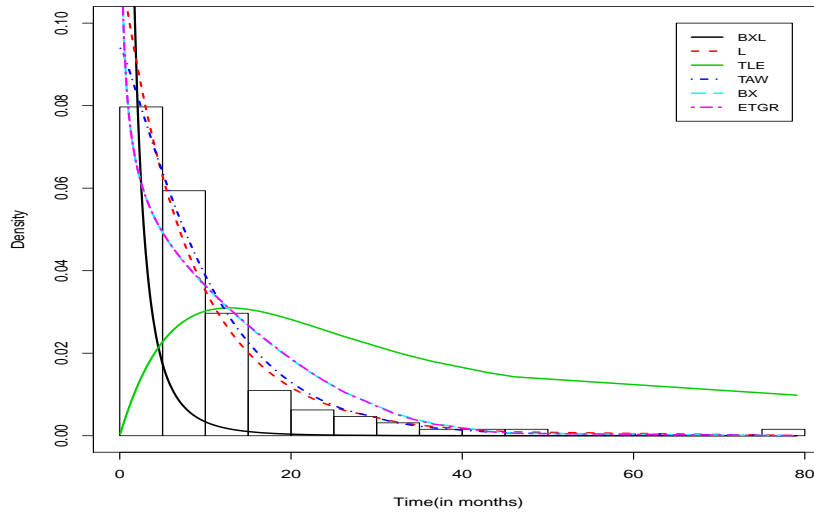


Fig. 3. Estimated pdfs and of the BXL model and competing models for the cancer data

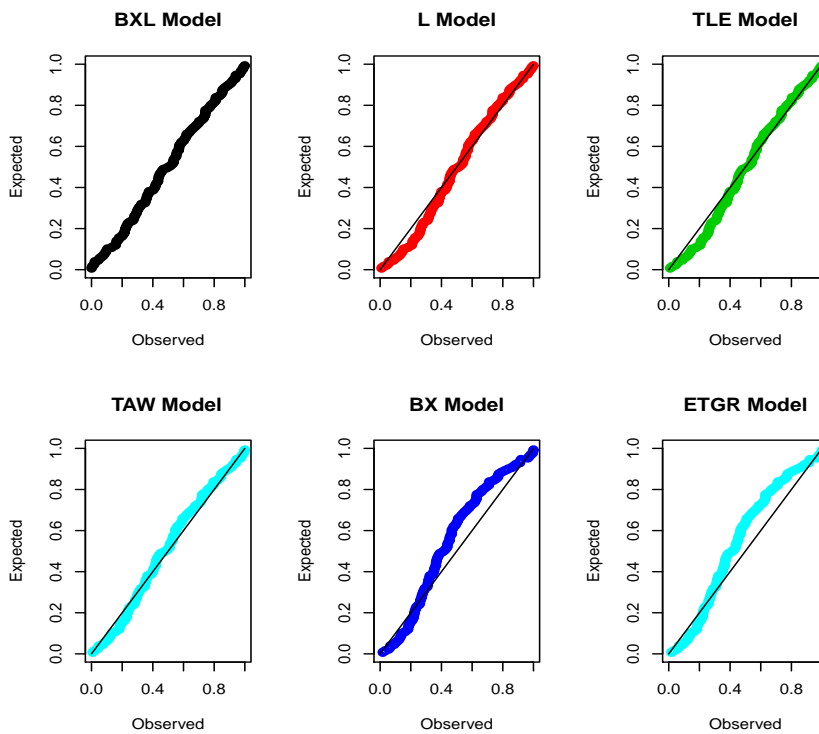


Fig. 4. Q-Q plots of the BXL model and its competing models for the cancer data

## 10. Conclusions

The idea of generating new extended models from the classical ones has been of great interest among researchers in the past decade. We present a new *Burr X-G* (BX-G) family of distributions



by adding one extra shape parameter. Many well-known distributions emerge as special cases of the proposed family by taking integer parameter values. We provide some mathematical properties of the new family. To illustrate a usefulness of proposed distribution we have studied the remission times (in months) of a random sample of 128 bladder cancer patients. Although time to remission depends on many covariates we have shown that the remission time can be modeled using BXL distribution.

### Acknowledgements

The authors would like to thank the editor and anonymous reviewers for carefully reading the manuscript and making valuable suggestions.

### References

- [1] Afify A. Z., Alizadeh, M., Yousof, H. M., Aryal, G. and Ahmad, M. (2016). The transmuted geometric-G family of distributions: theory and applications. *Pak. J. Statist.*,32, 139-160.
- [2] Afify A. Z., Cordeiro, G. M., Yousof, H. M., Alzaatreh, A. and Nofal, Z. M. (2016). The Kumaraswamy transmuted-G family of distributions: properties and applications. *Journal of Data Science*,14,245-270.
- [3] Afify, A. Z., Nofal, Z. M. and Ebraheim, A. N.(2015) Exponentiated transmuted generalized Rayleigh distribution: a new four parameter Rayleigh distribution, *PPakistan Journal of Statistics and Operation Research* **11**, 115–134.
- [4] Alizadeh, M., Cordeiro, G. M., Brito, E., (2015). The beta Marshall-Olkin family of distributions. *Journal of Statistical Distributions and Applications*, 2:4, 18-page.
- [5] Alizadeh, M., Cordeiro, G.M., Mansoor, M., Zubair, M., and Hamedani, G.G. (2015). The Kumaraswamy Marshal-Olkin family of distributions. *Journal of the Egyptian Mathematical Society*, **23**, 546-557.
- [6] Alizadeh, M., Emadi, M., Doostparast, M., Cordeiro, G.M., Ortega, E.M.M., Pescim, R.R. (2015). Kumaraswamy odd log-logistic family of distributions: Properties and applications. *Hacet. J. Math. Stat.*, forthcoming.
- [7] Burr, I.W. (1942) Cumulative frequency functions, *Ann. Math. Statist.*, **18**, 215–232.
- [8] 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.
- [9] Cordeiro, G.M. and de Castro, M. (2011). A new family of generalized distributions. *Journal of Statistical Computation and Simulation*, **81**, 883-898.
- [10] Cordeiro, G. M., Hashimoto, E. M., and Ortega, E. M. (2014). McDonald Weibull model. *Statistics: A Journal of Theoretical and Applied Statistics*, **48**, 256-278.
- [11] Cordeiro, G. M., Alizadeh, M., Tahir, M. H., Mansoor, M., Bourguignon, M., and Hamedani, G.G. (2015). The beta odd log-logistic family of distributions. *Hacet. J. Math. Stat.*, forthcoming.
- [12] Cordeiro, G. M., Ortega, E. M., Popović B. V. and Pescim, R. R. (2014). The Lomax generator of distributions: Properties, minification process and regression model, **247**, 465–486.
- [13] Elbatal, I. and Aryal, G. (2013). On the transmuted additive Weibull distribution, *Austrian Journal of Statistics*, **42**(2), 117–132.
- [14] Eugene, N., Lee, C. and Famoye, F. (2002). Beta-normal distribution and its applications. *Communications in Statistics-Theory and Methods*, **31**, 497-512.
- [15] Flor de Santana, T. N., Ortega, E. M. M., Cordeiro, G. M., (2012). Kumaraswamy log-logistic distribution. *Journal of Statistical Theory and Applications*, **11**, 265-291.
- [16] Glänzel, W., (1987). A characterization theorem based on truncated moments and its application to some distribution families, *Mathematical Statistics and Probability Theory (Bad Tatzmannsdorf, 1986)*, Vol. B, Reidel, Dordrecht, 75-84.
- [17] Glänzel, W., (1990). Some consequences of a characterization theorem based on truncated moments, *Statistics: A Journal of Theoretical and Applied Statistics*, **21**, 613-618.

- [18] Gupta, R. C., Gupta, P. L. and Gupta, R. D. (1998). Modeling failure time data by Lehmann alternatives. *Commun. Stat. Theory Methods*, **27**, 887-904.
- [19] Hamedani, G. G. (2013), On certain generalized gamma convolution distributions II, Technical Report, No. 484, MSCS, Marquette University.
- [20] Lee, E. T. and Wang J. W. (2003). *Statistical Methods for Survival Data Analysis*, 3rd ed., Wiley, New York.
- [21] Marinho, P.R. D., Bourguignon, M. and Dias, C. R. B. *Adequacy of probabilistic models and generation of pseudo-random numbers*, R package-AdequacyModel, Available from: <http://cran.r-project.org/web/packages/AdequacyModel/AdequacyModel.pdf>
- [22] Marshall, A. W., Olkin, I. (1997). A new methods for adding a parameter to a family of distributions with application to the Exponential and Weibull families. *Biometrika*, **84**, 641-652.
- [23] Nadarajah, S. and Kotz, S. (2006). The exponentiated type distributions. *Acta Applicandae Mathematicae*, **92**, 97-111.
- [24] Nofal, Z. M., Afify, A. Z., Yousof, H. M. and Cordeiro, G. M. (2015). The generalized transmuted-G family of distributions. *Communications in Statistics-Theory and Methods*, forthcoming.
- [25] Tian, Y., Tian, M. and Zhu, Q. (2014) Transmuted Linear Exponential Distribution: A New Generalization of the Linear Exponential Distribution, *Communications in Statistics - Simulation and Computation*, **43**(10), 2661–2677.
- [26] Yousof, H. M., Afify, A. Z., Alizadeh, M., Butt, N. S., Hamedani, G. G., Ali, M. M. (2015). The transmuted exponentiated generalized-G family of distributions, *Pak. J. Stat. Oper. Res.*, **11**, 441-464.

## Appendix A.

**Theorem A.1.** *Let  $(\Omega, \mathcal{F}, \mathbf{P})$  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  $X : \Omega \rightarrow H$  be a continuous random variable with the distribution function  $F$  and let  $q_1$  and  $q_2$  be two real functions defined on  $H$  such that*

$$\mathbf{E}[q_2(X) \mid X \geq x] = \mathbf{E}[q_1(X) \mid X \geq x] \eta(x), \quad x \in H,$$

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

$$F(x) = \int_a^x C \left| \frac{\eta'(u)}{\eta(u)q_1(u) - q_2(u)} \right| \exp(-s(u)) du,$$

*where the function  $s$  is a solution of the differential equation  $s' = \frac{\eta' q_1}{\eta q_1 - q_2}$  and  $C$  is the normalization constant, such that  $\int_H dF = 1$ .*