Pak. J. Statist.
2017 Vol. 33(2), 95-116

THE BETA WEIBULL-G FAMILY OF DISTRIBUTIONS: THEORY, CHARACTERIZATIONS AND APPLICATIONS

Haitham M. Yousof[1], Mahdi Rasekhi[2], Ahmed Z. Afify[1], Indranil Ghosh[3], Morad Alizadeh[4] and G.G. Hamedani[5]

[1] Department of Statistics, Mathematics and Insurance, Benha University Egypt. Email: haitham.yousof@fcom.bu.edu.eg
ahmed.afify@fcom.bu.edu.eg

[2] Department of Statistics, Malayer University, Malayer, Iran. Email: rasekhimahdi@gmail.com

[3] Department of Mathematics and Statistics, University of North Carolina Wilmington, USA. Email: ghoshi@uncw.edu

[4] Department of Statistics, Faculty of Sciences, Persian Gulf University Bushehr, 75169, Iran. Email: moradalizadeh78@gmail.com

[5] Department of Mathematics, Statistics and Computer Science Marquette University, USA. Email: gholamhoss.hamedani@marquette.edu

ABSTRACT

We propose and study a new class of continuous distributions called the beta Weibull-G family which extends the Weibull-G family introduced by Bourguignon et al. (2014). Some of its mathematical properties including explicit expressions for the ordinary and incomplete moments, generating function, order statistics and probability weighted moments 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 two applications to real data sets.

KEY WORDS

Beta-G Family, Generating Function, Maximum Likelihood Estimation, Moments, Order Statistics, Weibull-G Family.

1. INTRODUCTION

As of late, there has been an extraordinary enthusiasm for growing more flexible distributions through extending the classical distributions by introducing additional shape parameters to the baseline model. Many generalized families of distributions have been proposed and studied over the last two decades for modeling data in many applied areas such as economics, engineering, biological studies, environmental sciences, medical sciences and finance. So, several classes of distributions have been constructed by extending common families of continuous distributions. These generalized distributions give more flexibility by adding one (or more) shape parameters to the baseline model. They were pioneered by Gupta et al. (1998) who proposed the exponentiated-G class, which consists of raising the cumulative distribution function (cdf) to a positive power parameter. Many other classes can be cited such as the Marshall-Olkin-G by Marshall and Olkin (1997), beta generalized-G by Eugene et al. (2002), Kumaraswamy-G by Cordeiro and de Castro (2011), exponentiated generalized-G by Cordeiro et al. (2013), T-X by Alzaatreh et al. (2013), Lomax-G by Cordeiro et al. (2014), beta Marshall-Olkin-G by Alizadeh et al. (2015), transmuted exponentiated generalized-G by Yousof et al. (2015), Kumaraswamy transmuted-G by Afify et al. (2016b), transmuted geometric-G by Afify et al. (2016a), Burr X-G by Yousof et al. (2016), odd-Burr generalized-G by Alizadeh et al. (2016a), transmuted Weibull G by Alizadeh et al. (2016b), exponentiated transmuted-G by Merovc et al. (2016), generalized transmuted-G by Nofal et al. (2017), exponentiated generalized-G Poisson by Aryal and Yousof (2017) and beta transmuted-H families by Afify et al. (2017), among others.

The produced families have pulled in numerous scientists and analysts to grow new models in light of the fact that the computational and diagnostic offices accessible in most symbolic computation software platforms. Several mathematical properties of the extended distributions may be easily explored using mixture forms of exponentiated-G (exp-G) distributions.

Let $g(x;\phi)$ and $G(x;\phi)$ denote the probability density function (pdf) and cdf of a baseline model with parameter vector ϕ and consider the Weibull cdf $V(t) = 1 - e^{-t^{\lambda}}$ (for $t > 0$) with positive parameter λ. Based on this density, Bourguignon et al. (2014) replaced the argument x by $G(x;\phi)/\overline{G}(x;\phi)$, where $\overline{G}(x;\phi) = 1 - G(x;\phi)$, and defined the cdf of their Weibull-G (W-G) class by

$$H(x;\lambda;\phi) = \int_0^{\frac{G(x;\phi)}{\overline{G}(x;\phi)}} \lambda\, t^{\lambda-1}\, e^{-t^{\lambda}} dt = 1 - \exp\left[-\left(\frac{G(x;\phi)}{\overline{G}(x;\phi)}\right)^{\lambda}\right]. \tag{1.1}$$

The W-G density function is given by

$$h(x;\lambda;\phi) = \lambda\, g(x;\phi) \frac{G(x;\phi)^{\lambda-1}}{\overline{G}(x;\phi)^{\lambda+1}} \exp\left[-\left(\frac{G(x;\phi)}{\overline{G}(x;\phi)}\right)^{\lambda}\right]. \tag{1.2}$$

In this paper, we define and study a new family of distributions by adding two extra shape parameters in equation (1.1) to provide more flexibility to the generated family. In fact, based on the beta-generalized (B-G) family proposed by Eugene et al. (2002), we construct a new family called the *beta Weibull-G* (BW-G) family and give a comprehensive description of some of its mathematical properties. In fact, the BW-G family provides better fits than at least four other families in two applications. We hope that the new family will attract wider applications in reliability, engineering and other areas of research. For an arbitrary baseline cdf $H(x;\phi)$, Eugene et al. (2002) defined the B-G family via the cdf and pdf

$$F(x;a,b;\phi) = I_{H(x)}(a,b) = \frac{1}{B(a,b)} \int_0^{H(x;\phi)} t^{a-1}(1-t)^{b-1} dt, \tag{1.3}$$

and

$$f(x;a,b;\phi) = \frac{1}{B(a,b)} h(x;\phi) H(x;\phi)^{a-1}\{1 - H(x;\phi)\}^{b-1}, \tag{1.4}$$

respectively, where $h(x;\phi) = dH(x;\phi)/dx$, a and b are two additional positive shape parameters, $I_y(a,b) = B_y(a,b)/B(a,b)$ is the incomplete beta function ratio, $I_y(a,b) = \int_0^y t^{a-1}(1-t)^{b-1}dt$ is the incomplete beta function, $B(a,b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$ and $\Gamma(.)$ is the gamma function. Clearly, for $a = b = 1$, we obtain the baseline distribution. The additional parameters a and b aim to govern skewness and tail weight of the generated distribution. An attractive feature of this family is that a and b can afford greater control over the weights in both tails and in the center of the distribution. In this paper, we generalize the W-G family by incorporating two additional shape parameters to yield a more flexible generator. Henceforward $g(x) = g(x;\phi)$, $G(x) = G(x;\phi)$ and so on. The cdf of the BW-G family is defined by

$$F(x) = I_{\left\{1-\exp\left[-\left(\frac{G(x;\phi)}{\overline{G}(x;\phi)}\right)^{\lambda}\right]\right\}}(a,b), \tag{1.5}$$

henceforward $G(x) = G(x;\phi)$, $\overline{G}(x) = \overline{G}(x;\phi)$, $g(x) = g(x;\phi)$. The corresponding pdf of (1.5) is given by

$$f(x) = \frac{\lambda\, g(x) G(x)^{\lambda-1}}{B(a,b)\overline{G}(x)^{\lambda+1}} \exp\left[-b\left(\frac{G(x)}{\overline{G}(x)}\right)^{\lambda}\right]\left\{1-\exp\left[-\left(\frac{G(x)}{\overline{G}(x)}\right)^{\lambda}\right]\right\}^{a-1}. \tag{1.6}$$

Here, a random variable X having the density function (1.6) is denoted by $X\sim$BW-G(a,b,λ,ϕ). The quantile function (qf) of X, say $Q(u) = F^{-1}(u)$, can be obtained by inverting (5). Let $G^{-1}(.) = Q_G(.)$ denote the qf of G, then, if $V\sim$Beta(a,b), then

$$X_V = G^{-1}\left\{\frac{[-\log(1-V)]^{1/\lambda}}{1+[-\log(1-V)]^{1/\lambda}}\right\}$$

has cdf (1.5). Some special cases of the BW-G family are given below

- For $b = 1$, the BW-G class reduces to the exponentiated Weibull-G (EW-G) family.
- For $a = 1$, we have the generalized Weibull-G (GW-G) class.
- The BW-G family reduces to the Weibull-G (W-G) family proposed by Bourguignon et al. (2014) when $a = b = 1$.

Consider the series

$$(1-z)^{b-1} = \sum_{i=0}^{\infty} \frac{(-1)^i\,\Gamma(b)}{i!\;\Gamma(b-i)}\, z^i, \tag{1.7}$$

which holds for $|z| < 1$ and $b > 0$ real non-integer. The pdf in (1.6) can be rewritten as

$$f(x) = \frac{\lambda\, g(x)}{B(a,b)} \frac{G(x)^{\lambda-1}}{\overline{G}(x)^{\lambda+1}} \exp\left[-b\left(\frac{G(x)}{\overline{G}(x)}\right)^{\lambda}\right]\underbrace{\left\{1-\exp\left[-\left(\frac{G(x)}{\overline{G}(x)}\right)^{\lambda}\right]\right\}^{a-1}}_{A},$$

applying (1.7) for A, we obtain

$$(x) = \frac{\lambda\, g(x) \frac{G(x)^{\lambda-1}}{\bar{G}(x)^{\lambda+1}}}{B(a,b)} \sum_{i=0}^{\infty} \frac{(-1)^i \Gamma(a)}{i!\,\Gamma(a-i)} \exp\left[-(i+b)\left(\frac{G(x)}{\bar{G}(x)}\right)^{\lambda}\right].$$

Using the power series expansion, the last equation can be expressed as

$$f(x) = \sum_{i,j=0}^{\infty} \frac{\lambda\,(i+b)^j h(x)}{B(a,b)} \frac{(-1)^{i+j}\Gamma(a)}{i!\,j!\,\Gamma(a-i)} \underbrace{\frac{G(x)^{\lambda(j+1)-1}}{\bar{G}(x)^{\lambda(j+1)+1}}}_{B}.$$

Applying (1.7) for B, we arrive at

$$f(x) = \sum_{j,k=0}^{\infty} \mathbf{t}_{j,k}\, \pi_{\lambda(j+1)+k}(x), \tag{1.8}$$

where $\pi_{\lambda(j+1)+k}(x) = \lambda(j+1) + k\; g(x)\; G(x)^{\lambda(j+1)+k-1}$ is the exp-G pdf with power parameter $\lambda(j+1)+k > 0$ and

$$\mathbf{t}_{j,k} = \frac{\lambda(-1)^j \Gamma(a)\Gamma(\lambda(j+1)+k+1)}{j!\,k!\,B(a,b)[\lambda(j+1)+k]\Gamma(\lambda(j+1)+1)} \sum_{i=0}^{\infty} \frac{(i+b)^j(-1)^i}{i!\,\Gamma(a-i)}.$$

Using equation (1.8) and generalized binomial expansion twice and changing the summation over m and s we can write

$$f(x) = \sum_{m=0}^{\infty} \mathbf{w}_m\, \pi_m\,(x), \tag{1.9}$$

where

$$\mathbf{w}_m = \sum\nolimits_{s=m}^{\infty} \sum\nolimits_{j,k=0}^{\infty} (-1)^{m+s} \binom{\lambda(j+1)+k}{s} \binom{s}{m}\, \mathbf{t}_{j,k}.$$

Thus, several mathematical properties of the BW-G family can be obtained simply from those properties of the exp-G family. Equation (1.9) is the main result of this section. The cdf of the BW-G family can also be expressed as a mixture of exp-G densities. By integrating (1.9), we obtain the same mixture representation

$$F(x) = \sum_{j,k=0}^{\infty} \mathbf{t}_{j,k}\, \mathbf{\Pi}_{\lambda(j+1)+k}(x) = \sum_{m=0}^{\infty} \mathbf{w}_m\, \mathbf{\Pi}_m\,(x), \tag{1.10}$$

where $\mathbf{\Pi}_m\,(x)$ is the cdf of the exp-G family with power parameter m**.** The properties of exp-G distributions have been studied by many authors in recent years, see Mudholkar and Srivastava (1993) and Mudholkar et al. (1995) for exponentiated Weibull (EW), Gupta et al. (1998) for exponentiated Pareto (EPa), Gupta and Kundu (1999) for exponentiated exponential (EE), Shirke and Kakade (2006) for exponentiated log-normal (ELN) and Nadarajah and Gupta (2007) for exponentiated gamma distributions (EGa), among others. Equations (1.9) and (1.10) are the main results of this section.

The basic motivation for using the BW-G family can be given as follows. To define special models with all types of the hazard rate function (hrf); to build heavy-tailed models that are not longer-tailed for modeling real data sets; to reproduce a skewness for symmetrical distributions; to generate distributions with symmetric, right-skewed, left-skewed and reversed-J shaped; to convert the kurtosis of the new models more flexible compared to the baseline model; to generate a large number of special distributions; and to provide consistently better fits than other generated models under the same baseline distribution. That, is well-demonstrated by fitting the BW-N distribution to two real data sets in Section 7. However, we expect that there are other contexts in which the BW special models can produce worse fits than other generated distributions. Clearly, the results in Section 7 indicate that the new family is a very competitive class to other known generators with at most three extra shape parameters.

The rest of the paper is outlined as follows. In Section 2, we define two special models and provide the plots of their pdfs. In Section 3, we derive some of its mathematical properties including order statistics and their moments, entropies, ordinary and incomplete moments, moment generating function (mgf), stress-strength model, residual and reversed residual life functions. Some characterization results are provided in Section 4. Maximum likelihood estimation of the model parameters is addressed in Section 5. In Section 6, simulation results to assess the performance of the proposed maximum likelihood estimation procedure are discussed. In Section 7, we provide two applications to real data to illustrate the importance of the new family. Finally, some concluding remarks are presented in Section 8.

2. The BW-G SPECIAL MODELS

In this section, we provide two special models of the BW-G family. The pdf (1.6) will be most tractable when $g(x)$ and $G(x)$ have simple analytic expressions. These special models generalize some well-known distributions in the literature.

2.1 The BW-Weibull (BWW) Model

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

$$f(x) = \frac{\lambda\,\beta\alpha^{\beta}x^{\beta-1}\exp\left(-b\left\{\exp\left[(\alpha x)^{\beta}\right]-1\right\}^{\lambda}\right)}{B(a,b)\exp\left[-\lambda(\alpha x)^{\beta}\right]\left\{1-\exp\left[-(\alpha x)^{\beta}\right]\right\}^{1-\lambda}} \left[1-\exp\left(-\left\{\exp\left[(\alpha x)^{\beta}\right]-1\right\}^{\lambda}\right)\right]^{a-1}.$$

The BWW distribution includes the Weibull Weibull (WW) distribution when $a = b = 1$. For $b = 0$, we obtain the exponentiated Weibull Weibull (EWW) distribution. For $a = 1$, we have the generalized Weibull Weibull (GWW) distribution. For $\beta = 1$ and $\beta = 2$, we have the BW exponential and BW Rayleigh distributions, respectively. The BWW density and hrf plots for selected parameter values are displayed in Figure 1.

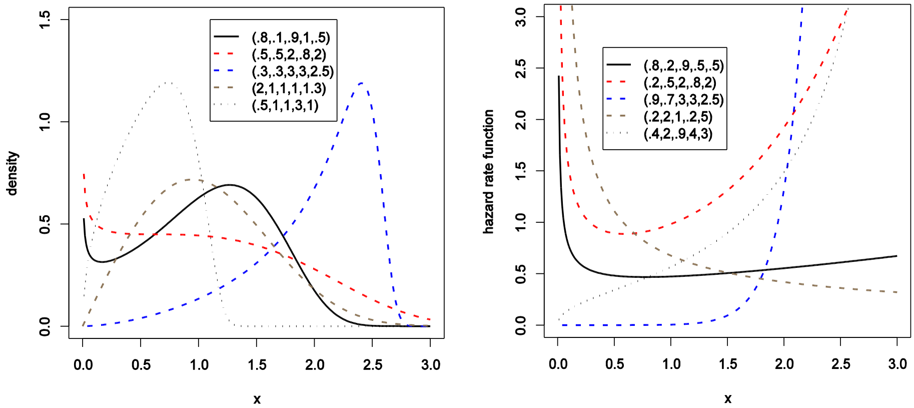

Figure 1: Plots of the BWW pdf and hrf for Selected Values of Parameters

2.2 The BW-Normal (BWN) Model

The cdf and pdf (for $-\infty < x < \infty$) of the normal distribution with $-\infty < \mu < \infty$ and $\sigma > 0$ are $G(x) = \boldsymbol{\Phi}\left(\frac{x-\mu}{\sigma}\right)$ and $g(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$, respectively. Then, the pdf of the BWN distribution becomes

$$f(x) = \frac{\lambda\,\Phi\left(\frac{x-\mu}{\sigma}\right)}{B(a,b)} \frac{\left[\boldsymbol{\Phi}\left(\frac{x-\mu}{\sigma}\right)\right]^{\lambda-1}}{\left[1-\boldsymbol{\Phi}\left(\frac{x-\mu}{\sigma}\right)\right]^{\lambda+1}} \exp\left[-b\left(\frac{\boldsymbol{\Phi}\left(\frac{x-\mu}{\sigma}\right)}{1-\boldsymbol{\Phi}\left(\frac{x-\mu}{\sigma}\right)}\right)^{\lambda}\right] \left\{1-\exp\left[-\left(\frac{\boldsymbol{\Phi}\left(\frac{x-\mu}{\sigma}\right)}{1-\boldsymbol{\Phi}\left(\frac{x-\mu}{\sigma}\right)}\right)^{\lambda}\right]\right\}^{a-1}.$$

The BWN distribution includes the Weibull normal (WN) distribution when $a = b = 1$. For $b = 0$, we obtain the exponentiated Weibull normal (EWN) distribution. For $a = 1$, we have the generalized Weibull normal (GWN) distribution. The importance of the BWN model is that the normal distribution is symmetric, but the BWN becomes skewed. The BWN density plots for selected parameter values are displayed in Figure 2.

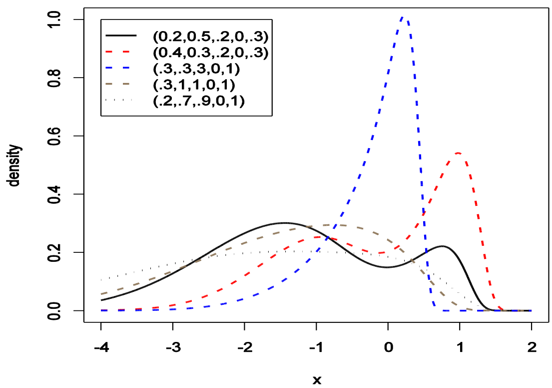

Figure 2: Plots of the BWN pdf for Selected Values of Parameters

3. PROPERTIES

3.1 Order Statistics

Order statistics make their appearance in many areas of statistical theory and practice. Suppose $X_1, \ldots, X_n$ is a random sample from any BW-G distribution. Let $X_{i:n}$ denote the ith order statistic. The pdf of $X_{i:n}$ can be expressed as

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

where $c = \frac{1}{B(i,n-i+1)}$. We use the result 0.314 of Gradshteyn and Ryzhik (2000) for a power series raised to a positive integer n (for $n \geq 1$), $\left(\sum_{i=0}^{\infty} a_i\, u^i\right)^n = \sum_{i=0}^{\infty} c_{n,i}\, u^i$, where the coefficients $c_{n,i}$ (for $i = 1,2,\ldots$) are determined from the recurrence equation (with $c_{n,0} = a_0^n$)

$$c_{n,i} = (i\, a_0)^{-1} \sum_{m=1}^{i} [m(n+1) - i]\, a_m\, c_{n,i-m}. \tag{3.1}$$

We can demonstrate that the pdf of the ith order statistic of any BW-G distribution can be expressed as

$$f_{i:n}(x) = \sum_{l,k=0}^{\infty} \mathbf{b}_{l,k}\, \pi_{l+k+1}(x), \tag{3.2}$$

where $\mathbf{b}_{l,k} = \frac{n!\,(l+1)\,(i-1)!\,\mathbf{w}_{l+1}}{(l+k+1)} \sum_{j=0}^{n-i} \frac{(-1)^j f_{j+i-1,k}}{(n-i-j)!\, j!}$, $\mathbf{w}_r$ is given in Section 2 and the quantities $f_{j+i-1,k}$ can be determined with $f_{j+i-1,0} = w_0^{j+i-1}$ and recursively for $k \geq 1$

$$f_{j+i-1,k} = (k\, \mathbf{w}_0)^{-1} \sum_{m=1}^{k} [m\,(j+i) - k]\, \mathbf{w}_m\, f_{j+i-1,k-m}.$$

We can obtain the ordinary and incomplete moments, generating function and mean deviations of the BW-G order statistics from equation (3.2) and some properties of the exp-G model. For the BW-W model the moments of $X_{i:n}$ can be expressed as

$$E\left(X_{i:n}^q\right) = \sum_{l,k,h=0}^{\infty} \frac{(l+k+1)(-1)^h \mathbf{b}_{l,k}}{\alpha^q (h+1)^{(q+\beta)/\beta}} \binom{l+k-1}{h} \Gamma\left(1 + \frac{q}{\beta}\right), \forall\, q > -\beta.$$

3.2 Entropy Measures

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

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

Using the pdf (1.6), we can write

$$f(x)^{\delta}=\sum_{j,k=0}^{\infty} m_{j,k}\, g(x)^{\delta} G(x)^{\lambda j+\delta(\lambda-1)+k},$$

where

$$m_{j,k}=\left(\frac{\lambda}{B(a,b)}\right)^{\delta}\sum_{i=0}^{\infty}\frac{(-1)^{i+j}[(\delta b+i)]^{j}\Gamma(\delta(a-1)+1)\Gamma(\lambda j+\delta(\lambda+1)+k)}{i!\,j!\,k!\,\Gamma(\delta(a-1)-i+1)\Gamma\big(\lambda j+\delta(\lambda+1)\big)}.$$

Then, the Rényi entropy of the BW-G family is given by

$$I_{\delta}(X)=(1-\delta)^{-1}\log\left\{\sum_{j,k=0}^{\infty} m_{j,k}\int_{-\infty}^{\infty} g(x)^{\delta} G(x)^{\lambda j+\delta(\lambda-1)+k}dx\right\}.$$

The q-entropy, say $H_q(X)$, can be obtained as

$$H_q(X)=(q-1)^{-1}\log\left\{1-\left[\sum_{j,k=0}^{\infty} m_{j,k}\int_{-\infty}^{\infty} g(x)^{q} G(x)^{\lambda j+q(\lambda-1)+k}dx\right]\right\},$$

where $q>0$, $q\neq 1$. The Shannon entropy of a random variable X, say SI, is defined by $SI=E\{-[\log f(X)]\}$. It is the special case of the Rényi entropy when $\delta\uparrow 1$.

3.3 Moments, Generating Function and Incomplete Moments

The r^{th} moment of X, say μ_r', follows from (1.9) as

$$\mu_r'=E(X^r)=\sum\nolimits_{m=0}^{\infty}\mathbf{w}_m\, E(Y_m^r).$$

Henceforth, Y_m denotes the exp-G distribution with power parameter m.

The n^{th} central moment of X, say M_n, is given by

$$M_n=E(X-\mu_1')^n=\sum\nolimits_{r=0}^{n}\sum\nolimits_{m=0}^{\infty}(-1)^{n-r}\,\mathbf{w}_m\binom{n}{r}\mu_r'^{(n-r)}\,E(Y_{a+k}^r).$$

The cumulants (κ_n) of X follow recursively from $\kappa_n=\mu_n'-\sum_{r=0}^{n-1}\binom{n-1}{r-1}\kappa_r\,\mu_{n-r}'$, where $\kappa_1=\mu_1'$, $\kappa_2=\mu_2'-\mu_1'^2, \kappa_3=\mu_3'-3\mu_2'\mu_1'+\mu_1'^3$, etc. The skewness and kurtosis measures can be calculated from the ordinary moments using well-known relationships.

The mgf $M_X(t)=E(e^{t\,X})$ of X is given by $M_X(t)=\sum_{m=0}^{\infty}\mathbf{w}_m\, M_m(t)$, where $M_m(t)$ is the mgf of Y_m. Hence, $M_X(t)$ can be determined from the exp-G generating function. The s^{th} incomplete moment, say $I_s(t)$, of X can be expressed from (1.9) as

$$I_s(t)=\int_{-\infty}^{t}x^s f(x)dx=\sum_{m=0}^{\infty}\mathbf{w}_m\int_{-\infty}^{t}x^s\,\pi_m(x)dx. \tag{3.3}$$

A general formula for the first incomplete moment, $I_1(t)$, can be derived from (3.3) (with $s=1$) as $I_1(t)=\sum_{m=0}^{\infty}\mathbf{w}_m\, v_m(x)$, where $v_m(x)=\int_{-\infty}^{t}x\,\pi_m(x)dx$ is the first incomplete moment of the exp-G distribution.

This equation for $I_1(t)$ can be applied to construct Bonferroni and Lorenz curves defined for a given probability π by $B(\pi) = I_1(q)/(\pi\mu_1')$ and $L(\pi) = I_1(q)/\mu_1'$, respectively, where $\mu_1' = E(X)$ and $q = Q(\pi)$ is the qf of X at π. The mean deviations about the mean $[\delta_1 = E(|X - \mu_1'|)]$ and about the median $[\delta_2 = E(|X - M|)]$ of X are given by $\delta_1 = 2\mu_1' F(\mu_1') - 2I_1(\mu_1')$ and $\delta_2 = \mu_1' - 2I_1(M)$, respectively, where $\mu_1' = E(X)$, $M = Median(X) = Q(0.5)$ is the median and $F(\mu_1')$ is easily calculated from (1.5). For the BWW model

$$\mu_r' = \sum_{m,h=0}^{\infty} \frac{m(-1)^h \mathbf{w}_m}{\alpha^r (h+1)^{(r+\beta)/\beta}} \binom{m-1}{h} \Gamma\left(1 + \frac{r}{\beta}\right), \forall\, r > -\beta$$

and

$$I_s(t) = \sum_{m,h=0}^{\infty} \frac{m(-1)^h \mathbf{w}_m}{\alpha^s (h+1)^{(s+\beta)/\beta}} \binom{m-1}{h} \gamma\left(1 + \frac{s}{\beta}, \left(\frac{\alpha}{t}\right)^{\beta}\right), \forall\, s > -\beta.$$

3.4 Stress-Strength Model

Stress-strength model is the most broadly approach utilized for dependability estimation. This model is used in many applications of physics and engineering such as strength failure and system collapse. In stress-strength modeling, $R = \Pr(X_2 < X_1)$ 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 $X_1 > X_2$. R can be considered as a measure of system performance and naturally arise in electrical and electronic systems. Other interpretation can be that, the reliability, say R, of the 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 BW-G$(a_1, b_1, \lambda_1, \phi)$ and BW-G$(a_2, b_2, \lambda_2, \phi)$ distributions.

The reliability is given by $R = \int_0^\infty f_1(a_1, b_1, \lambda_1; \phi) F_2(a_2, b_2, \lambda_2; \phi) dx$. Then

$$R = \sum_{m,c=0}^{\infty} \mho_{m,c}.$$

where

$$\mho_{m,c} = \sum_{s=m}^{\infty} \sum_{d=c}^{\infty} \sum_{j,k,w,l=0}^{\infty} (-1)^{m+s+c+d} \binom{\lambda_1(j+1)+k}{s} \binom{\lambda_2(w+1)+l}{d} \binom{s}{m} \binom{d}{c} \mathbf{t}_{j,k,w,l}$$

and

$$\mathbf{t}_{j,k,w,l} \frac{\lambda_1 \lambda_2 (-1)^{j+w} \Gamma(a_1)\Gamma(a_2)}{c(m+c) j!\, k!\, w!\, l!\, B(a_1, b_1) B(a_2, b_2)} \sum_{i,h=0}^{\infty} \frac{(i+b_1)^j (h+b_2)^w (-1)^{i+h}}{i!\, h!\, \Gamma(a_1 - i)\Gamma(a_2 - h)} \frac{\Gamma(\lambda_1(j+1)+k+1)\Gamma(\lambda_2(w+1)+l+1)}{\Gamma(\lambda_1(j+1)+1)\Gamma(\lambda_2(w+1)+1)[\lambda_1(j+1)+k][\lambda_2(w+1)+l]}.$$

3.5 Moment of Residual and Reversed Residual Lifes

The n^{th} moment of the residual life, say $m_n(t) = E[(X-t)^n | X > t]$, $n = 1,2,...$, uniquely determines $F(x)$. The n^{th} moment of the residual life of X is given by $m_n(t) = \frac{1}{1-F(t)} \int_t^{\infty} (x-t)^n dF(x)$.

Subsequently, $m_n(t) = \frac{1}{1-F(t)} \sum_{m=0}^{\infty} \mathbf{w}_m \sum_{r=0}^{n} \binom{n}{r} (-t)^{n-r} \int_t^{\infty} x^r \pi_m(x) dx$.

For the BWW model we have

$$m_n(t) = \frac{1}{1-F(t)} \sum_{m,h=0}^{\infty} \sum_{r=0}^{n} \frac{m(-1)^h (-t)^{n-r} \mathbf{w}_m \binom{n}{r} \binom{m-1}{h}}{\alpha^n (h+1)^{\frac{(n+\beta)}{\beta}}} \gamma \left(1 + \frac{n}{\beta}, \left(\frac{\alpha}{t}\right)^{\beta}\right), \forall\, n > -\beta.$$

Another important function is the mean residual life 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 $n = 1,2,...$ 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 $M_n(t) = \frac{1}{F(t)} \sum_{m=0}^{\infty} \mathbf{w}_m \sum_{r=0}^{n} (-1)^r \binom{n}{r} t^{n-r} \int_0^t x^r \pi_m(x) dx$.

For the BWW model

$$M_n(t) = \frac{1}{F(t)} \sum_{m,h=0}^{\infty} \sum_{r=0}^{n} \frac{m(-1)^{r+h} t^{n-r} \mathbf{w}_m \binom{n}{r} \binom{m-1}{h}}{\alpha^n (h+1)^{\frac{(n+\beta)}{\beta}}} \gamma \left(1 + \frac{n}{\beta}, \left(\frac{\alpha}{t}\right)^{\beta}\right), \forall\, n > -\beta.$$

The mean inactivity time (MIT) or mean waiting time 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 BW-G family of distributions can be obtained easily by setting $n = 1$ in the above equation.

4. CHARACTERIZATIONS

The problem of characterizing probability distributions is important for applied scientists who would like to know if their proposed model is the right one for their data. This section deals with various characterizations of BW-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.

4.1 Characterizations based on two Truncated Moments

In this subsection we present characterizations of BW-G distribution in terms of a simple relationship between two truncated moments. Our first characterization result employs a theorem due to Glänzel (1987), see Theorem 1 of Appendix A. Note that the result holds also when the interval H is not closed. Moreover, it could be also applied when the cdf F does not have a closed form. As shown in Glänzel (1990), this characterization is stable in the sense of weak convergence. Here is our first characterization.

Proposition 4.1

Let $X: \Omega \to (0, \infty)$ be a continuous random variable and let $q_1(x) = \left\{1 - \exp\left[-\left(\frac{G(x)}{\overline{G}(x)}\right)^{\lambda}\right]\right\}^{1-a}$ and $q_2(x) = q_1(x)\exp\left[-\left(\frac{G(x)}{\overline{G}(x)}\right)^{\lambda}\right]$ for $x > 0$. The random variable X belongs to BW-G family (1.6) if and only if the function η defined in Theorem1 has the form

$$\eta(x) = \frac{b}{b+1}\exp\left[-\left(\frac{G(x)}{\overline{G}(x)}\right)^{\lambda}\right], x > 0. \tag{4.1}$$

Proof.

Let X be a random variable with pdf (6), then

$$\left(1 - F(x)\right)E[q_1(x)|X \geq x] = \frac{1}{bB(a,b)}\exp\left[-b\left(\frac{G(x)}{\overline{G}(x)}\right)^{\lambda}\right], x > 0,$$

and

$$\left(1 - F(x)\right)E[q_2(x)|X \geq x] = \frac{1}{(b+1)B(a,b)}\exp\left[-(b+1)\left(\frac{G(x)}{\overline{G}(x)}\right)^{\lambda}\right], x > 0,$$

and finally

$$\eta(x)q_1(x) - q_2(x) = -\frac{1}{b+1}q_1(x)\frac{1}{bB(a,b)}\exp\left[-\left(\frac{G(x)}{\overline{G}(x)}\right)^{\lambda}\right] < 0\ for\ x > 0.$$

Conversely, if η is given as above, then

$$s'(x) = \frac{\eta'(x)q_1(x)}{\eta(x)q_1(x) - q_2(x)} = \frac{\lambda b g(x) G(x)^{\lambda-1}}{\overline{G}(x)^{\lambda+1}}, x > 0,$$

and hence, $s(x) = b\left(\frac{G(x)}{\overline{G}(x)}\right)^{\lambda}, x > 0$. Now, in view of Theorem1, X has density (1.6).

Corollary 4.1

Let $X: \Omega \to (0, \infty)$ be a continuous random variable and let $q_1(x)$ be as in Proposition 4.1. The pdf of X is (6) if and only if there exist functions q_2 and η defined in Theorem 1 satisfying the differential equation

$$\frac{\eta'(x)q_1(x)}{\eta(x)q_1(x) - q_2(x)} = \frac{\lambda b g(x) G(x)^{\lambda-1}}{\overline{G}(x)^{\lambda+1}}, x > 0. \tag{4.2}$$

Remark 4.1

The general solution of the differential equation in Corollary 4.1 is

$$\eta(x) = \exp\left[b\left(\frac{G(x)}{\overline{G}(x)}\right)^{\lambda}\right] \left\{-\int \frac{\lambda b g(x) G(x)^{\lambda-1}}{\overline{G}(x)^{\lambda+1}} \exp\left[-b\left(\frac{G(x)}{\overline{G}(x)}\right)^{\lambda}\right] \left(q_1(x)\right)^{-1} q_2(x) dx + D\right\},$$

where D is a constant. Note that a set of functions satisfying the differential equation (4.2) is given in Proposition 4.1 with $D = 0$. However, it should be also noted that there are other triplets (q_1, q_2, η) satisfying the conditions of Theorem1.

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

$$\frac{f'(x)}{f(x)} = \frac{h_F'(x)}{h_F(x)} - h_F(x). \tag{4.3}$$

For many univariate continuous distributions, this is the only characterization available in terms of the hazard function. The following characterization establish a non-trivial characterization for BW-G distribution in terms of the hazard function when $a = 1$, which is not of the trivial form given in (4.3). Clearly, we assume that $G(x)$ is twice differentiable.

Proposition 4.2

Let $X: \Omega \to (0, \infty)$ be a continuous random variable. Then for $a = 1$, the pdf of X is (1.6) if and only if its hazard function $h_F(x)$ satisfies the differential equation

$$h_F'(x) - \frac{g'(x)}{g(x)} h_F(x) = \frac{\lambda b g(x) G(x)^{\lambda-2}\left\{\lambda + G(x) - \overline{G}(x)\right\}}{\overline{G}(x)^{\lambda+2}}, \tag{4.4}$$

with the boundary condition $h_F(0) = 0$ for $\lambda > 1$.

Proof.

If X has pdf (1.6), then clearly (4.4) holds. Now, if (4.4) holds, then

$$\frac{d}{dx}\left\{\left(g(x)\right)^{-1} h_F(x)\right\} = \lambda b \frac{d}{dx}\left[\left(\frac{G(x)^{\lambda-1}}{\overline{G}(x)^{\lambda+1}}\right)\right],$$

or, equivalently, $h_F(x) = \frac{\lambda b g(x) G(x)^{\lambda-1}}{\overline{G}(x)^{\lambda+1}}$, which is the hazard function of the BW-G distribution.

4.3 Characterization based on Truncated Moment of Certain Function of the Random Variable

The following proposition has already appeared in (Hamedani, Technical Report, 2013), so we will just state them here which can be used to characterize BW-G distribution.

Proposition 4.3

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 $\delta \neq 1$,

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

if and only if

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

Remark 4.3

It is easy to see that for certain functions $\psi(x)$ on $(0,\infty)$, Proposition 4.3 provides a characterization of BW-G distribution for $a=1$.

5. ESTIMATION

Several approaches for parameter estimation were proposed in the literature but the maximum likelihood method is the most commonly employed. The maximum likelihood estimators (MLEs) enjoy desirable properties and can be used when constructing confidence intervals and also in test statistics. The normal approximation for these estimators in large sample theory is easily handled either analytically or numerically. So, we consider the estimation of the unknown parameters for this family from complete samples only by maximum likelihood. Here, we determine the MLEs of the parameters of the new family of distributions from complete samples only.

Let $x_1,\ldots,x_n$ be a random sample from the BW-G family with parameters a,b,λ and ϕ,where ϕ is a $q\times 1$ baseline parameter vector. Let $\mathbf{\Psi}=(a,b,\lambda,\phi^{\mathrm{T}})^{\mathrm{T}}$ a $(q+3)\times 1$ parameter vector. Then, the log-likelihood function for $\mathbf{\Psi}$, say $\ell=\ell(\mathbf{\Psi})$, is given by

$$\ell = -n\log[B(a,b)] + n\log\lambda + \sum_{i=1}^{n}\log g(x_i;\phi) + (\lambda-1)\sum_{i=1}^{n}\log G(x_i;\phi)$$
$$-(\lambda+1)\sum_{i=1}^{n}\log\overline{G}(x_i;\phi) - b\sum_{i=1}^{n}s_i^{\lambda} + (a-1)\sum_{i=1}^{n}\log z_i, \qquad (5.1)$$

where $s_i=\frac{G(x_i;\phi)}{\overline{G}(x_i;\phi)}$ and $z_i=\left[1-\exp\left(-s_i^{\lambda}\right)\right]$. Equation (5.1) can be maximized either directly by using the R (optim function), SAS (PROC NLMIXED) or Ox program (sub-routine MaxBFGS) or by solving the nonlinear likelihood equations obtained by differentiating (5.1). The score vector components, say

$$\mathbf{U}(\mathbf{\Psi}) = \frac{\partial\ell}{\partial\mathbf{\Psi}} = \left(\frac{\partial\ell}{\partial a},\frac{\partial\ell}{\partial b},\frac{\partial\ell}{\partial\lambda},\frac{\partial\ell}{\partial\phi_r}\right)^{\mathrm{T}} = \left(U_a,U_b,U_\lambda,U_{\phi_r}\right)^{\mathrm{T}},$$

are given by

$$U_a = -n[\psi(a)-\psi(a+b)] + \sum_{i=1}^{n}\log z_i,$$

$$U_b = -n[\psi(b) - \psi(a+b)] - \sum_{i=1}^{n} s_i^{\lambda},$$

$$U_\lambda = \frac{n}{\lambda} + \sum_{i=1}^{n} \log G(x_i;\phi) - \sum_{i=1}^{n} \log \overline{G}(x_i;\phi) - b\theta \sum_{i=1}^{n} s_i^{\lambda} \log s_i + (a-1) \sum_{i=1}^{n} \frac{m_i}{z_i},$$

and (for $r = 1, \ldots, q$)

$$U_{\phi_r} = \sum_{i=1}^{n} \frac{g_r'(x_i;\phi)}{g(x_i;\phi)} + (\lambda - 1) \sum_{i=1}^{n} \frac{G_r'(x_i;\phi)}{G(x_i;\phi)}$$
$$+(\lambda + 1) \sum_{i=1}^{n} \frac{G_r'(x_i;\phi)}{\overline{G}(x_i;\phi)} - b\lambda \sum_{i=1}^{n} s_i^{\lambda-1} p_{i,r} + (a-1) \sum_{i=1}^{n} \frac{w_i}{z_i},$$

where $\psi(a)$ is the digamma function, $g_r'(x_i;\phi) = \partial g(x_i;\phi)/\partial\phi_r$, $G'(x_i;\phi) = \partial G(x_i;\phi)/\partial\phi_r$, $p_{i,r} = \frac{\overline{G}(x_i;\phi)G_r'(x_i;\phi) + G(x_i;\phi)G_r'(x_i;\phi)}{\overline{G}(x_i;\phi)^2}$, $m_i = s_i^{\lambda} \log s_i \exp(-s_i^{\lambda})$ and $w_i = \lambda z_i s_i^{\lambda-1} \exp(-s_i^{\lambda})$.

Setting the nonlinear system of equations $U_a = U_b = U_\lambda = 0$ and $U_{\phi_k} = \mathbf{0}$ (for $k = 1 = \cdots, q$) and solving them simultaneously yields the MLEs $\widehat{\mathbf{\Psi}} = (\hat{a}, \hat{b}, \hat{\lambda}, \hat{\phi}^{\mathrm{T}})^{\mathrm{T}}$. To solve these equations, it is usually more convenient to use nonlinear optimization methods such as the quasi-Newton algorithm to numerically maximize $\ell(\mathbf{\Psi})$. For interval estimation of the parameters, we can evaluate numerically the elements of the $(q+3) \times (q+3)$ observed information matrix $J(\mathbf{\Psi}) = \left(-\frac{\partial^2 \ell}{\partial\theta_r\, \theta_s}\right)$.

Under standard regularity conditions when $n \to \infty$, the distribution of $\widehat{\mathbf{\Psi}}$ can be approximated by a multivariate normal $N_p\left(0, J(\widehat{\mathbf{\Psi}})^{-1}\right)$ distribution to construct approximate confidence intervals for the parameters. Here, $J(\widehat{\mathbf{\Psi}})$ is the total observed information matrix evaluated at $\widehat{\mathbf{\Psi}}$. The method of the re-sampling bootstrap can be used for correcting the biases of the MLEs of the model parameters. Good interval estimates may also be obtained using the bootstrap percentile method. We can compute the maximum values of the unrestricted and restricted log-likelihoods to obtain likelihood ratio (LR) statistics for testing some sub-models of the BW-G distribution.

Hypothesis tests of the type $H_0\colon \omega = \omega_0$ versus $H_1\colon \omega \neq \omega_0$, where ω is a vector formed with some components of $\mathbf{\Psi}$ and ω_0 is a specified vector, can be performed using LR statistics. For example, the test of $H_0\colon a = b = \lambda = 1$ versus $H_1\colon H_0$ is not true is equivalent to comparing the BW-G and G distributions and the LR statistic is given by $w = 2\{\ell(\hat{a}, \hat{b}, \hat{\lambda}, \hat{\phi}) - \ell(1,1,1,\hat{\phi})\}$, where $\hat{a}$, $\hat{b}$, $\hat{\lambda}$ and $\hat{\phi}$ are the MLEs under H and $\hat{\phi}$ is the estimate under H_0. The elements of $J(\mathbf{\Psi})$ are given in the Appendix B.

6. SIMULATION STUDY

In order to assess the performance of the MLEs, a small simulation study is performed using the statistical software R through the package (stats4), command MLE, for a particular member of the BW-G family of distributions, the BWE (Beta Weibull-exponential) distribution. However, one can also perform in SAS by PROC NLMIXED procedure. The number of Monte Carlo replications was $15{,}000$. For maximizing the log-likelihood function, we use the MaxBFGS subroutine with analytical derivatives. The evaluation of the estimates was performed based on the following quantities for each sample size: the empirical mean squared errors (MSEs) are calculated using the R package from the Monte Carlo replications. The MLEs are determined for each simulated data, say, $(\hat{\alpha}_i, \hat{\lambda}_i, \hat{a}, \hat{b})$ for $i = 1,2,\cdots,15{,}000$ and the biases and MSEs are computed by

$$bias_h(n) = \frac{1}{15000}\sum_{i=1}^{15000} \left(\hat{h}_i - h\right),$$

and

$$MSE_h(n) = \frac{1}{15000}\sum_{i=1}^{15000} (\hat{h}_i - h)^2, \text{for } h = \alpha, \lambda, a, b$$

We consider the sample sizes at $n = 50{,}100$ and 200 and consider different values for the parameters (I: $\alpha = \lambda = a = b = 0.5$, II: $\alpha = 0.5$, $\lambda = 0.5$, $a = 1.3$ and $b = 1.5$, III: $\alpha = 0.7$, $\lambda = 0.8$, $a = 0.8$ and $b = 0{,}9$, IV: $\alpha = 0.9$, $\lambda = 0.7$, $a = 1.2$ and $b = 1.4$, V: $\alpha = 1$, $\lambda = 1.5$, $a = 0.9$ and $b = 0.6$, VI: $\alpha = 1.5$, $\lambda = 2$, $a = 0.6$ and $b = 0.8$). The empirical results are given in Table 1. The figures in Table 1 indicate that the estimates are quite stable and, more important, are close to the true values for the these sample sizes. Furthermore, as the sample size increases, the MSEs decreases as expected.

7. APPLICATION

In this section, we illustrate the applicability of the BW-G family to two real data sets. We focus on the BWN distribution presented in Subsection 2.3. The method of maximum likelihood is used to estimate the model parameters. We shall fit the BWN model using two real data sets and compare it with other existing competitive distributions. In order to compare the fits of the distributions, we consider various measures of goodness-of-fit including the maximized log-likelihood under the model $(-\hat{\ell})$, Anderson-Darling (A^*) and Cramér-Von Mises (W^*) statistics. The smaller these statistics show better model for fitting. The first data set, referred to as D2, consists of lifetimes of 43 blood cancer patients (in days) from one of the ministry of Health Hospitals in Saudi Arabia, see Abouammoh et al. (1994). These data are: 115, 181, 255, 418, 441, 461, 516, 739, 743, 789, 807, 865, 924, 983, 1025, 1062, 1063, 1165, 1191, 1222, 1222, 1251, 1277, 1290, 1357, 1369, 1408, 1455, 1478, 1519, 1578, 1578, 1599, 1603, 1605, 1696, 1735, 1799, 1815 ,1852, 1899, 1925, 1965. Table 2 shows summary statistics for this data set.

Table 1
Bias and MSE of the Estimates under the Method of Maximum Likelihood

n	Actual Value	Bias				MSE			
		$\hat{\alpha}$	$\hat{\lambda}$	$\hat{a}$	$\hat{b}$	$\hat{\alpha}$	$\hat{\lambda}$	$\hat{a}$	$\hat{b}$
50	I	−0.4173	0.4198	0.3554	−0.3932	0.0518	0.0460	0.0538	0.953
	II	−0.7738	0.3239	−0.2145	−0.3426	0.0180	0.0426	0.0978	0.626
	III	0.4891	−0.2460	−0.6227	0.4821	0.0154	0.1207	0.1065	0.167
	IV	0.1883	0.9799	−0.5091	0.0563	0.0481	0.0222	0.0442	0.1138
	V	0.1780	−0.4983	−0.4292	−0.5454	0.4271	0.0282	0.0917	0.4956
	VI	−0.0841	−0.3634	−0.4059	−0.2203	0.9537	0.0183	0.0737	0.583
100	I	0.0717	0.3617	0.0499	0.0729	0.0216	0.0230	0.0248	0.3132
	II	0.5182	0.1845	0.0843	0.1153	0.0084	0.0229	0.0457	0.5781
	III	0.3165	0.1594	0.0505	−0.3298	0.0065	0.0591	0.0441	0.1587
	IV	0.1375	−0.0493	0.1309	−0.0318	0.0181	0.0105	0.0203	0.0955
	V	0.1258	0.4753	0.0866	0.2424	0.0646	0.0131	0.0281	0.1472
	VI	0.0343	0.1737	0.2243	−0.1509	0.4015	0.0089	0.0367	0.3381
200	I	−0.04609	−0.02898	−0.0368	−0.0475	0.0094	0.0114	0.0111	0.0842
	II	−0.0512	−0.1110	−0.0355	0.0023	0.0041	0.0102	0.0229	0.0182
	III	−0.0730	−0.0527	0.0467	−0.02271	0.0039	0.0315	0.0241	0.0364
	IV	−0.1023	−0.0208	−0.0986	−0.0235	0.0086	0.0053	0.0102	0.0217
	V	−0.0783	−0.0519	−0.0169	0.0026	0.0270	0.0071	0.0123	0.0154
	VI	0.0078	−0.0691	0.0660	−0.0853	0.1368	0.0040	0.0149	0.01035

The second data set is IQ data set for 52 non-white males hired by a large insurance company in 1971 given in Roberts (1988). These data are: 91, 102, 100, 117, 122, 115, 97, 109, 108, 104, 108, 118, 103, 123, 123, 103, 106,102, 118, 100, 103, 107, 108, 107, 97, 95, 119, 102, 108, 103, 102, 112, 99, 116, 114, 102, 111, 104, 122, 103, 111, 101, 91, 99, 121, 97, 109, 106, 102, 104, 107, 955. Table 1 presents summary statistics for this data set.

Table 2
Summary Statistics of Two Data Sets

Data	Size	Arithmetic Mean	Standard Deviation	Skewness	Kurtosis
D2	43	1192	506.635	−0.44	−0.72
IQ	52	106.65	8.30	0.37	−0.56

For the first data set, BWN is compared with other generalized normal models such as the Beta Normal (BN) (Eugene et al., 2002), Kumaraswamy Normal (KN) (Cordeiro and de Castro, 2011), Weibull Normal (WN) (Bourguignon et al., 2014), Gamma Normal (GN) (Alzaatreh et al., 2014).

In the second application, BWN is compared with the Skew Symmetric Component Normal (SSCN) (Rasekhi et al, 2015), BN, KN, WN. Its worth to mention, Rasekhi et al (2015) illustrated that the SSCN model for this data set is more better than flexible skew symmetric normal distributions such as FGSN (Ma and Genton 2004), SN (Azzalini, 1985), ESN (Mudholkar and Hutson 2000), FS (Fernandez and Steel 1998). Also Kumar and Anusree (2014) used this data for comparison of four classes of skew normal distribution.

The maximum likelihood estimates of the parameters are obtained and the goodness of fit tests are reported in Tables 3 and 4. Figure 3 shows fitted distributions on histogram of two data sets and Figures 4 and 5 show fitted cdf on empirical cdf of two data sets. It is clear from Tables 3 and 4 and Figures 3, 4 and 5 that the BWN model provides the best fits to both data sets.

Table 3
The MLEs of the Parameters, SEs in Parentheses and the Goodness-of-Fit Statistics for D2

Model	Estimates					$-2\hat{\ell}$	W^*	A^*
$B(a, b, \lambda, \mu, \sigma)$	8.51 (0.75)	3.94 (0.33)	0.08 $(8e^{-3})$	984.45 (56.59)	254.51 (16.44)	325.01	0.016	0.143
$BN(\alpha, \beta, \mu, \sigma)$	0.02 $(3e^{-3})$	9.77 (9.90)	2151.33 (47.05)	109.77 (8.20)		325.19	0.022	0.199
$KN(\alpha, \beta, \mu, \sigma)$	0.01 $(2e^{-3})$	0.51 (0.07)	1825.28 (48.95)	73.44 (7.46)		325.44	0.021	0.182
$WN(\alpha, \beta, \mu, \sigma)$	0.40 (0.06)	0.43 (0.03)	1036.70 (51.61)	355.08 (21.41)		325.59	0.028	0.213
$GN(\alpha, \beta, \mu, \sigma)$	5.45 (0.33)	0.15 (0.01)	1013.45 (76.33)	1246.13 (130.37		328.25	0.069	0.491

Table 4
The MLEs of the Parameters and SEs in Parentheses and the Goodness-of-Fit Statistics for IQ

Model	Estimates					$-2\hat{\ell}$	W^*	A^*
$BW(a, b, \lambda, \mu, \sigma)$	0.66 (0.09)	1.14 (0.17)	0.05 $(5e^{-3})$	113.90 (0.67)	1.44 (0.08)	178.06	0.040	0.251
$SSCN(\alpha, \lambda, \mu, \sigma)$	2.61 (1.49)	−0.58 (0.14)	112.36 (0.81)	6.40 (0.41)		179.45	0.043	0.295
$BN(\alpha, \beta, \mu, \sigma)$	0.02 $(3e^{-3})$	0.10 (0.02)	113.94 (0.84)	1.45 (0.10)		180.04	0.083	0.456
$KN(\alpha, \beta, \mu, \sigma)$	2.65 (0.70)	0.24 (0.03)	94.00 (1.05)	5.90 (0.45)		182.19	0.068	0.449
$WN(\alpha, \beta, \mu, \sigma)$	1.53 (0.21)	0.10 (0.01)	113.85 (0.79)	2.43 (0.15)		178.64	0.047	0.291

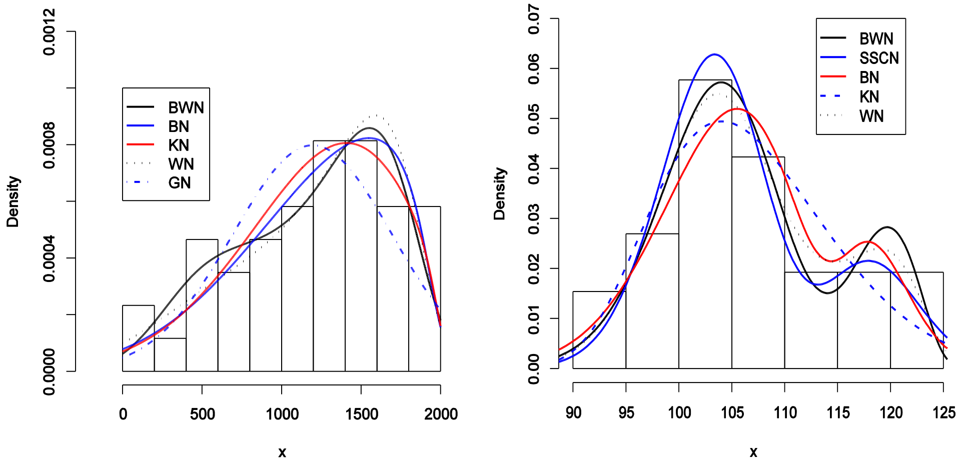

Figure 3: (Left Panel): Fitted Models on Histogram of D2 Data Set, (Right Panel): Fitted Models on Histogram of IQ Data Set.

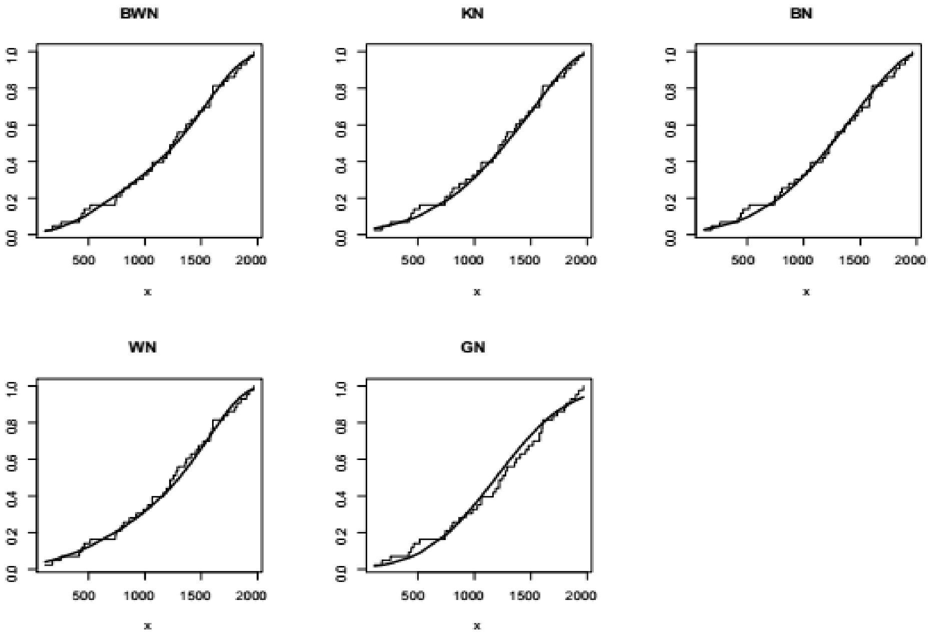

Figure 4: Fitted cdfs on Empirical cdf of D2 Data Set

The strengths of the proposed distributions evident from the two data applications are: their ability to provide better fits (to the D2 and IQ data sets) than four other distributions.

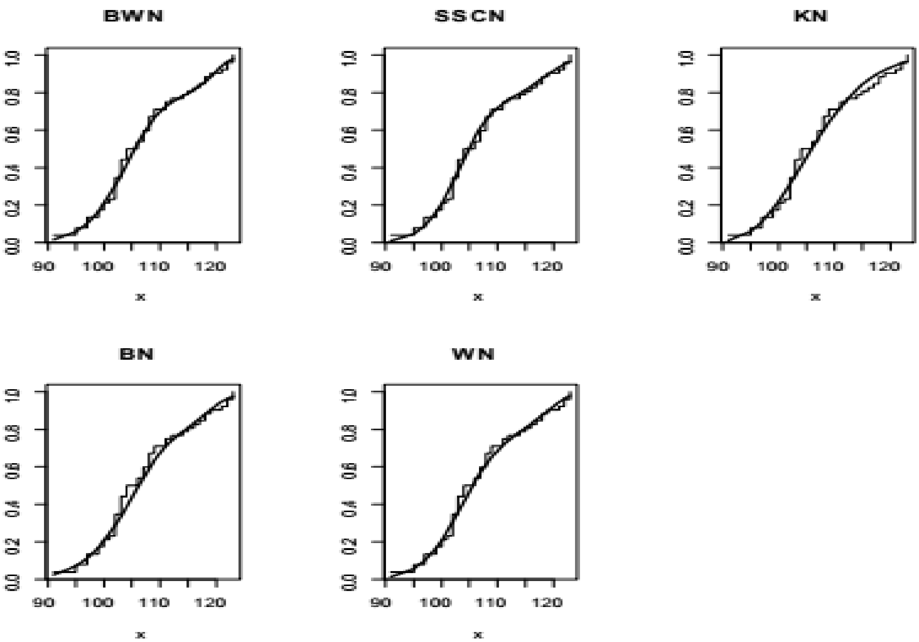

Figure 5: Fitted cdfs on Empirical cdf of IQ Data Set

8. CONCLUSIONS

We define a new class of models, named the beta Weibull-G (BW-G) family of distributions by adding two shape parameters, which generalizes the Weibull-G family. Many well-known distributions emerge as special cases of the BW-G family by using special parameter values. The mathematical properties of the new family including order statistics, entropies, explicit expansions for the order statistics, ordinary and incomplete moments, generating function, stress-strength model, residual and reversed residual lifes. Some characterizations for the new family are presented. The model parameters are estimated by the maximum likelihood method and the observed information matrix is determined. It is shown, by means of two real data sets, that special cases of the BW-G family can give better fits than other models generated by well-known families.

REFERENCES

1. Abouammoh, A.M., Abdulghani, S.A. and Qamber, I.S. (1994). On partial ordering and testing of new better than renewal used class. *Reliability Eng. Systems Safety*, 43, 37-41.
2. Afify A.Z., Alizadeh, M., Yousof, H.M., Aryal, G. and Ahmad, M. (2016a). The transmuted geometric-G family of distributions: theory and applications. *Pak. J. Statist.*, 32, 139-160.
3. Afify A.Z., Cordeiro, G.M., Yousof, H.M., Alzaatreh, A. and Nofal, Z.M. (2016b). The Kumaraswamy transmuted-G family of distributions: properties and applications. *J. Data Sci.*, 14, 245-270.

4. Afify, A.Z., Yousof, H.M. and Nadarajah, S. (2017). The beta transmuted-H family of distributions: properties and applications. *Stasistics and its Inference*, 10, 505-520.
5. Alzaatreh, A., Famoye, F. and Lee, C. (2014). The gamma-normal distribution: Properties and applications. *Computational Statistics and Data Analysis*, 69, 67-80.
6. Alzaatreh, A., Lee, C. and Famoye, F. (2013). A new method for generating families of continuous distributions. *Metron*, 71, 63-79.
7. Alizadeh, M., Cordeiro, G.M., Nascimento, A.D.C. Lima M.D.S. and Ortega, E.M.M. (2016a). Odd-Burr generalized family of distributions with some applications. *Journal of Statistical Computation and Simulation*, 83, 326-339.
8. Alizadeh, M., Rasekhi, M., Yousof, H.M. and Hamedani, G.G. (2016b). The transmuted Weibull G family of distributions. *Hacettepe Journal of Mathematics and Statistics*, Forthcoming.
9. Aryal, G.R. and Yousof, H.M. (2017). The exponentiated generalized-G Poisson family of distributions. *Stochastic and Quality Control*, forthcoming.
10. Azzalini, A. (1985). A class of distributions which include the normal. *Scandinavian Journal of Statistics*, 12, 171-178.
11. Bourguignon, M., Silva, R.B. and Cordeiro, G.M. (2014). The Weibull-G family of probability distributions. *Journal of Data Science*, 12, 53-68.
12. Cordeiro, G.M. and Bager, R.S.B. (2012). The beta power distribution. *Brazilian Journal of Probability and Statistics*, 26, 88-112.
13. Cordeiro, G.M. and de Castro, M. (2011). A new family of generalized distributions. *Journal of Statistical Computation and Simulation*, 81, 883-898.
14. 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.
15. Cordeiro, G.M., Ortega, E.M.M., Popovicʹ B.V. and Pescim, R.R. (2014). The Lomax generator of distributions: Properties, minification process and regression model. *Applied Mathematics and Computation*, 247, 465-486.
16. Eugene, N., Lee, C. and Famoye, F. (2002). Beta-normal distribution and its applications. *Commun. Stat. Theory Methods*, 31, 497-512.
17. Fernandez, C. and Steel, M.F.J. (1998). On Bayesian modeling of fat tails and skewness. *Journal of the American Statistical Association*, 93, 359-371.
18. 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.
19. Glänzel, W. (1990). Some consequences of a characterization theorem based on truncated moments. *Statistics: A Journal of Theoretical and Applied Statistics*, 21, 613-618.
20. Gradshteyn, I.S. and Ryzhik, I.M. (2000). *Table of Integrals, Series and Products* (sixth edition). San Diego: Academic Press.
21. 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.
22. Kotz, S. and Seier, E. (2007). Kurtosis of the Topp Leone distributions. *Interstat*, 1-15.
23. Kumaraswamy, P. (1980). A generalized probability density function for double bounded random processes. *Journal of Hydrology*, 46, 79-88.

24. Kumar, C.S. and Anusree, M.R. (2014). On modified generalized skew normal distribution and some of its properties. *Journal of Statistical Theory and Practice*, 9, 1-17.
25. Lemonte, A.J., Barreto-Souza, W. and Cordeiro, G.M. (1013). The exponentiated Kumaraswamy distribution and its log-transform. *Brazilian Journal of Probability and Statistics*, 27, 31-53.
26. Marshall, A.W. and 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.
27. Ma, Y. and Genton, M.G. (2004). Flexible class of skew-symmetric distributions. *Scandinavian Journal of Statistics*, 31, 459-468.
28. Merovci, F., Alizadeh, M., Yousof, H.M. and Hamedani G.G. (2016). The exponentiated transmuted-G family of distributions: theory and applications. *Commun. Stat. Theory Methods*, Forthcoming.
29. Mudholkar, G.S. and Hutson, A.D. (2000). The epsilon-skew-normal distribution for analyzing near normal data. *Journal of Statistical Planning and Inference*, 83, 291-309.
30. Mudholkar, G.S. and Srivastava, D.K. (1993). Exponentiated Weibull family for analysing bathtub failure rate data. *IEEE Transactions on Reliability*, 42, 299-302.
31. Mudholkar, G.S., Srivastava, D.K. and Freimer, M. (1995). The exponentiated Weibull family: a reanalysis of the bus-motor-failure data. *Technometrics*, 37, 436-445.
32. Nadarajah, S. and Gupta, A.K. (2007). The exponentiated gamma distribution with application to drought data. *Calcutta Statistical Association Bulletin*, 59, 29-54.
33. Nofal, Z.M., Afify, A.Z., Yousof, H.M. and Cordeiro, G.M. (2015). The generalized transmuted-G family of distributions. *Commun. Stat. Theory Methods*, 46, 4119-4136.
34. Rasekhi, M., Chinipardaz, R. and Alavi S.M.R. (2015). A flexible generalization of the skew normal distribution based on a weighted normal distribution. *Statistical Methods and Applications*, Published Online, DOI: 10.1007/s10260-015-0337-4.
35. Roberts, H.V. (1988). *Data Analysis for Managers with Minitab*. Redwood City, CA:Scientific Press.
36. Shirke, D.T. and Kakade, C.S. (2006). On exponentiated lognormal distribution. *International Journal of Agricultural and Statistical Sciences*, 2, 319-326.
37. Yousof, H.M., Afify, A.Z., Alizadeh, M., Butt, N.S., Hamedani, G.G. and Ali, M.M. (2015). The transmuted exponentiated generalized-G family of distributions, *Pak. J. Stat. Oper. Res.*, 11, 44-464.
38. Yousof, H.M., Afify, A.Z., Hamedani, G.G. and Aryal, G. (2016). the Burr X generator of distributions for lifetime data. *Journal of Statistical Theory and Applications*, 16, 1-19.

APPENDIX A

Theorem 1

Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a given probability space and let $H = [d, e]$ be an interval for some $d < e$ $(d = -\infty, e = \infty$ mightaswellbeallowed$)$. Let $X: \Omega \to H$ 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

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

is defined with some real function η. 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 η, particularly

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

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 $\int_H dF = 1$.

APPENDIX B

The elements of the observed information matrix are:

$$U_{aa} = -n\left\{\frac{\Gamma''(a)}{\Gamma(a)} - \psi^2(a) - \Gamma(a+b)[\psi'(a+b) + \psi^2(a+b)]^2 + \psi^2(a+b)\right\},$$

$$U_{ab} = n\psi'(a+b), U_{a\lambda} = \sum_{i=1}^n \frac{m_i}{z_i},$$

$$U_{\lambda\lambda} = \frac{-n}{\lambda^2} - b\sum_{i=1}^n \frac{(\log s_i)^2}{s_i^{-\lambda}} + (a-1)\sum_{i=1}^n \frac{z_i\left(\frac{\partial m_i}{\partial \lambda}\right) - m_i^2}{z_i^2}, U_{a\phi_r} = -\sum_{i=1}^n \frac{w_i}{z_i},$$

$$U_{\lambda\phi_r} = \sum_{i=1}^n \frac{G_r'(x_i;\phi)}{G(x_i;\phi)} + (a-1)\sum_{i=1}^n \frac{z_i(\partial m_i/\partial \phi_r) - m_i w_i}{z_i^2}$$
$$+ \sum_{i=1}^n \frac{G'(x_i;\phi)}{\overline{G}(x_i;\phi)} - b\sum_{i=1}^n \frac{p_{i,r}[1 + \lambda \log s_i]}{s_i^{1-\lambda}},$$

and (for $r = 1, \ldots, q$)

$$U_{\phi_r\phi_r} = (\lambda-1)\sum_{i=1}^n \frac{G(x_i;\phi)G_r''(x_i;\phi) - [G_r'(x_i;\phi)]^2}{G(x_i;\phi)^2}$$
$$+(\lambda+1)\sum_{i=1}^n \frac{\overline{G}(x_i;\phi)G_r''(x_i;\phi) + [G_r'(x_i;\phi)]^2}{\overline{G}(x_i;\phi)^2}$$
$$-b\lambda\sum_{i=1}^n \frac{\left(\frac{\partial p_{i,r}}{\partial \phi_r}\right) + \frac{(\lambda-1)p_{i,r}^2}{s_i}}{s_i^{1-\lambda}} + \sum_{i=1}^n \frac{g(x_i;\phi)g_r''(x_i;\phi) - [g_r'(x_i;\phi)]^2}{g(x_i;\phi)^2}$$
$$+(a-1)\sum_{i=1}^n \frac{z_i\left(\frac{\partial w_i}{\partial \phi_r}\right) - w_i^2}{z_i^2},$$

where $g_r''(x_i;\phi) = \partial^2 g(x_i;\phi)/\partial \phi_r^2$ and $G_r''(x_i;\phi) = \partial^2 G(x_i;\phi)/\partial \phi_r^2$.