The Kumaraswamy-G Poisson Family of Distributions

Manoel Wallace A. Ramos[1,*], Pedro Rafael D. Marinho[2], Gauss M. Cordeiro[2]

Ronaldo V. da Silva [2] and G. G. Hamedani [3]

[1]Instituto Federal de Educação, Ciência e Tecnologia da Paraíba,

João Pessoa, PB, Brazil

[2]Departamento de Estatística, Universidade Federal de Pernambuco,

Recife, PE, Brazil

[3]Department of Mathematics, Statistics and Computer Science,

Marquette University, Milwaukee,WI 53201-1881, USA.

For any baseline continuous G distribution, we propose a new generalized family called the Kumaraswamy-G Poisson (denoted with the prefix "Kw-GP") with three extra positive parameters. Some special distributions in the new family such as the Kw-Weibull Poisson, Kw-gamma Poisson and Kw-beta Poisson distributions are introduced. We derive some mathematical properties of the new family including the ordinary moments, generating function and order statistics. The method of maximum likelihood is used to fit the distributions in the new family. We illustrate its potentiality by means of an application to a real data set.

Keywords: Kumaraswamy distribution; Kumaraswamy-G Poisson distribution; Poisson distribution; Maximum likelihood estimation.

1. Introduction

In recent years, several ways of generating new distributions from classic ones were developed and discussed. The beta-generated family was proposed by Eugene et al. (2002) and further discussed in Zografos and Balakrishnan (2009), who introduced the gamma-generated family of distributions. More recently, Cordeiro and de Castro (2011) defined the *Kumaraswamy-G* ("Kw-G") family as follows. If $G(x)$ denotes the cumulative distribution function (cdf) of a random variable, the Kw-G

*Correspondence: Manoel Wallace A. Ramos, Instituto Federal de Educação, Ciência e Tecnologia da Paraíba, Avenida 1º de Maio, 720 Jaguaribe, 58015430 João Pessoa, PB, Brazil. E-mail: wallace.ifpb@gmail.com

cdf is given by

$$H_{a,b}(x) = 1 - [1 - G(x)^a]^b, \tag{1.1}$$

where $a > 0$ and $b > 0$ are two additional shape parameters to the G distribution, whose role is to govern skewness and tail weights. The probability density function (pdf) corresponding to (1.1) is given by

$$h_{a,b}(x) = a\,b\,g(x)\,G(x)^{a-1}[1 - G(x)^a]^{b-1}, \tag{1.2}$$

where $g(t) = dG(t)/dt$. Equation (1.2) does not involve any special function, such as the incomplete beta function, as is the case of the *beta-G* distribution proposed by Eugene *et al.* (2002). The Kw-G distribution is obtained by adding two shape parameters a and b to the G distribution. The generalization (1.2) contains distributions with unimodal and bathtub shaped hazard rate functions. It also contemplates a broad class of models with monotonic hazard rate function.

We now provide a physical interpretation of the proposed model. Suppose that a system has N subsystems functioning independently at a given time, where N is a truncated Poisson random variable with probability mass function (pmf)

$$p_n = Pr(N = n) = \frac{\lambda^n}{(e^\lambda - 1)\,n!} \tag{1.3}$$

for $n = 1, 2, \ldots$ Suppose that the failure time of each subsystem has the Kw-G distribution defined by the cdf (1.1) for $x > 0$. Further, let Y_i denote the failure time of the ith subsystem and X denote the time to failure of the first out of the N functioning subsystems. We can write $X = \min\{Y_1, \ldots, Y_N\}$. The conditional cdf of X given N is

$$F(x \mid N) = 1 - P(X > x \mid N) = 1 - P(Y_1 > x)^N = 1 - [1 - H_{a,b}(x)]^N.$$

So, the unconditional cdf of X (for $x > 0$) can be expressed as

$$F(x) = \frac{1}{(e^\lambda - 1)} \sum_{n=1}^{\infty} \frac{\lambda^n \left\{1 - [1 - H_{a,b}(x)]^n\right\}}{n!}$$

and then

$$F(x) = \frac{1 - \exp[-\lambda H_{a,b}(x)]}{1 - e^{-\lambda}}. \tag{1.4}$$

Equation (1.4) is called the *Kumaraswamy-G Poisson* ("Kw-GP" for short) family of distributions. Several new models can be generated by considering special distributions for G. The corresponding pdf and hazard rate function (hrf) are

$$f(x) = \frac{\lambda\,h_{a,b}(x)\exp[-\lambda H_{a,b}(x)]}{1 - e^{-\lambda}} \tag{1.5}$$

and

$$\vartheta(x) = \frac{\lambda\, h_{a,b}(x)\exp[-\lambda H_{a,b}(x)]}{-\mathrm{e}^{-\lambda} + \exp[-\lambda H_{a,b}(x)]},\tag{1.6}$$

respectively. Equation (1.5) has three extra parameters $\lambda > 0$, $a > 0$ and $b > 0$ to the parameters of the G distribution. Note that when $x \to \infty$ we have to $H_{a,b}(x) \to 1$. So we have to $\lim_{x\to\infty} F(x) = 1$. If G is continuous right, have $F(x)$ is also continuous right. Note also that for $x \to -\infty$, with $x > 0$ implies $H_{a,b}(x) \to 0$, ie, $\lim_{x\to\infty} F(x) = 0$. Thus, X has cumulative distribution function $F(x)$. Deriving (1.4) have (1.5), ie, (1.5) is the probability density function of the X.

We shall refer to (1.5) as the Kw-GP pdf. If $\lambda \to 0$, $f(x)$ tends to $h_{a,b}(x)$. A random variable with density function (1.5) is denoted by $X \sim$ Kw-GP$(\lambda, a, b, \boldsymbol{\eta})$, where $\boldsymbol{\eta}$ is the vector of the baseline parameters.

The rest of the paper is organized as follows. In Section 2, we present three special models of the Kw-GP family corresponding to the Weibull, gamma and beta distributions. Section 3 provides some general useful expansions for the Kw-GP density function. Moments of the new family are derived in Section 4. Generating function and quantile function are derived in Sections 5 and 6, respectively. The Rényi and Shannon entropies and the reliability are determined in Sections 7 and 8, respectively. Characterizations of the Kw-GP distribution are described in Section 9. Order statistics are studied in Section 10. Maximum likelihood estimation is investigated in Section 11. An application to a real data set is performed in Section 12. Some conclusions are given in Section 13.

2. Special Kw-GP distributions

The Kw-GP family of densities (1.5) allows for greater flexibility of its tails and can be widely applied in many areas of engineering and biology. It will be most tractable when the cdf $G(x)$ and pdf $g(x)$ have simple analytic expressions. Now, we provide some special Kw-GP distributions.

2.1. *Kw-Weibull Poisson (Kw-WP)*

The Weibull cdf with parameters $\beta > 0$ and $c > 0$ is $G(x) = 1 - \mathrm{e}^{-(\beta x)^c}$ for $x > 0$. Correspondingly, the Kw-WP pdf, say Kw-WP$(\lambda, a, b, c, \beta)$, reduces to

$$
\begin{aligned}
f(x) = {}& \frac{\lambda\, a b c\, \beta^c}{\mathrm{e}^\lambda - 1} x^{c-1} \left[1 - \mathrm{e}^{-(\beta x)^c}\right]^{a-1} \left\{1 - \left[1 - \mathrm{e}^{-(\beta x)^c}\right]^a\right\}^{b-1} \times \\
& \exp\left[\lambda \left\{1 - \left[1 - \mathrm{e}^{-(\beta x)^c}\right]^a\right\}^b - (\beta x)^c\right].
\end{aligned}\tag{2.1}
$$

For $c = 1$, we obtain the Kw-exponential Poisson distribution. The Kw-WP distribution for $a = 1$, $c = 1$ and $\lambda \to 0$ corresponds to the exponential distribution with parameters $\beta^* = b\beta$. The hrf corresponding to (2.1) is given by

$$\vartheta(x) = -\frac{\lambda\,abc\,\beta^c\,x^{c-1}\left[1-e^{-(\beta x)^c}\right]^a\left\{1-\left[1-e^{-(\beta x)^c}\right]^a\right\}^{b-1}}{\left[e^{(\beta x)^c}-1\right]\left\{\exp\left[-\lambda\left\{1-\left[1-e^{-(\beta x)^c}\right]^a\right\}^b\right]-1\right\}}.$$

Plots of the Kw-WP density and hrf for selected parameter values are displayed in Figure 1.
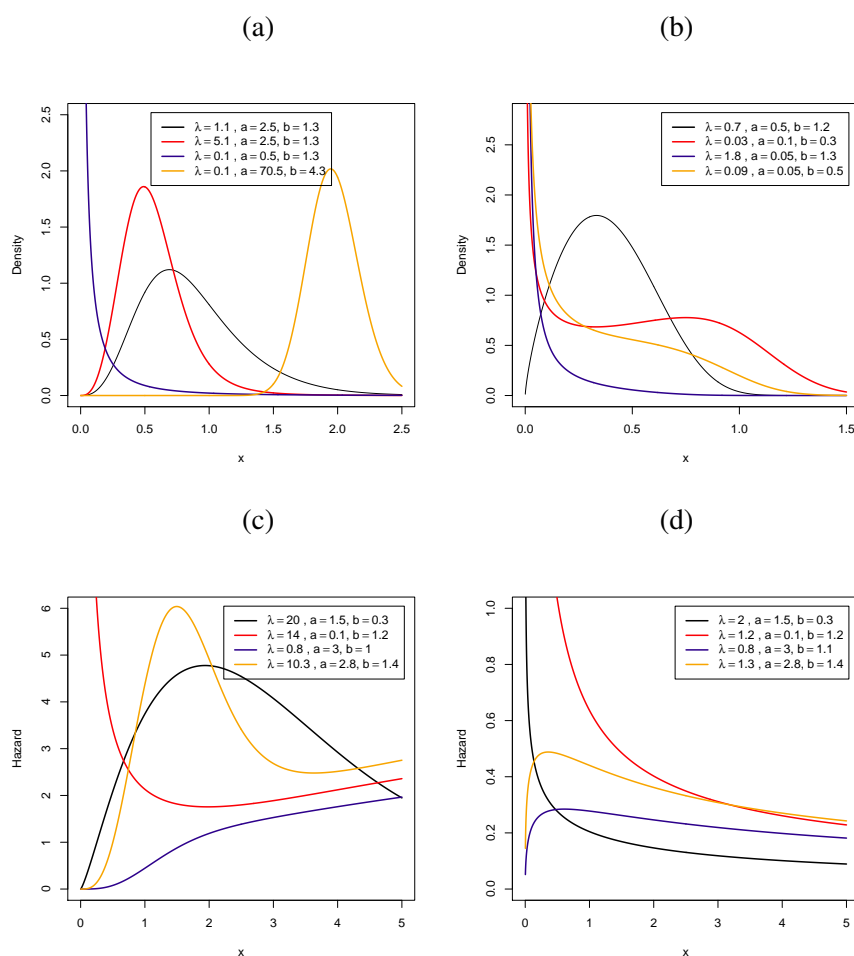
(a) (b)

(c) (d)

Fig. 1. Plots of the density and hrf of the Kw-WP distribution for some parameter values. (a) For $c = 1.5$ and $\beta = 1.2$. (b) For $c = 1.5$ and $\beta = 0.4$. (c) For $c = 1.5$ and $\beta = 0.7$. (d) For $c = 0.5$ and $\beta = 0.6$.

2.2. *Kw-gamma Poisson (Kw-GaP)*

Let Y be a gamma random variable with cdf $G(y) = \gamma(\alpha, \beta x)/\Gamma(\alpha)$ for $y, \alpha, \beta > 0$, where $\Gamma(\cdot)$ is the gamma function and $\gamma(\alpha, y) = \int_0^y t^{\alpha-1} e^{-t} dt$ is the incomplete gamma function. The density

function of a random variable X following the Kw-GaP distribution, say $X \sim \text{Kw-GaP}(\lambda, a, b, \alpha, \beta)$, can be expressed as

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

For $\alpha = 1$, we obtain the Kw-exponential Poisson distribution. For $a = 1$, $\alpha = 1$ and $\lambda \to 0$, the Kw-GaP reduces to the exponential distribution with parameter $\beta^* = b\beta$. The Kw-GaP hrf is given by

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

Plots of the Kw-GaP density and hrf for some parameter values are displayed in Figure 2.

2.3. *Kw-beta Poisson (Kw-BP)*

Consider the beta distribution with positive shape parameters $\alpha > 0$ and $\beta > 0$ and pdf and cdf (for $0 < x < 1$) given by

$$g(x) = \frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1} \qquad \text{and} \qquad G(x) = I_x(\alpha,\beta) = \frac{B_x(\alpha,\beta)}{B(\alpha,\beta)},$$

where $B_x(\alpha,\beta) = \int_0^x w^{\alpha-1}(1-w)^{\beta-1}dw$ is the incomplete beta function, $B(\alpha,\beta) = \int_0^1 w^{\alpha-1}(1-w)^{\beta-1}dw$ is the beta function and $I_x(\alpha,\beta) = \frac{B_x(\alpha,\beta)}{B(\alpha,\beta)}$ is the regularized incomplete beta function. The density of a random variable X having the Kw-BP distribution, say $X \sim \text{Kw-BP}(\lambda, a, b, \alpha, \beta)$, can be expressed as

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

Plots of the Kw-BP density for selected parameter values are displayed in Figure 3.

3. Expansion for the density function

The pdf (1.5) can be expressed as a linear combination of Kw-G density functions. Using the power series for the exponential function, we can rewrite (1.5) as

$$f(x) = h_{a,b}(x) \sum_{k=0}^{\infty} t_k H_{a,b}(x)^k, \tag{3.1}$$
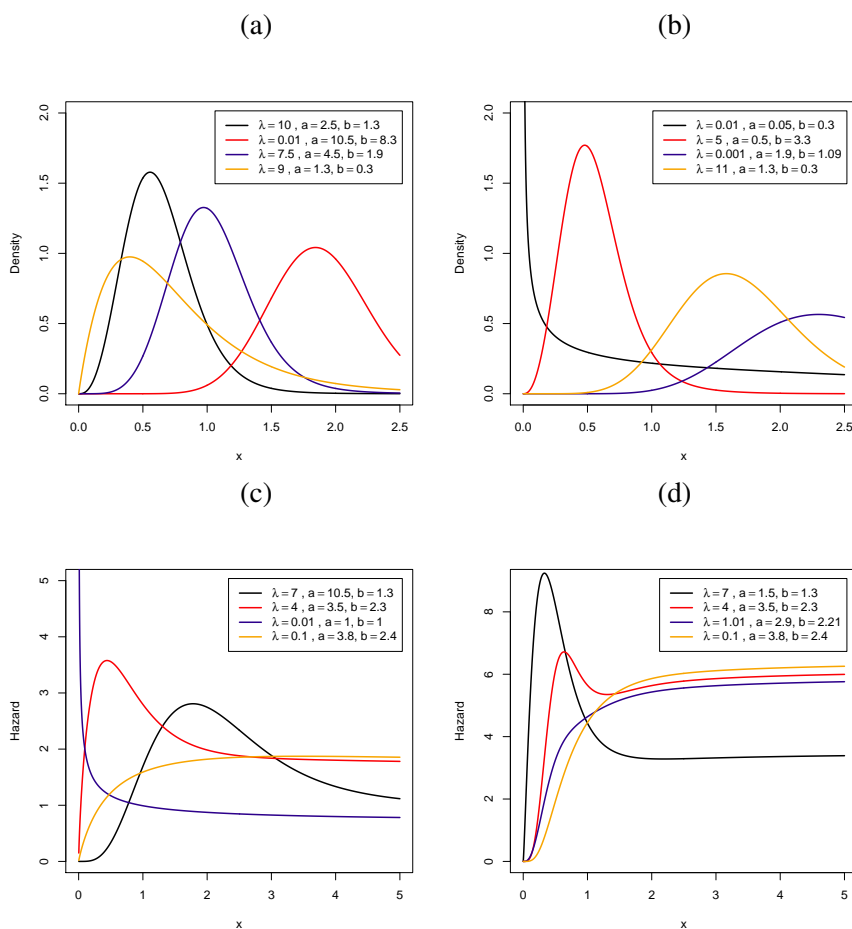
(a)

(b)

(c)

(d)

Fig. 2. Plots of the density and hrf of the Kw-GaP distribution for some parameter values. (a) For $\alpha = 7$ and $\beta = 3.2$. (b) For $\alpha = 0.5$ and $\beta = 0.7$. (c) For $\alpha = 0.5$ and $\beta = 0.7$. (d) For $\alpha = 1.5$ and $\beta = 2.7$.

where

$$t_k = \frac{(-1)^k \lambda^{k+1}}{(1 - e^{-\lambda}) k!}.$$

Substituting (1.1) and (1.2) in equation (3.1) and using the binomial expansion, we obtain

$$f(x) = \sum_{k=0}^{\infty} \omega_k h_{a,(j+1)b}(x), \tag{3.2}$$

where

$$\omega_k = \frac{\lambda^{k+1}}{1 - e^{-\lambda}} \sum_{j=0}^{k} \frac{(-1)^{k+j}}{(j+1)!(k-j)!}.$$
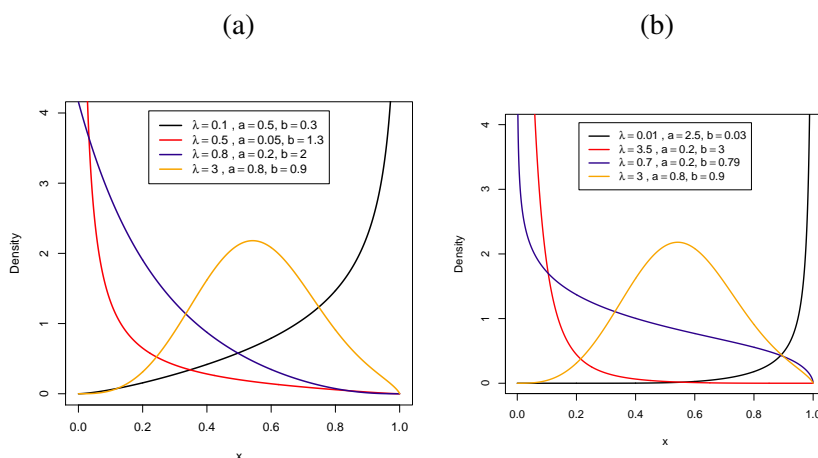
(a)

(b)

Fig. 3. Plots of the Kw-BP density function for some parameter values. (a) For $\alpha = 1.5$ and $\beta = 0.7$. (b) For $\alpha = 3.5$ and $\beta = 1.5$.

It was checked using the Mathematica software (Wolfram, 2003) that $\sum_{k=0}^{\infty} \omega_k = 1$. Note that the finite series $\sum_{j=0}^{k} (-1)^{k+j} / [(j+1)!(k-j)!] = (-1)^k / [(1+k)k!]$. So we need to show that

$$\frac{1}{1-e^{-\lambda}} \sum_{k=0}^{\infty} \frac{(-1)^k \lambda^{1+k}}{(1+k)k!} = \frac{1}{1-e^{-\lambda}} \sum_{k=0}^{\infty} \frac{(-1)^k \lambda^{1+k}}{(k+1)!} = 1.$$

Note also that expanding $(1-e^{-\lambda})$ in Taylor series around zero, we have

$$(1-e^{-\lambda}) = \lambda - \frac{\lambda^2}{2} + \frac{\lambda^3}{6} - \frac{\lambda^4}{24} + \frac{\lambda^5}{120} - \frac{\lambda^6}{720} + \frac{\lambda^7}{5040} + \cdots.$$

If we expand the series term by term, we arrive precisely in the expansion just above. Therefore, we conclude that $\sum_{k=0}^{\infty} \omega_k = 1$.

The cdf corresponding to (3.2) can be expressed as

$$F(x) = \sum_{k=0}^{\infty} \omega_k H_{a,(j+1)b}(x). \tag{3.3}$$

Based on equations (3.2) and (3.3) some structural properties as moments and function generator of moments of the Kw-GP family of distributions can be obtained from well-established properties of the Kw-G distribution. Equations (3.2) and (3.3) can also be expressed as linear combinations of exponentiated-G ("exp-G") distributions.

For an arbitrary baseline cdf $G(x)$, we write $Z \sim$ exp-G(α) if Z follows the exp-G distribution with power parameter α. The cumulative distribution and density functions of Z are given by

$$\Pi_\alpha(x) = G(x)^\alpha \quad \text{and} \quad \pi_\alpha(x) = \alpha \, g(x) \, G(x)^{\alpha-1},$$

respectively. Substituting (1.1) in equation (3.2) and using the binomial expansion, we obtain

$$f(x) = \sum_{l=0}^{\infty} v_l \, \pi_{(l+1)a}(x), \tag{3.4}$$

where

$$v_l = v_l(\lambda, b) = \frac{\lambda \, b}{(1 - e^{-\lambda})(l+1)!} \sum_{k=0}^{\infty} \sum_{j=0}^{k} \frac{(-1)^{k+l+j} \lambda^k \, \Gamma[(j+1)b]}{j!(k-j)!\Gamma[(j+1)b-l]}.$$

Integrating (3.4), we can write

$$F(x) = \sum_{l=0}^{\infty} v_l \, \Pi(x)^{(l+1)a}, \tag{3.5}$$

where $\Pi(x)^{(l+1)a}$ denotes the exp-G cdf with power parameter $(l+1)a$. Equation (3.4) reveals that the Kw-GP density function is a linear combination of exp-G densities. Thus, some structural properties of the Kw-GP family of distributions such as the ordinary and incomplete moments and generating function can be obtained from well-established properties of the exp-G distributions. Equations (3.2)-(3.5) are the main results of this section.

4. Moments

Hereafter, we shall assume that G is the cdf of a random variable Y and that F is the cdf of a random variable X having density function (1.5). The moments of the Kw-GP distribution can be determined from the (r, s)th probability weighted moment (PWM) of Y defined by

$$\tau_{r,s} = E[Y^r \, G(Y)^s] = \int_{-\infty}^{\infty} y^r \, G(y)^s \, g(y) \, dy. \tag{4.1}$$

In fact, from equation (3.4), we can write

$$E(X^r) = \sum_{l=0}^{\infty} v_l^* \, \tau_{r,(l+1)a-1}, \tag{4.2}$$

where

$$v_l^* = v_l^*(\lambda, a, b) = \frac{\lambda \, a \, b}{(1 - e^{-\lambda}) \, l!} \sum_{k=0}^{\infty} \sum_{j=0}^{k} \frac{(-1)^{k+l+j} \lambda^k \, \Gamma[(j+1)b]}{j!(k-j)!\Gamma[(j+1)b-l]}.$$

Thus, the moments of any Kw-GP distribution can be expressed as an infinite weighted sum of the baseline PWMs. A second formula for $\tau_{r,s}$ can be based on the parent quantile function $Q_G(x) = G^{-1}(x)$. Setting $G(x) = u$, we obtain

$$\tau_{r,s} = \int_0^1 Q_G(u)^r \, u^s \, du. \tag{4.3}$$

The ordinary moments of several Kw-GP distributions can be calculated directly from equations (4.2) and (4.3). For example, the moments of the Kw-exponential Poisson (with parameter $\beta > 0$)

are given by

$$E(X^r) = r! \beta^r \sum_{l,m=0}^{\infty} \frac{(-1)^{r+m} v_l^* \binom{(l+1)a-1}{m}}{(m+1)^{r+1}}.$$

5. Moment generating function

The moment generating function (mgf) of X, say $M(t) = E(e^{tX})$, comes from (3.4) as an infinite weighted sum

$$M(t) = \sum_{l=0}^{\infty} v_l M_l(t), \tag{5.1}$$

where $M_l(t)$ is the mgf of $Y_l \sim \exp\text{-}G((l+1)a)$. Hence, for several Kw-GP distributions, $M(t)$ can be immediately determined from the mgf of the G distribution. The mgf of Y_l is given by

$$M_l(t) = (l+1)a \int_0^{\infty} e^{tx} g(x) G(x)^{(l+1)a-1} dx.$$

Setting $G(x) = u$, we can write $M_l(t)$ in terms of the baseline quantile function $Q_G(x)$

$$M_l(t) = (l+1)a \int_0^1 \exp[tQ_G(u)] u^{(l+1)a-1} du. \tag{5.2}$$

The generating function of some Kw-GP distributions can follow immediately from equations (5.1) and (5.2). For example, the mgf's of the Kw-exponencial Poisson (with parameter $\beta > 0$ and $t < \beta^{-1}$) and Kw-standard logistic Poisson (for $t < 1$), where $G(x) = (1 + e^{-x})^{-1}$, are determined as

$$M(t) = \sum_{l=0}^{\infty} v_l^* B[(l+1)a, 1 - t\beta]$$

and

$$M(t) = \sum_{l=0}^{\infty} v_l^* B[1 - t, (l+1)a + t],$$

respectively.

6. Quantile function

The Kw-GP quantile function, say $Q(u) = F^{-1}(u)$, is straightforward to be computed by inverting (1.4) provided a closed-form expression for the quantile function $Q_H(u) = H_{a,b}^{-1}(u)$ is available. From equation (1.4), we can write

$$Q(u) = Q_H\left(-\lambda^{-1}\log[1-(1-e^{-\lambda})u]\right), \qquad (6.1)$$

where

$$H_{a,b}^{-1}(u) = G^{-1}[1-(1-u)^{1/b}]^{1/a}.$$

The quantiles of X are easily obtained from equation (6.1) and the quantiles of the baseline G distribution.

7. Entropies

The entropy of a random variable X with density function $f(x)$ is a measure of variation of the uncertainty. Two popular entropy measures are due to Shannon and Rényi (Shannon, 1951, Rényi, 1961). A large value of the entropy indicates the greater uncertainty in the data. The Rényi entropy is defined by (for $\gamma > 0$ and $\gamma \neq 1$)

$$I_R(\gamma) = \frac{1}{(1-\gamma)}\log\int_0^\infty f^\gamma(x)dx.$$

Based on the pdf (1.5), the Rényi entropy of the Kw-GP distribution is given by

$$\begin{aligned}
I_R(\gamma) &= \frac{1}{(1-\gamma)}\log\int_0^\infty \frac{\lambda^\gamma h_{a,b}^\gamma(x)e^{-\lambda\gamma H_{a,b}(x)}}{(1-e^{-\lambda})^\gamma}dx \\
&= -\frac{\log[(1-e^{-\lambda})^\gamma]}{(1-\gamma)} + \frac{1}{(1-\gamma)}\log\left[\lambda^\gamma\sum_{k=0}^\infty \frac{(-\lambda\gamma)^k}{k!}I_k\right],
\end{aligned} \qquad (7.1)$$

where

$$I_k = \int_0^\infty h_{a,b}^\gamma(x)H_{a,b}^k(x)dx.$$

The integral I_k can be determined numerically for most baseline G distributions.

The Shannon entropy is given by

$$E\{-\log[f(X)]\} = \log\left(\frac{\lambda}{1-e^{-\lambda}}\right) - E\{\log[h_{a,b}(X)]\} + \lambda\,E\,[H_{a,b}(X)].$$

A general expression for $E[H_{a,b}(X)]$ follows by using (3.1) and setting $H_{a,b}(x) = u$

$$\begin{aligned}
E[H_{a,b}(X)] &= \int_0^\infty \sum_{k=0}^\infty t_k\,h_{a,b}(x)H_{a,b}(x)^{k+1}dx \\
&= \sum_{k=0}^\infty t_k\int_0^1 u^{k+1}du = \sum_{k=0}^\infty \frac{t_k}{k+2}.
\end{aligned}$$

The quantity $\delta_X = E\{\log[h_{a,b}(X)]\}$ may be calculated for special forms of $h_{a,b}(x)$. Thus, we obtain

$$E\{-\log[f(X)]\} = \log\left(\frac{\lambda}{1-e^{-\lambda}}\right) - \delta_X + \lambda \sum_{k=0}^{\infty} \frac{t_k}{k+2}. \tag{7.2}$$

8. Reliability

The component fails at the instant that the random stress X_2 applied to it exceeds the random strength X_1, and the component will function satisfactorily whenever $X_1 > X_2$. Hence, $R = P(X_2 < X_1)$ is a measure of component reliability. It has many applications especially in the area of engineering. We derive the reliability R when X_1 and X_2 have independent Kw-GP$(\lambda_1, a_1, b_1, \boldsymbol{\eta})$ and Kw-GP$(\lambda_2, a_2, b_2, \boldsymbol{\eta})$ distributions with the same parameter vector $\boldsymbol{\eta}$ for G. The reliability is defined by

$$R = \int_0^{\infty} f_1(x) F_2(x) dx.$$

The pdf of X_1 and cdf of X_2 are obtained from equations (3.4) and (3.5) as

$$f_1(x) = g(x;\boldsymbol{\eta}) \sum_{l=0}^{\infty} v_l^*(\lambda_1, a_1, b_1) \, G(x;\boldsymbol{\eta})^{(l+1)a_1-1} \quad \text{and} \quad F_2(x) = \sum_{l=0}^{\infty} v_l(\lambda_2, b_2) \, G(x;\boldsymbol{\eta})^{(l+1)a_2},$$

where

$$v_l^*(\lambda_1, a_1, b_1) = \frac{\lambda_1 a_1 b_1}{(1-e^{-\lambda_1})\, l!} \sum_{k=0}^{\infty} \sum_{j=0}^{k} \frac{(-1)^{k+l+j}\, \lambda_1^k\, \Gamma[(j+1)b_1]}{j!(k-j)!\,\Gamma[(j+1)b_1-l]}$$

and

$$v_l(\lambda_2, b_2) = \frac{\lambda_2 b_2}{(1-e^{-\lambda_2})(l+1)!} \sum_{k=0}^{\infty} \sum_{j=0}^{k} \frac{(-1)^{k+l+j}\, \lambda_2^k\, \Gamma[(j+1)b_2]}{j!(k-j)!\,\Gamma[(j+1)b_2-l]}.$$

Hence,

$$R = \sum_{l=0}^{\infty} v_l^*(\lambda_1, a_1, b_1)\, v_l(\lambda_2, b_2) \int_0^{\infty} g(x;\boldsymbol{\eta})\, G(x;\boldsymbol{\eta})^{(l+1)(a_1+a_2)-1} dx.$$

Setting $u = G(x)$, we have

$$R = \sum_{l=0}^{\infty} v_l^*(\lambda_1, a_1, b_1)\, v_l(\lambda_2, b_2) \int_0^{1} u^{(l+1)(a_1+a_2)-1} dx.$$

Hence, the reliability of the Kw-GP family reduces to

$$R = \sum_{l=0}^{\infty} \frac{v_l^*(\lambda_1, a_1, b_1)\, v_l(\lambda_2, b_2)}{(l+1)(a_1+a_2)}. \tag{8.1}$$

For the special case $\lambda_1 = \lambda_2$, $a_1 = a_2$ and $b_1 = b_2$, we have $R = 1/2$.

9. Characterization results

Characterizations of distributions are important to many researchers in the applied fields. An investigator will be vitally interested to know if their model fits the requirements of a particular distribution. To this end, one will depend on the characterizations of this distribution which provide conditions under which the underlying distribution is indeed that particular distribution. Various characterizations of distributions have been established in many different directions. Here, we present characterizations of the Kw-GP distribution with pdf (1.5) in terms of a simple relationship between two truncated moments.

9.1. *Characterizations based on truncated moments*

In this subsection we present characterizations of Kw-GP distribution in terms of a simple relationship between two truncated moments. Our characterization results presented here will employ an interesting result due to Glänzel (1987) (Theorem 9.1, below). The advantage of the characterizations given here is that, *cdf F* need not have a closed form and are given in terms of an integral whose integrand depends on the solution of a first order differential equation, which can serve as a bridge between probability and differential equation.

Theorem 9.1. *Let* $(\Omega, \mathscr{F}, \mathbf{P})$ *be a given probability space and let* $I = [c, d]$ *be an interval for some* $c < d$ $(c = -\infty, d = \infty$ *might as well be allowed). Let* $X : \Omega \to I$ *be a continuous random variable with the distribution function* F *and let* q_1 *and* q_2 *be real functions defined on* I *such that*

$$\frac{E\left[q_2\left(X\right) \mid X \geq x\right]}{E\left[q_1\left(X\right) \mid X \geq x\right]} = \eta\left(x\right), \quad x \in I,$$

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* I. *Finally, assume that the equation* $\eta q_1 = q_2$ *has no real solution in the interior of* I. *Then* F *is uniquely determined by the functions* q_1, q_2 *and* η, *particularly*

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

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

Proposition 9.1. *Let* $X : \Omega \to (0, \infty)$ *be a continuous random variable and let* $q_1(x) = e^{\lambda H_{a,b}(x)}$ *and* $q_2(x) = q_1(x)\left[1 - (G(x))^a\right]$ *for* $x \in (0, \infty)$. *Then, pdf of* X *is* (1.5) *if and only if the function*

η defined in Theorem 9.1 has the form

$$\eta(x) = \frac{b}{b+1}\left[1 - (G(x))^a\right], \quad x > 0. \tag{9.1}$$

Proof. Proof. Let X have *pdf* (1.5), then

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

and

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

and

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

Conversely, if $\eta(x)$ is given by (9.1), then

$$s'(x) = \frac{\eta'(x)q_1(x)}{\eta(x)q_1(x) - q_2(x)} = \frac{abg(x)(G(x))^{a-1}}{\left[1 - (G(x))^a\right]},$$

from which we obtain

$$s(x) = -\log\left\{\left[1 - (G(x))^a\right]^b\right\}, \quad x > 0.$$

Now, in view of Theorem 9.1, X has *cdf* (1.4) and *pdf* (1.5). $\qquad\square$

Corollary 9.1. *Let* $X : \Omega \to (0, \infty)$ *be a continuous random variable and let* q_1 *be as in Proposition 9.1. Then, pdf of* X *is* (1.5) *if and only if there exist function* q_2 *and* η *defined in Theorem 9.1 satisfying the following differential equation*

$$s'(x) = \frac{\eta'(x)q_1(x)}{\eta(x)q_1(x) - q_2(x)} = \frac{abg(x)(G(x))^{a-1}}{\left[1 - (G(x))^a\right]}, \quad x > 0. \tag{9.2}$$

Remark 9.1. (a) The general solution of the differential equation (9.2) is

$$\eta(x) = \left[1 - (G(x))^a\right]^{-b}\left[-\int q_2(x)\,abg(x)(G(x))^{a-1}\left[1 - (G(x))^a\right]^{b-1}e^{-\lambda H_{a,b}(x)}dx + D\right],$$

for $x > 0$, where D is a constant. One set of function (q_1, q_2, η) satisfying the above equation is given in Proposition 9.1 for $D = 0$.

(b) Clearly there are other triplet of functions (q_1, q_2, η) satisfying the conditions of Theorem 9.1. We presented one such pair in Proposition 9.1.

10. Order Statistics

Order statistics make their appearance in many areas of statistical theory and practice. The density $f_{i:n}(x)$ of the ith order statistic, for $i = 1, \ldots, n$, from independent and identically distributed random variables $X_1, \ldots, X_n$ from the Kw-GP distribution is given by

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

We can write from equations (1.4) and (1.5)

$$f_{i:n}(x) = \frac{1}{B(i, n-i+1)} \sum_{r=0}^{n-i} \sum_{s=0}^{r+i-1} (-1)^{r+s} \binom{n-i}{r} \binom{r+i-1}{s}$$
$$\times \frac{\lambda\, h_{a,b}(x)\, e^{-\lambda(s+1)H_{a,b}(x)}}{(1-e^{-\lambda})^{r+1}}.$$

Hence,

$$f_{i:n}(x) = \sum_{r=0}^{n-i} \sum_{s=0}^{r+i-1} p_{r,s}\, f_{\lambda(s+1)}(x), \tag{10.1}$$

where

$$p_{r,s} = \frac{(-1)^{r+s}[1 - e^{-\lambda(s+1)}]}{(s+1)\,(1-e^{-\lambda})^{r+1}\, B(i, n-i+1)}$$

and $f_{\lambda(s+1)}(x)$ is the Kw-GP density function with parameters $\lambda(s+1)$, a and b. Equation (10.1) is the main result of this section. It reveals that the pdf of the Kw-GP order statistics is a double linear combination of Kw-GP density functions. So, several mathematical quantities of the Kw-GP order statistics (ordinary, incomplete and factorial moments, mgf, mean deviations and several others) can be obtained from those quantities of the Kw-GP family of distributions.

11. Maximum Likelihood Estimation

We determine the maximum likelihood estimates (MLEs) of the parameters of the Kw-GP distribution from complete samples only. Let $x_1, \ldots, x_n$ be a random sample of size n from the Kw-GP$(\lambda, a, b, \boldsymbol{\eta})$ distribution, where $\boldsymbol{\eta}$ is a $p \times 1$ vector of unknown parameters in the parent distribution $G(x; \boldsymbol{\eta})$. The log-likelihood function for the vector of parameters $\theta = (\lambda, a, b, \boldsymbol{\eta}^T)^T$ can be

expressed as

$$l(\theta) = n\log(\lambda \, ab) - n\log(1 - e^{-\lambda}) + \sum_{i=1}^{n} \log[g(x_i; \boldsymbol{\eta})] + (a-1) \sum_{i=1}^{n} \log[G(x_i; \boldsymbol{\eta})]$$

$$+ (b-1) \sum_{i=1}^{n} \log[1 - G(x_i; \boldsymbol{\eta})^a] - \lambda \sum_{i=1}^{n} \{1 - [1 - G(x_i; \boldsymbol{\eta})^a]^b\}. \tag{11.1}$$

The log-likelihood can be maximized either directly by using the SAS (Proc NLMixed), the MaxBFGS routine in the matrix programming language Ox (see, Doornik, 2006) or by solving the nonlinear likelihood equations obtained by differentiating (11.1). The components of the score vector $U(\theta)$ are

$$U_\lambda(\theta) = \frac{n}{\lambda} - \frac{n e^{-\lambda}}{1 - e^{-\lambda}} - \sum_{i=1}^{n} \{1 - [1 - G(x_i; \boldsymbol{\eta})^a]^b\},$$

$$U_a(\theta) = \frac{n}{a} + \sum_{i=1}^{n} \log[G(x_i; \boldsymbol{\eta})] + (b-1) \sum_{i=1}^{n} - \frac{G(x_i; \boldsymbol{\eta})^a \log[G(x_i; \boldsymbol{\eta})]}{1 - G(x_i; \boldsymbol{\eta})^a} - \lambda \sum_{i=1}^{n} b \, G(x_i; \boldsymbol{\eta})^a$$

$$\times [1 - G(x_i; \boldsymbol{\eta})^a]^{b-1} \log[G(x_i; \boldsymbol{\eta})],$$

$$U_b(\theta) = \frac{n}{b} + \sum_{i=1}^{n} \log[1 - G(x_i; \boldsymbol{\eta})^a] + \lambda \sum_{i=1}^{n} [1 - G(x_i; \boldsymbol{\eta})^a]^b \log[1 - G(x_i; \boldsymbol{\eta})^a],$$

$$U_{\boldsymbol{\eta}_j}(\theta) = \sum_{i=0}^{n} \left\{ \frac{[\dot{g}(x_i; \boldsymbol{\eta})]_{\eta_j}}{g(x_i; \boldsymbol{\eta})} + (a-1) \frac{[\dot{G}(x_i; \boldsymbol{\eta})]_{\eta_j}}{G(x_i; \boldsymbol{\eta})} - (b-1) \frac{a \, G(x_i; \boldsymbol{\eta})^{a-1} [\dot{G}(x_i; \boldsymbol{\eta})]_{\eta_j}}{1 - G(x_i; \boldsymbol{\eta})^a} \right.$$

$$\left. - \lambda \, ab \, G(x_i; \boldsymbol{\eta})^{a-1} [1 - G(x_i; \boldsymbol{\eta})^a]^{b-1} [\dot{G}(x_i; \boldsymbol{\eta})]_{\eta_j} \right\},$$

where $[\dot{g}(x_i; \boldsymbol{\eta})]_{\eta_j} = \frac{\partial g(x_i; \boldsymbol{\eta})}{\partial \eta_j}$ and $[\dot{G}(x_i; \boldsymbol{\eta})]_{\eta_j} = \frac{\partial G(x_i; \boldsymbol{\eta})}{\partial \eta_j}$ for $j = 1, \ldots, p$. For interval estimation on the model parameters, we require the observed information matrix $J(\theta)$. Let $\widehat{\theta}$ be the MLE of θ. Under standard regular conditions (Cox and Hinkley, 1974) which are fulfilled for the proposed model whenever the parameters are in the interior of the parameter space, we can approximate the distribution of $\sqrt{n}(\widehat{\theta} - \theta)$ by the multivariate normal $N_{(p+3)}(0, K(\theta)^{-1})$, where $K(\theta) = \lim_{n \to \infty} J_n(\theta)$ is the information matrix and p is the number of parameters of the G distribution. Based on the approximate multivariate normal $N_{(p+3)}(0, J(\widehat{\theta})^{-1})$ distribution of $\widehat{\theta}$, where $J(\widehat{\theta})$ is the observed information matrix evaluated at $\widehat{\theta}$, we can construct approximate confidence regions for the model parameters. For the general case, there is no closed formula for the estimators of maximum likelihood. In a practical problem, obtaining the maximum likelihood estimates for a particular G must be obtained numerically.

12. Application

Here, we present an application of the five parameter Kw-WP distribution to a real data set for illustrative purposes. This application indicates the flexibility of the new distribution in modeling positive data. The computations were performed using the package AdequacyModel in R developed by the authors Cícero Dias and Pedro Marinho. The real data set represents the remission times (in months) of a random sample of 128 bladder cancer patients (Lee and Wang, 2003):

0.08, 2.09, 3.48, 4.87, 6.94, 8.66, 13.11, 23.63, 0.20, 2.23, 3.52, 4.98, 6.97, 9.02, 13.29, 0.40, 2.26, 3.57, 5.06, 7.09, 9.22, 13.80, 25.74, 0.50, 2.46, 3.64, 5.09, 7.26, 9.47, 14.24, 25.82, 0.51, 2.54, 3.70, 5.17, 7.28, 9.74, 14.76, 26.31, 0.81, 2.62, 3.82, 5.32, 7.32, 10.06, 14.77, 32.15, 2.64, 3.88, 5.32, 7.39, 10.34, 14.83, 34.26, 0.90, 2.69, 4.18, 5.34, 7.59, 10.66, 15.96, 36.66, 1.05, 2.69, 4.23, 5.41, 7.62, 10.75, 16.62, 43.01, 1.19, 2.75, 4.26, 5.41, 7.63, 17.12, 46.12, 1.26, 2.83, 4.33, 5.49, 7.66, 11.25, 17.14, 79.05, 1.35, 2.87, 5.62, 7.87, 11.64, 17.36, 1.40, 3.02, 4.34, 5.71, 7.93, 11.79, 18.10, 1.46, 4.40, 5.85, 8.26, 11.98, 19.13, 1.76, 3.25, 4.50, 6.25, 8.37, 12.02, 2.02, 3.31, 4.51, 6.54, 8.53, 12.03, 20.28, 2.02, 3.36, 6.76, 12.07, 21.73, 2.07, 3.36, 6.93, 8.65, 12.63, 22.69.

Table 1 lists the maximum likelihood estimates (and the corresponding standard errors in parentheses) of the unknown parameters of the Kw-WP, Kumaraswamy-Weibull (Kw-W), beta Weibull (BW), exponentiated Weibull (EW) and Weibull (W) models for the remission times data. It is important to emphasize that the EW model is very popular for the analysis of lifetime data.

Table 1. MLEs (standard errors in parentheses).

Distributions	Estimates				
Kw-WP (λ, a, b, c, β)	3.63748	1.84467	2.14003	0.79822	0.03756
	(2.58073)	(2.76562)	(11.12893)	(1.06864)	(0.11321)
Kw-W (a, b, c, β)	4.12327	2.94308	0.45855	0.21630	
	(5.83511)	(8.10936)	(0.51350)	(0.24837)	
BW (a, b, c, β)	2.73456	0.90765	0.66618	0.32174	
	(1.59435)	(1.49643)	(0.24362)	(0.43221)	
EW (a, c, β)	2.79620	0.65439	0.29890		
	(1.26290)	(0.13459)	(0.16875)		
W (c, β)	1.04782	0.10459			
	(0.06757)	(0.00933)			

Next, we shall apply formal goodness-of-fit tests in order to verify which distribution fits better to these data. We consider the Cramér-von Mises (W^*) and Anderson-Darling (A^*) statistics described in Chen and Balakrishnan (1995). In general, the smaller the values of these statistics, the better the fit to the data.

The statistics A^* and W^* for all the models are listed in Table 2 for the remission time data. First, the proposed Kw-WP model fits these data better than the other models according to the statistics A^* and W^*. The Kw-WP model may be an interesting alternative to other models available in the literature for modeling positive real data. The QQ-plots for the Kw-WP, Kw-W, EW and W distributions are displayed in Figure 4. Based on these plots, the Kw-WP model outperforms the other models.

Table 2. Statistics A^* and W^*.

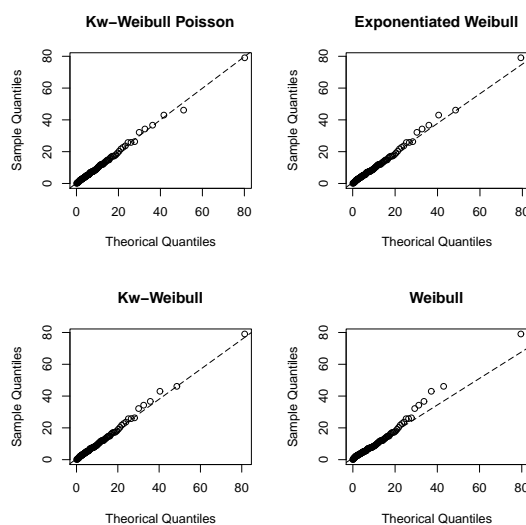| Distributions	A^*	W^*
Kw-WP	**0.14942**	**0.02250**
Kw-W	0.27317	0.04147
BW	0.28824	0.04362
EW	0.28847	0.04367
W	0.78648	0.13137

Fig. 4. QQ-plots.

13. Conclusions

We propose a new family of continuous distributions called the Kumaraswamy-G Poisson ("Kw-GP" for short) family of distributions. Based on this generalization, several new models can be generated by considering special distributions for G. We demonstrate that the probability density function (pdf) of any Kw-GP distribution can be expressed as a linear combination of exponentiated-G density functions. This result allows us to derive general explicit expressions for several measures of the Kw-GP distributions such as the ordinary, incomplete moments and generating function. Further, we demonstrate that the pdf of the Kw-GP order statistics can be expressed as a linear combination of Kw-GP density functions. We discuss maximum likelihood estimation and inference on the model parameters based on Cramér-von Mises and Anderson-Darling statistics. An application of the new class of distributions to a real data set demonstrates the potentiality of the new family. We hope this generalization may attract several applications in statistics.

References

[1] Chen, G., Balakrishnan, N. (1995). A general purpose approximate goodness-of-fit test. *Journal of Quality Technology*, **27**, 154-161.

[2] Cordeiro, G. M. and de Castro M. (2011). A new family of generalized distributions. *Journal of Statistical Computation and Simulation*, **81**, 883-898.

[3] Cox, D.R., Hinkley, D.V. (1974). *Theoretical Statistics*. Chapman and Hall, London.

[4] Doornik, J.A. (2006). *An Object-Oriented Matrix Language − Ox 4*, 5th ed. Timberlake Consultants Press: London.

[5] Eugene, N., Lee, C. and Famoye, F. (2002). Beta-normal distribution and its applications. *Communications in Statistics-Theory and Methods*, **31**, 497-512.

[6] 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.*

[7] Lee, E.T., Wang, J.W. (2003). *Statistical Methods for Survival Data Analysis*, third ed. Wiley, New York.

[8] R Development Core Team. (2012). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna.

[9] Rényi, A. (1961). On measures of entropy and information, In: Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability, Volume I, pp. 547-561, University of California Press, Berkeley.

[10] Shannon, C.E. (1951). Prediction and entropy of printed English. *The Bell System Technical Journal*, **30**, 50-64.

[11] Wolfram, S. (2003). *The Mathematica Book*. 5th ed. Cambridge University Press, London.

[12] Zografos, K. and Balakrishnan, N. (2009). On families of beta- and generalized gamma-generated distributions and associated inference. *Statistical Methodology*, **6**, 344-362.