Journal of Statistical Theory and Applications
Vol. **18(4)**, December 2019, *pp.* 329–342
DOI: https://doi.org/10.2991/jsta.d.191030.001; ISSN 1538-7887
https://www.atlantis-press.com/journals/jsta

# The Generalized Kumaraswamy-G Family of Distributions

Zohdy M. Nofal[1], Emrah Altun[2], Ahmed Z. Afify[1,*], M. Ahsanullah[3]

[1] *Department of Statistics, Mathematics and Insurance, Benha University, Benha, Egypt*
[2] *Department of Statistics, Bartin University, Bartin, Turkey*
[3] *Department of Management Sciences, Rider University, Lawrenceville, NJ, USA*

**ARTICLE INFO**

*Article History*

Received 06 Jun 2018
Accepted 19 Jan 2019

*Keywords*

Kumaraswamy-G Family
Maximum Likelihood
Order Statistics
Regression

2000 Mathematics Subject Classification: 62E10

**ABSTRACT**

We propose a new class of continuous distributions called the *generalized Kumaraswamy-G* family which extends the Kumaraswamy-G family defined by Cordeiro and de Castro [1]. Some special models of the new family are provided. Some of its mathematical properties including explicit expressions for the ordinary and incomplete moments, generating function, Rényi entropy, order statistics and characterizations are derived. The new location-scale regression model is introduced based on the new generated distribution. The maximum likelihood is used for estimating the model parameters. The flexibility of the generated family is illustrated by means of two applications to real data sets.



## 1. INTRODUCTION

Recently, the interest in developing more flexible generators remains strong. Many generalized distributions have been developed over the past decades for modeling data in several areas such as biological studies, environmental sciences, economics, engineering, finance and medical sciences. There has been an increased interest in defining new generated families of univariate distributions by introducing additional shape parameters to the baseline model. For example, the Marshall-Olkin-$G$ [2], beta-$G$ [3], Kumaraswamy-$G$ (K-$G$) [1], transmuted geometric-$G$ [4], beta transmuted-$H$ [5] and the generalized transmuted-$G$ [6] families. However, in many applied areas, there is a clear need for extending forms of the classical models.

The generated distributions have attracted several statisticians to develop new models because the computational and analytical facilities available 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.

Consider a baseline cumulative distribution function (cdf) $G(x;\varphi)$ and probability density function (pdf) $g(x;\varphi)$ depending on a parameter vector $\varphi$, where $\varphi = (\varphi_k) = (\varphi_1, \varphi_2, ...)$. Thus, Cordeiro and de Castro [1] defined the K-$G$ family by the cdf and pdf given by

$$F(x;a,b,\varphi) = 1 - \left[1 - G(x;\varphi)^a\right]^b \tag{1}$$

and

$$f(x;a,b,\varphi) = abg(x;\varphi)\,G(x;\varphi)^{a-1}\left[1 - G(x;\varphi)^a\right]^{b-1}, \tag{2}$$

respectively, where $g(x) = dG(x)/dx$ and $a$ and $b$ are two additional positive shape parameters. 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. Further details can be found in Cordeiro and de Castro [1].

In this paper, we define and study a new family of distributions by adding one extra shape parameter in (1) to provide more flexibility to the generated family. To this end, we construct a new generator so-called the *generalized Kumaraswamy-G* (GK-$G$) family and give a

*Corresponding author. Email: ahmed.afify@fcom.bu.edu.eg

comprehensive description of some of its mathematical properties. We hope that the new model will attract wider applications in reliability, engineering and other areas of research.

The cdf of the GK-$G$ family is defined (for $x > 0$) by

$$F(x; a, b, \alpha, \varphi) = \frac{1 - \left[1 - \alpha G(x; \varphi)^a\right]^b}{1 - (1 - \alpha)^b}. \tag{3}$$

The corresponding pdf of (3) is given by

$$f(x; a, b, \alpha, \varphi) = \frac{\alpha a b g(x; \varphi)}{1 - (1 - \alpha)^b} G(x; \varphi)^{a-1} \left[1 - \alpha G(x; \varphi)^a\right]^{b-1}, \tag{4}$$

where $0 < \alpha \leq 1$, $a > 0$ and $b > 0$ are shape parameters.

Henceforth, a random variable $X$ having the density function (4) is denoted by $X \sim$GK-$G(a, b, \alpha, \varphi)$.

The hazard rate function (hrf) of $X$, say $\tau(x)$, is given by

$$\tau(x) = \frac{\alpha a b g(x; \varphi) G(x; \varphi)^{a-1} \left[1 - \alpha G(x; \varphi)^a\right]^{b-1}}{\left[1 - \alpha G(x; \varphi)^a\right]^b - (1 - \alpha)^b}.$$

Some special cases of the new family are listed in Table 1.

The rest of the paper is outlined as follows. In Section 2, three special models of GK-G family including Weibull, log-logistic and gamma are presented. In Section 3, some of mathematical properties of the proposed family including linear representation, ordinary and incomplete moments, mean deviations, moment generating function (mgf), Rényi entropy and order statistics are obtained. Maximum likelihood estimation of the model parameters is investigated in Section 4. In Section 5, we provide a simulation study to evaluate the performance of the maximum likelihood method in estimating the parameters of the GK-G family. The log-generalized Kumaraswamy-Weibull regression model is defined in Section 6. Section 7 is devoted to applications to prove empirically the flexibility of the proposed models. Finally, some concluding remarks are given in Section 8.

## 2. SPECIAL MODELS

In this section, we provide four special models of the GK-$G$ family, namely, GK-Weibull, GK-log logistic and GK-gamma distributions. These sub-models generalize important existing distributions in the literature.

### 2.1. The GK-Weibull (GKW) Distribution

The Weibull (W) distribution, with positive parameters $\lambda$ and $\beta$, has pdf and cdf given (for $x > 0$) by $g(x) = \lambda \beta x^{\beta-1} e^{-\lambda x^\beta}$ and $G(x) = 1 - e^{-\lambda x^\beta}$, respectively. Then, the GKW pdf reduces to

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

The GKW distribution reduces to the GK-exponential (GKE) distribution when $\beta = 1$. Also, when $a = b = 0$, it reduces to the W distribution. Figure 1 displays some possible shapes of the density and hazard rate functions of this distribution.

**Table 1** Sub-models of the GK-$G$ family.

| $a$ | $b$ | $\alpha$ | Reduced Model | Authors |
|---|---|---|---|---|
| $a$ | $b$ | 1 | K-$G$ family | Cordeiro and de Castro [1] |
| 1 | $b$ | $\alpha$ | Ex-$G$ family | New |
| $a$ | 1 | − | exp-$G$ family | Gupta *et al.* [7] |
| 1 | 1 | − | $G(x; \varphi)$ | – |

## 2.2. The GK-Log Logistic (GKLL) Distribution

The log-logistic (LL) distribution with positive parameters $\lambda$ and $\beta$ has pdf and cdf given by $g(x) = \beta\lambda^{-\beta}x^{\beta-1}\left[1+\left(\frac{x}{\lambda}\right)^{\beta}\right]^{-2}$ (for $x > 0$) and $G(x) = 1-\left[1+\left(\frac{x}{\lambda}\right)^{\beta}\right]^{-1}$, respectively. Then, the pdf of the GKLL distribution is given by

$$
\begin{aligned}
f(x) = & \frac{\alpha ab\beta\lambda^{-\beta}x^{\beta-1}}{1-(1-\alpha)^{b}}\left[1+\left(\frac{x}{\lambda}\right)^{\beta}\right]^{-2}\left\{1-\left[1+\left(\frac{x}{\lambda}\right)^{\beta}\right]^{-1}\right\}^{a-1} \\
& \times\left(1-\alpha\left\{1-\left[1+\left(\frac{x}{\lambda}\right)^{\beta}\right]^{-1}\right\}^{a}\right)^{b-1}.
\end{aligned}
$$

The GKLL model reduces to the LL distribution when $a = b = 1$. Plots of the density and hazard rate functions of the GKLL distribution are displayed in Figure 2 for some parameter values.

## 2.3. The GK-Gamma (GKGa) Distribution

By taking $G(x)$ and $g(x)$ in (4) to be the cdf $G(x) = \gamma(\lambda, x/\beta)/\Gamma(\lambda)$ and the pdf $g(x) = x^{\lambda-1}e^{-x/\beta}/\beta^{\lambda}\Gamma(\lambda)$ of the gamma (Ga) distribution, where $\lambda > 0$ is a shape parameter and $\beta > 0$ is a scale parameter. Then, the pdf of the GKGa (for $x > 0$) reduces to

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

This distribution reduces to the Ga distribution if $a = b = 1$. For $\lambda = 1$, we obtain the GK-exponenial (GKE) distribution. Figure 3 displays plots of the density and hazard rate functions for the GKGa distribution for selected parameter values.

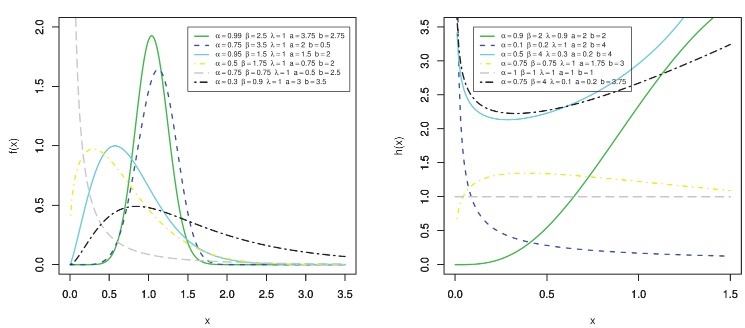


**Figure 1** pdf (left) and hrf (right) plots of GK-Weibull (GKW) distribution.

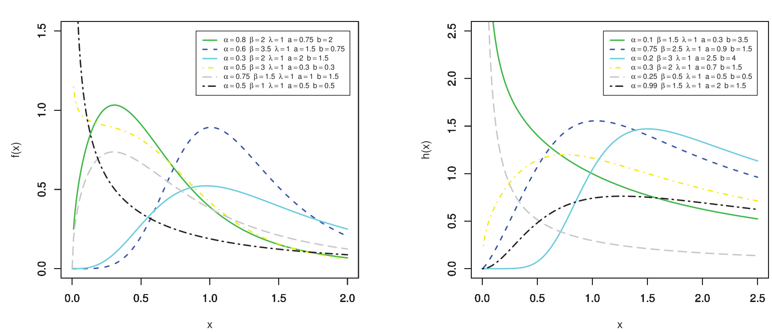


**Figure 2** pdf (left) and hrf (right) plots of GK-log logistic (GKLL) distribution.

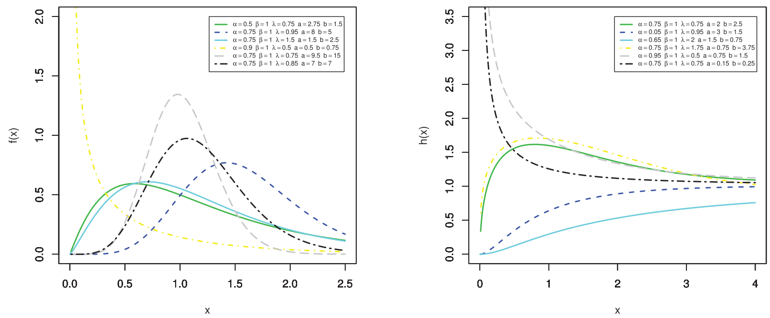


**Figure 3** pdf (left) and hrf (right) plots of GK-gamma (GKGa) distribution.

# 3. MATHEMATICAL PROPERTIES

## 3.1. Linear Representation

In this section, we provide a useful representation for the GK-$G$ pdf. Consider the power series, for $|z| < 1$ and $\rho > 0$ real non-integer,

$$(1-z)^{\rho-1} = \sum_{k=0}^{\infty} (-1)^k \binom{\rho-1}{k} z^k. \tag{5}$$

After applying the power series (5) to (4), we obtain

$$f(x) = \frac{b}{1-(1-\alpha)^b} \sum_{k=0}^{\infty} (-1)^k \alpha^{k+1} \binom{b-1}{k} ag(x)\, G(x)^{a(k+1)-1}.$$

Further, we can write the last equation as

$$f(x) = \sum_{k=0}^{\infty} \upsilon_k\, h_{a(k+1)}(x), \tag{6}$$

where

$$\upsilon_k = \frac{(-1)^k \alpha^{k+1}}{1-(1-\alpha)^b} \binom{b}{k+1}$$

and $h_{a(k+1)}(x) = a(k+1)\ g(x)\ G(x)^{a(k+1)-1}$ is the exp-$G$ density with power parameter $a(k+1) > 0$.

Thus, several mathematical properties of the GK-$G$ family can be derived from those properties of the exp-$G$ family. For example, the ordinary and incomplete moments and mgf of $X$ can be obtained directly from those of the exp-$G$ class.

The cdf of the GK-$G$ family can also be expressed as a mixture of exp-$G$ densities. By integrating (6), we obtain the same linear representation

$$F(x) = \sum_{k=0}^{\infty} \upsilon_k\, H_{a(k+1)}(x),$$

where $H_{(k+1)a}(x)$ is the cdf of the exp-$G$ family with power parameter $(k+1)\,a$.

The formulae derived throughout the paper can be easily handled in most symbolic computation software platforms such as Maple, Mathematica and Matlab because of their ability to deal with analytic expressions of formidable size and complexity. Established explicit expressions to evaluate statistical measures can be more efficient than computing them directly by numerical integration. We have noted that the infinity limit in these sums can be substituted by a large positive integer such as 50 for most practical purposes.

## 3.2. Quantile Function

The quantile function (qf) of $X$, say $Q(u) = F^{-1}(u)$, can be obtained by inverting (3) numerically and it is given by

$$Q(u) = G^{-1}\left\{\alpha^{-1}\left[1-(1-ud)^{\frac{1}{b}}\right]\right\}^{\frac{1}{a}},$$

where $d = 1-(1-\alpha)^b$.

### 3.3. Moments

Hereafter, $Y_{(k+1)a}$ denotes the exp-$G$ distribution with power parameter $a\,(k+1)$. The $r$th moment of $X$, say $\mu_r'$, follows from (6) as

$$\mu_r' = E\left(X^r\right) = \sum_{k=0}^{\infty} \upsilon_k \, E\left(Y_{a(k+1)}^r\right).$$

### 3.4. Generating Function

Here, we provide two formulae for the mgf $M_X\,(t) = E\left(\mathrm{e}^{t\,X}\right)$ of $X$. Clearly, the first one can be derived from (6 ) as

$$M_X\,(t) = \sum_{k=0}^{\infty} \upsilon_k \; M_{a(k+1)}\,(t)\,,$$

where $M_{a(k+1)}\,(t)$ is the mgf of $Y_{a(k+1)}$. Hence, $M_X\,(t)$ can be determined from the exp-$G$ generating function.

A second formula for $M_X\,(t)$ follows from (6) as

$$M_X\,(t) = \sum_{k=0}^{\infty} \upsilon_k \; \tau\,(t,k)\,,$$

where $\tau\,(t,k) = \int_0^1 \exp\left[t\;Q_G\,(u)\right]\; u^{a(k+1)-1} du$ and $Q_G\,(u)$ is the qf corresponding to $G\,(x;\phi)$, i.e., $Q_G\,(u) = G^{-1}\,(u;\phi)$.

### 3.5. Incomplete Moments

The $s$th incomplete moment, say $\varphi_s\,(t)$, of $X$ can be expressed from (6) as

$$\varphi_s\,(t) = \int_{-\infty}^{t} x^s f(x)\,dx = \sum_{k=0}^{\infty} \upsilon_k \int_{-\infty}^{t} x^s \; h_{a(k+1)}\,(x)\,dx. \tag{7}$$

The mean deviations about the mean $\left[\delta_1 = E\left(\left|X - \mu_1'\right|\right)\right]$ and about the median $\left[\delta_2 = E\left(\left|X - M\right|\right)\right]$ of $X$ are given by $\delta_1 = 2\mu_1' F\left(\mu_1'\right) - 2\varphi_1\left(\mu_1'\right)$ and $\delta_2 = \mu_1' - 2\varphi_1\,(M)$, respectively, where $\mu_1' = E\,(X)$, $M = Median\,(X) = Q\,(0.5)$ is the median, $F\left(\mu_1'\right)$ is easily evaluated from (3) and $\varphi_1\,(t)$ is the first incomplete moment given by (7) with $s = 1$.

Now, we provide two ways to determine $\delta_1$ and $\delta_2$. First, a general equation for $\varphi_1\,(t)$ can be derived from (7) as

$$\varphi_1\,(t) = \sum_{k=0}^{\infty} \upsilon_k \; J_{a(k+1)}\,(t)\,,$$

where $J_{a(k+1)}\,(t) = \int_{-\infty}^{t} x\,h_{a(k+1)}\,(x)\,dx$ is the first incomplete moment of the exp-$G$ distribution.

A second general formula for $\varphi_1\,(t)$ is given by

$$\varphi_1\,(t) = \sum_{k=0}^{\infty} \upsilon_k \; \nu_k\,(t)\,,$$

where $\nu_k\,(t) = a\,(k+1)\;\int_0^{G(t)}\; Q_G\,(u)\;\, u^{a(k+1)-1} du$ can be computed numerically.

These equations for $\varphi_1\,(t)$ can be applied to construct Bonferroni and Lorenz curves defined for a given probability $\pi$ by $B\,(\pi) = \varphi_1\left(q\right) / \left(\pi\mu_1'\right)$ and $L\,(\pi) = \varphi_1\left(q\right) / \mu_1'$, respectively, where $\mu_1' = E\,(X)$ and $q = Q\,(\pi)$ is the qf of $X$ at $\pi$. These curves are very useful in economics, reliability, demography, insurance and medicine.

### 3.6. Entropies

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

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

Using the pdf (4), we can write

$$f(x)^{\theta} = \left(\frac{\alpha ab}{d}\right)^{\theta} g(x)^{\theta}\, G(x)^{\theta(a-1)} \left[1-\alpha G(x)^{a}\right]^{\theta(b-1)}.$$

Applying the power series (5) to the last term, we obtain

$$\left[1-\alpha G(x)^{a}\right]^{\theta(b-1)} = \sum_{k=0}^{\infty} (-1)^{k} \binom{\theta(b-1)}{k} \alpha^{k} G(x)^{ak}$$

$$\begin{aligned} f(x)^{\theta} &= \left(\frac{ab}{d}\right)^{\theta} \sum_{k=0}^{\infty} (-1)^{k} \alpha^{k+\theta} \binom{\theta(b-1)}{k} g(x)^{\theta}\, G(x)^{a(k+\theta)-\theta} \\ &= \sum_{k=0}^{\infty} \eta_k\, g(x)^{\theta}\, G(x)^{a(k+\theta)-\theta}, \end{aligned}$$

where

$$\eta_k = (-1)^{k} \alpha^{k+\theta} \left(\frac{ab}{d}\right)^{\theta} \binom{\theta(b-1)}{k}.$$

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

$$I_{\theta}(X) = \frac{1}{1-\theta} \log\left[\sum_{k=0}^{\infty} \eta_k \int_{-\infty}^{\infty} g(x)^{\theta}\, G(x)^{a(k+\theta)-\theta}\, dx\right].$$

## 3.7. Order Statistics

Order statistics make their appearance in many areas of statistical theory and practice. Let $X_1, \ldots, X_n$ be a random sample from the GK-$G$ family. The pdf of $X_{i:n}$ can be written as

$$f_{i:n}(x) = \frac{f(x)}{\mathrm{B}(i, n-i+1)} \sum_{j=0}^{n-i} (-1)^{j} \binom{n-i}{j} F(x)^{j+i-1}, \tag{8}$$

where $B(\cdot,\cdot)$ is the beta function. Based on (3), we have

$$F^{j+i-1}(x) = \sum_{l=0}^{j+i-1} \frac{(-1)^{l}}{s^{j+i-1}} \binom{j+i-1}{l} \left[1-\alpha G(x;\varphi)^{a}\right]^{l\,b},$$

where $s = 1-(1-\alpha)^{b}$.

Using (4) and the above equation, we can write

$$\begin{aligned} f(x)\, F(x)^{j+i-1} &= \alpha ab \sum_{l=0}^{j+i-1} \frac{(-1)^{l}}{s^{j+i}} \binom{j+i-1}{l} g(x;\varphi)\, G(x;\varphi)^{a-1} \\ &\quad \times \left[1-\alpha G(x;\varphi)^{a}\right]^{b(l+1)-1}. \end{aligned}$$

After a power series expansion, the last equation reduces to

$$\begin{aligned} f(x)\, F(x)^{j+i-1} &= ab \sum_{l=0}^{j+i-1} \sum_{k=0}^{\infty} \frac{(-1)^{l+k} \alpha^{k+1}}{s^{j+i}} \binom{j+i-1}{l} \\ &\quad \times \binom{b(l+1)-1}{k} g(x;\varphi)\, G(x;\varphi)^{a(k+1)-1}. \end{aligned} \tag{9}$$

Then, we have

$$f(x)\, F^{j+i-1}(x) = \sum_{k=0}^{\infty} d_k \quad h_{a(k+1)}(x), \tag{10}$$

where

$$d_k = \sum_{l=0}^{j+i-1} \frac{(-1)^{l+k}\, b\alpha^{k+1}}{a\,(k+1)\, s^{j+i}} \binom{j+i-1}{l} \binom{b\,(l+1)-1}{k}.$$

Substituting (10) in (8), the pdf of $X_{i:n}$ can be expressed as

$$f_{i:n}(x) = \sum_{k=0}^{\infty} \sum_{j=0}^{n-i} \frac{(-1)^j \binom{n-i}{j}}{\mathrm{B}\,(i, n-i+1)}\, d_k\, h_{a(k+1)}(x),$$

where $h_{a(k+1)}(x)$ is the exp-$G$ density with power parameter $a\,(k+1)$.

(10) reveals that the density function of the GK-$G$ order statistics is a linear combinations of exp-$G$ densities. So, based on (10), we can derive the properties of $X_{i:n}$ from those properties of $Y_{a(k+1)}$.

For example, the $q$th moments of $X_{i:n}$ is given by

$$E\left(X_{i:n}^q\right) = \sum_{k=0}^{\infty} \sum_{j=0}^{n-i} \frac{(-1)^j \binom{n-i}{j}}{\mathrm{B}\,(i, n-i+1)}\, d_k\, E\left(Y_{(k+1)a}\right). \qquad (11)$$

## 4. CHARACTERIZATIONS

Here, we provide two characterization theorems. We will use the following two Lemmas to prove our main results.

**Assumptin A.**

Suppose the random variable $X$ has an absolutely continuous cdf $F(x)$ and pdf $f(x)$. Let $\gamma = \sup\{X|F(x) > 0\}$ and $\delta = \inf\{X|F(x) > 1\}$.

**Lemma 1.** Suppose $X$ be a random variable having the assumption A. Let

$$E\left(X|X \le x\right) = m(x)\,\tau(x),$$

where $m(x)$ is a continuous differentiable function with the condition

$$\int_{\gamma}^{x} \frac{u - m'(u)}{m(u)} du < \infty \text{ for all } x,$$

$\gamma < x < \delta$ and $\tau(x) = f(x)/F(x)$.Then

$$f(x) = c e^{\int_{\gamma}^{x} \frac{u - m'(u)}{m(u)} du},$$

where $c$ is determined such that $\frac{1}{c} = \int_{\gamma}^{\delta} f(x)\, dx.$

**Lemma 2.** Suppose $X$ be a random variable having the assumption A. Let

$$E\left(X|X \ge x\right) = n(x)\, r(x),$$

where $n(x)$ is a continuous differentiable function with the condition

$$\int_{\gamma}^{x} \frac{u - n'(u)}{n(u)} du < \infty \text{ for all } x,$$

$\gamma < x < \delta$ and $r(x) = f(x)/[1 - F(x)]$.Then

$$f(x) = c e^{-\int_{\gamma}^{x} \frac{u + n'(u)}{n(u)} du},$$

where $c$ is determined such that $\frac{1}{c} = \int_{\gamma}^{\delta} f(x)\, dx.$

**Theorem 1.** Suppose that $X$ is an absolutely continuous random variable with cdf $F(x)$ and pdf $f(x)$. We assume $\gamma = 0$, $\delta = \infty$ and $E(X) < \infty$. Then

$$E\left(X|X \leq x\right) = m(x)\,\tau(x),$$

where

$$m(x) = \frac{1}{g(x)\,G(x)^{\alpha-1}\left[1-\alpha G(x)^{a}\right]^{b-1}}\mu_1(x),$$

$$\mu_1(x) = \int_0^x ug(u)\,G(u)^{\alpha-1}\left[1-\alpha G(x)^{a}\right]^{b-1}du$$

and $\tau(x) = f(x)/F(x)$.

**Proof.** It is easy to show that if

$$f(x) = \frac{\alpha abg(u)\,G(x)^{\alpha-1}\left[1-\alpha G(x)^{a}\right]^{b-1}}{1-(1-\alpha)^{b}},$$

then

$$m(x) = \frac{1}{g(x)\,G(x)^{\alpha-1}\left[1-\alpha G(x)^{a}\right]^{b-1}}\mu_1(x).$$

We prove here the only if condition.

Suppose that

$$m(x) = \frac{1}{g(x)\,G(x)^{\alpha-1}\left[1-\alpha G(x)^{a}\right]^{b-1}}\mu_1(x),$$

$$\mu_1(x) = \int_0^x ug(u)\,G(u)^{\alpha-1}\left[1-\alpha G(u)^{a}\right]^{b-1}du.$$

$$\begin{aligned} mm'(x) \;=\; x - & \frac{m(x)}{g(x)\,G(x)^{\alpha-1}\left[1-\alpha G(x)^{a}\right]^{b-1}} \\ & \times\left(\left\{\left[g'(x)G(x)+(\alpha-1)g(x)\right]^{2}\right\}G(x)^{\alpha-2}\left[1-\alpha G(x)^{a}\right]^{b-1}\right. \\ & \left. -\;(b-1)\left[1-\alpha G(x)^{a}\right]^{b-2}\alpha aG(x)^{\alpha+a-2}g(x)^{2}\left[1-\alpha G(x)^{a}\right]^{b-1}\right). \end{aligned}$$

We have

$$\begin{aligned} \frac{x-m'(x)}{m(x)} \;=\; & \frac{1}{g(x)\,G(x)^{\alpha-1}\left[1-\alpha G(x)^{a}\right]^{b-1}} \\ & \times\left(\left\{\left[g'(x)G(x)+(\alpha-1)g(x)\right]^{2}\right\}G(x)^{\alpha-2}\left[1-\alpha G(x)^{a}\right]^{b-1}\right. \\ & \left. -\;(b-1)\left[1-\alpha G(x)^{a}\right]^{b-2}\alpha aG(x)^{\alpha+a-2}g(x)^{2}\left[1-\alpha G(x)^{a}\right]^{b-1}\right) \end{aligned}$$

Thus by Lemma 1

$$\begin{aligned} \frac{f'(x)}{f(x)} \;=\; & \frac{1}{g(x)\,G(x)^{\alpha-1}\left[1-\alpha G(x)^{a}\right]^{b-1}} \\ & \times\left(\left\{\left[g'(x)G(x)+(\alpha-1)g(x)\right]^{2}\right\}G(x)^{\alpha-2}\left[1-\alpha G(x)^{a}\right]^{b-1}\right. \\ & \left. -\;(b-1)\left[1-\alpha G(x)^{a}\right]^{b-2}\alpha aG(x)^{\alpha+a-2}g(x)^{2}\left[1-\alpha G(x)^{a}\right]^{b-1}\right). \end{aligned}$$

On integrating both sides of the above equation, we obtain

$$f(x) = cg(x)\,G(x)^{\alpha-1}\left[1-\alpha G(x)^{a}\right]^{b-1}.$$

Using the boundary condition $\int_0^\infty f(x)\,dx = 1$, we obtain $c = \dfrac{\alpha ab}{1-(1-\alpha)^{b}}$. □

**Theorem 2.** Suppose that $X$ is an absolutely continuous random variable with cdf $F(x)$ and pdf $f(x)$. We assume $\gamma = 0$, $\delta = \infty$ and $E(X) < \infty$. Then

$$E\left(X|X \geq x\right) = n(x)\,r(x),$$

where

$$n(x) = \frac{1}{g(x)\,G(x)^{\alpha-1}\left[1-\alpha G(u)^{a}\right]^{b-1}}\mu_1^*(x),$$

$$\mu_1^*(x) = \int_x^\infty ug(u)\,G(u)^{\alpha-1}\left[1-\alpha G(u)^{a}\right]^{b-1}du$$

and $r(x) = f(x)/\left[1-F(x)\right]$.

**Proof.** The if condition is easy to show. We will prove here the only if condition.

If

$$n(x) = \frac{1}{g(x)\,G(x)^{\alpha-1}\left[1-\alpha G(u)^{a}\right]^{b-1}}\mu_1^*(x),$$

□

$$\mu_1^*(x) = \int_x^\infty ug(u)\,G(u)^{\alpha-1}\left[1-\alpha G(u)^{a}\right]^{b-1}du.$$

Then

$$\begin{aligned} n'(x) = & -x - \frac{n(x)}{g(x)\,G(x)^{\alpha-1}\left[1-\alpha G(x)^{a}\right]^{b-1}} \\ & \times \left(\left\{\left[g'(x)G(x)+(\alpha-1)g(x)\right]^2\right\}G(x)^{\alpha-2}\left[1-\alpha G(x)^{a}\right]^{b-1}\right. \\ & \left. -\ (b-1)\left[1-\alpha G(x)^{a}\right]^{b-2}\alpha aG(x)^{\alpha+a-2}g(x)^2\left[1-\alpha G(x)^{a}\right]^{b-1}\right). \end{aligned}$$

Thus

$$\begin{aligned} -\frac{x+n'(x)}{n(x)} = & \frac{1}{g(x)\,G(x)^{\alpha-1}\left[1-\alpha G(x)^{a}\right]^{b-1}} \\ & \times \left(\left\{\left[g'(x)G(x)+(\alpha-1)g(x)\right]^2\right\}G(x)^{\alpha-2}\left[1-\alpha G(x)^{a}\right]^{b-1}\right. \\ & \left. -\ (b-1)\left[1-\alpha G(x)^{a}\right]^{b-2}\alpha aG(x)^{\alpha+a-2}g(x)^2\left[1-\alpha G(x)^{a}\right]^{b-1}\right). \end{aligned}$$

By Lemma 2, we have

$$\begin{aligned} \frac{f'(x)}{f(x)} = & \frac{1}{g(x)\,G(x)^{\alpha-1}\left[1-\alpha G(x)^{a}\right]^{b-1}} \\ & \times \left(\left\{\left[g'(x)G(x)+(\alpha-1)g(x)\right]^2\right\}G(x)^{\alpha-2}\left[1-\alpha G(x)^{a}\right]^{b-1}\right. \\ & \left. -\ (b-1)\left[1-\alpha G(x)^{a}\right]^{b-2}\alpha aG(x)^{\alpha+a-2}g(x)^2\left[1-\alpha G(x)^{a}\right]^{b-1}\right). \end{aligned}$$

On integrating both sides of the above equation, we obtain

$$f(x) = cg(x)\,G(x)^{\alpha-1}\left[1-\alpha G(x)^{a}\right]^{b-1}.$$

Using the boundary ondition $\int_0^\infty f(x)\,dx = 1$, we obtain $c = \dfrac{\alpha ab}{1-(1-\alpha)^{b}}$.

**Remark 1.** $m(x)$ and $n(x)$ can be given for the GKW, GKLL and GKGa distributions.

## 5. MAXIMUM LIKELIHOOD ESTIMATION

In this section, we determine the MLEs of the parameters of the new GK-G family from complete samples only. Let $x_1, \dots, x_n$ be a random sample from the GK-G family with parameters $\lambda$, $a$, $b$ and $\varphi$. Let $\theta = (a, b, \alpha, \varphi^{\intercal})^{\intercal}$ be the $(p \times 1)$ parameter vector. Then, the log-likelihood function for $\theta$, say $\ell = \ell(\theta)$, is given by

$$\begin{aligned}\ell &= n\log\alpha + n\log a + n\log b - n\log s + (a-1)\sum_{i=1}^{n}\log G(x_i;\varphi)\\ &+\sum_{i=1}^{n}\log g(x_i;\varphi) + (b-1)\sum_{i=1}^{n}\log\left[1-\alpha G(x_i;\varphi)^{a}\right],\end{aligned} \tag{12}$$

where $s = 1-(1-\alpha)^{b}$.

(12) 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 (12).

The score vector components, say $U(\theta) = \dfrac{\partial\ell}{\partial\theta} = \left(\dfrac{\partial\ell}{\partial a}, \dfrac{\partial\ell}{\partial b}, \dfrac{\partial\ell}{\partial\alpha}, \dfrac{\partial\ell}{\partial\varphi_k}\right)^{\intercal} = \left(U_a, U_b, U_\alpha, U_{\varphi_k}\right)^{\intercal}$, are available with the authors upon request.

Setting the nonlinear system of equations $U_a = U_b = U_\alpha = U_{\varphi_k} = 0$ and solving them simultaneously yields the MLE $\hat{\theta} = \left(\hat{a}, \hat{b}, \hat{\alpha}, \hat{\varphi}^{\intercal}\right)^{\intercal}$ of $\theta = (a, b, \alpha, \varphi^{\intercal})^{\intercal}$. These equations cannot be solved analytically and statistical software can be used to solve them numerically using iterative methods such as the Newton-Raphson type algorithms. For interval estimation of the model parameters, we require the observed information matrix whose elements are available with the corresponding author.

## 6. SIMULATION STUDY

In this subsection, a simulation study is conducted to examine the performance of the MLEs of the generalized Kumaraswamy normal (GKN) parameters. We generate 10,000 samples of size, $n$ = 50, 500 and 1,000 of the GKN model. The precision of the MLEs is discussed by means of the following measures: mean, mean square error (MSE), estimated average length (AL) and coverage probability (CP). The empirical study was conducted with software R. The empirical results are given in Table 2. The values in Table 1 indicate that the estimates are quite stable and, more importantly, are close to the true values for the these sample sizes. The simulation study shows that the maximum likelihood method is appropriate for estimating the GKN parameters. In fact, the means of the parameters tend to be closer to the true parameter values when n increases. This fact supports that the asymptotic normal distribution provides an adequate approximation to the finite sample distribution of the MLEs.

**Table 2** Simulation results of the GK-N distribution for several values of parameters.

| $\alpha$ | $a$ | $b$ | $\mu$ | $\sigma$ | $n$ | Mean | | | | | MSE | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | $\alpha$ | $a$ | $b$ | $\mu$ | $\sigma$ | $\alpha$ | $a$ | $b$ | $\mu$ | $\sigma$ |
| 0.5 | 0.5 | 2 | 0 | 1 | 50 | 0.3641 | 0.6933 | 2.2054 | −0.1111 | 1.0367 | 0.1296 | 0.3960 | 0.2200 | 0.3944 | 0.0857 |
| | | | | | 500 | 0.3991 | 0.5997 | 2.0905 | −0.0906 | 1.0334 | 0.0814 | 0.1150 | 0.0610 | 0.1732 | 0.0408 |
| | | | | | 1000 | 0.4669 | 0.5507 | 2.0448 | −0.0245 | 1.0224 | 0.0510 | 0.0547 | 0.0286 | 0.0811 | 0.0205 |
| 0.3 | 2 | 0.5 | 0 | 1 | 50 | 0.0517 | 2.2070 | 0.1416 | −0.0938 | 0.9765 | 1.0513 | 0.8443 | 0.4594 | 0.0919 | 0.0258 |
| | | | | | 500 | 0.2085 | 2.1592 | 0.3098 | −0.0988 | 0.9879 | 0.4034 | 0.3933 | 0.1505 | 0.0471 | 0.0118 |
| | | | | | 1000 | 0.1871 | 2.1492 | 0.3888 | −0.0642 | 0.9919 | 0.3583 | 0.3828 | 0.0755 | 0.0612 | 0.0089 |
| 0.7 | 1.5 | 2.5 | 0 | 1 | 50 | 0.4229 | 2.0211 | 2.8649 | −0.2270 | 1.1292 | 0.2230 | 0.8115 | 0.3188 | 0.3174 | 0.1674 |
| | | | | | 500 | 0.5629 | 1.8111 | 2.6869 | −0.1810 | 1.0252 | 0.0730 | 0.4305 | 0.1898 | 0.1404 | 0.0165 |
| | | | | | 1000 | 0.6727 | 1.5157 | 2.4998 | −0.0182 | 0.9933 | 0.0294 | 0.0253 | 0.0144 | 0.0108 | 0.0084 |

## 7. THE LOG-GENERALIZED KUMARASWAMY-WEIBULL (LGKW) REGRESSION MODEL

The GKW distribution with five parameters, $0 < \alpha \le 1$, $a > 0$, $b > 0$, $\lambda > 0$ and $\beta > 0$, introduced in Section 3.1. Let $X$ is a random variable following the GKW density function and $Y$ is defined by $Y = \log(X)$. The density function of $Y$ obtained by replacing $\lambda = 1/\sigma$ and $\beta = \exp(\mu)$ reduces to

$$f(y) = \frac{\frac{\alpha ab}{\sigma}\exp\left[\left(\frac{y-\mu}{\sigma}\right) - \exp\left(\frac{y-\mu}{\sigma}\right)\right]}{1-(1-\alpha)^b} \times \left\{1-\exp\left[-\exp\left(\frac{y-\mu}{\sigma}\right)\right]\right\}^{a-1}\left[1-\alpha\left\{1-\exp\left[-\exp\left(\frac{y-\mu}{\sigma}\right)\right]\right\}^a\right]^{b-1} \quad (13)$$

where $y \in \mathfrak{R}$, $\mu \in \mathfrak{R}$, $\sigma > 0$, $0 < \alpha \le 1$, $a > 0$ and $b > 0$. We refer to (13) as the LGKW distribution, say $Y \sim \text{LGKW}(\alpha, a, b, \sigma, \mu)$, where $\mu \in \mathfrak{R}$ is the location parameter, $\sigma > 0$ is the scale parameter and $\alpha$, $a$ and $b$ are shape parameters.

The corresponding survival function is

$$s(y) = \frac{\left[1-\alpha\left\{1-\exp\left[-\exp\left(\frac{y-\mu}{\sigma}\right)\right]\right\}^a\right]^b - (1-\alpha)^b}{1-(1-\alpha)^b} \quad (14)$$

and the hrf is simply $h(y) = f(y)/S(y)$. The standardized random variable $Z = (Y-\mu)/\sigma$ has density function

$$f(z) = \frac{\alpha ab\exp[z-\exp(z)]}{1-(1-\alpha)^b}\{1-\exp[-\exp(z)]\}^{a-1}\left[1-\alpha\{1-\exp[-\exp(z)]\}^a\right]^{b-1} \quad (15)$$

Parametric regression models to estimate univariate survival functions for censored data are widely used. A parametric model that provides a good fit to lifetime data tends to yield more precise estimates of the quantities of interest. Based on the LGKW density, we propose a linear location-scale regression model linking the response variable $y_i$ and the explanatory variable vector $\mathbf{v}_i^T = (v_{i1}, \dots, v_{ip})$ given by

$$y_i = \boldsymbol{v}_i^T\beta + \sigma z_i, \ \mathrm{i} = 1, \dots, \mathrm{n} \quad (16)$$

where the random error $z_i$ has density function (15), $\beta = (\beta_1, \dots, \beta_p)^T$, $\sigma > 0$, $0 < \alpha \le 1$, $a > 0$ and $b > 0$ are unknown parameters. The parameter $\mu_i = \boldsymbol{v}_i^T\beta$ is the location of $y_i$. The location parameter vector $\mu = (\mu_1, \dots, \mu_n)^T$ is represented by a linear model $\mu = V\beta$, where $V = (v_1, \dots, v_n)^T$ is a known model matrix.

Consider a sample $(y_1, v_1), \dots, (y_n, v_n)$ of $n$ independent observations, where each random response is defined by $y_i = \min\{\log(x_i), \log(c_i)\}$. We assume non-informative censoring such that the observed lifetimes and censoring times are independent. Let $F$ and $C$ be the sets of individuals for which $y_i$ is the log-lifetime or log-censoring, respectively. The log-likelihood function for the vector of parameters $\tau = (\alpha, a, b, \sigma, \beta^T)^T$ from model (16) has the form $l(\tau) = \sum_{i\in F} l_i(\tau) + \sum_{i\in C} l_i^{(c)}(\tau)$, where $l_i(\tau) = \log[f(y_i)]$, $l_i^{(c)}(\tau) = \log[S(y_i)]$, $f(y_i)$ is the density (13) and $S(y_i)$ is the survival function (14) of $Y_i$. Then, the total log-likelihood function for $\tau$ reduces to

$$\begin{aligned}\ell(\tau) = r\log\left(\frac{\alpha ab}{\sigma}\right) - r\log\left[1-(1-\alpha)^b\right] + \sum_{i\in F}(z_i - u_i) + \\ (a-1)\sum_{i\in F}\log\{1-\exp[-u_i]\} + (b-1)\sum_{i\in F}\log\left[1-\alpha\{1-\exp[-u_i]\}^a\right] \\ \sum_{i\in C}\log\left\{\frac{\left[1-\alpha\{1-\exp[-u_i]\}^a\right]^b - (1-\alpha)^b}{1-(1-\alpha)^b}\right\}\end{aligned} \quad (17)$$

where $u_i = \exp(z_i)$, $z_i = (y_i - v_i^T\beta)/\sigma$ and $r$ is the number of uncensored observations (failures) and $c$ is the number of the censored observations. The MLE $\hat{\tau}$ of the vector of unknown parameters can be evaluated by maximizing the log-likelihood (17). We use the statistical software R to determine the estimate $\hat{\tau}$.

Further, we can use the likelihood ratio (LR) statistic for comparing LGKW model with its sub-models. We consider the partition $\tau = (\tau_1^T, \tau_2^T)^T$, where $\tau_1$ is a subset of parameters of interest and $\tau_2$ is a subset of remaining parameters. The LR statistic for testing the null hypothesis $H_0 : \tau_1 = \tau_1^{(0)}$ versus the alternative hypothesis $H_1 : \tau_1 \ne \tau_1^{(0)}$ is given by $w = 2\{\ell(\hat{\tau}) - \ell(\tilde{\tau})\}$, where $\tilde{\tau}$ and $\hat{\tau}$ are the estimates under the null and alternative hypotheses, respectively. The statistic $w$ is asymptotically (as $n \to \infty$) distributed as $\chi_k^2$, where $k$ is the dimension of the subset of parameters $\tau_1$ of interest.

# 8. APPLICATIONS

## 8.1. First Application

In this section, we illustrate the fitting performance of GKGa distribution by means of real data sets. We compare the fitting performance of GKGa distribution with its sub-models. The sub-models of the GKGa distribution are given as follows: (i) Gamma distribution, (ii) exponentiated Gamma distribution, (iii) extended Gamma distribution (new), (iv) Kumaraswamy-Gamma distribution.

The used data set consists of prices ($\times 10^4$ dollars) of 428 new vehicles for the 2004 year (Kiplinger's Personal Finance, Dec 2003) (see for details Oluyede *et al.* [8]). The required computations are carried out using the R software. Summary statistics of used data set are presented in Table 3.

The measures of goodness-of-fit including the−log-likelihood function evaluated at the MLEs, Anderson-Darling ($A^*$) and Cramer-von Mises ($W^*$) are calculated to compare the fitted models. In general, the smaller the values of these statistics, the better the fit to the data.

Table 4 gives the parameter estimates and their corresponding errors, the $W^*$ and $A^*$ statistics, the minus log-likelihood values and $p$-values. Based on Table 4, it is clear that GKGa distribution provides the overall best fit and therefore could be chosen as the most adequate model among the considered models for modeling the used data set. Here, we also applied LR tests. The LR tests can be used for comparing the GKGa distribution with its sub-models. For example, the test of $H_0 : \alpha = 1$ against $H_1 : \alpha \neq 1$ is equivalent to comparing GKGa and K-Ga distributions with each other. For this test, the LR statistic can be calculated by the following relation

$$LR = 2\left[\ell\left(\hat{\alpha}, \hat{a}, \hat{b}, \hat{\lambda}, \hat{\beta}\right) - \ell\left(1, \hat{a}^*, \hat{b}^*, \hat{\lambda}^*, \hat{\beta}^*\right)\right],$$

where $\hat{a}^*$, $\hat{b}^*$, $\hat{\lambda}^*$ and $\hat{\beta}^*$ are the ML estimators of $a, b, \lambda$ and $\lambda$, respectively, obtained under $H_0$. Under the regularity conditions and if $H_0$ is assumed to be true, the LR test statistic converges in distribution to a chi square with $r$ degrees of freedom, where $r$ equals the difference between the number of parameters estimated under $H_0$ and the number of parameters estimated in general, (for $H_0 : \alpha = 1$, we have $r = 1$). Table 5 gives the LR statistics and the corresponding $p$-values for the first data set.

Based on Table 5, we reject all the null hypotheses and conclude that the GKGa fits the used data set better than the its sub-models according to the LR test.

We also plotted the fitted pdfs of the considered models for the sake of visual comparison, in Figure 4. Figure 4(a) represents that the GKGa fits the right skewed data very well. In addition, we presented the plots of the fitted density, cumulative and survival functions as well as the probability-probability (P-P) plot for the GKGa model in Figure 4(b). These plots reveal that the GKGa distribution is a suitable model for the data.

**Table 3** Descriptive statistics of turbocharger failure time data set ($\gamma_1$ and $\gamma_2$ are pearson skewness and kurtosis coefficients, respectively).

| Data set | Mean | Median | SD | $\gamma_1$ | $\gamma_2$ |
|---|---|---|---|---|---|
| Prices ($\times 10^4$ dollars) of 428 new vehicles | 3.3 | 2.7 | 1.9 | 2.8 | 16.7 |

**Table 4** Parameters estimates of proposed model and other competitive models.

| Models | $\alpha$ | $a$ | $b$ | $\lambda$ | $\beta$ | $A^*$ | $W^*$ | $-\ell$ | $p-value$ |
|---|---|---|---|---|---|---|---|---|---|
| Ga | 1 | 1 | 1 | 4.071 | 1.242 | 4.308 | 0.646 | 777.719 | 0.035 |
| | – | – | – | 0.267 | 0.086 | | | | |
| Ex-Ga | 0.005 | 1 | 681.384 | 4.247 | 0.848 | 1.668 | 0.234 | 758.5601 | 0.422 |
| | 0.011 | – | 93.441 | 0.294 | 0.115 | | | | |
| Exp-Ga | 1 | 111.785 | 1 | 0.078 | 0.602 | 1.175 | 0.156 | 754.536 | 0.550 |
| | – | 44.434 | – | 0.030 | 0.033 | | | | |
| K-Ga | 1 | 2.500 | 0.344 | 3.426 | 2.310 | 1.558 | 0.215 | 757.241 | 0.244 |
| | – | 0.011 | 0.017 | 0.005 | 0.005 | | | | |
| GKGa | 0.005 | 449.042 | 437.736 | 0.016 | 0.404 | 0.433 | 0.047 | 748.677 | 0.916 |
| | 0.024 | 94.829 | 44.201 | 0.040 | 0.063 | | | | |

**Table 5** LR tests results for first data set.

| Models | Hypotheses | LR Statistic $w$ | $p$Value |
|---|---|---|---|
| GKGa vs Ga | $H_0 = a = b = \alpha = 1$ | 58.084 | $< 0.0001$ |
| GKGa vs Ex-Ga | $H_0 = \alpha = 1$ | 19.766 | $< 0.0001$ |
| GKGa vs Exp-Ga | $H_0 = b = \alpha = 1$ | 11.718 | 0.003 |
| GKGa vs K-Ga | $H_0 = \alpha = 1$ | 17.128 | $< 0.0001$ |

## 8.2. Second Application

The dataset contains 100 observations on HIV+ subjects belonging to an Health Maintenance Organization(HMO). The HMO wants to evaluate the survival time of these subjects. In this hypothetical data set, subjects were enrolled from January 1, 1989 until December 31, 1991. Study follow-up then ended on December 31, 1995. This data set are reported in Hosmer and Lemeshow [9] and also can be found in R package *Bolstad2*. The variables involved in the study are: $y_i$ - observed survival time (in months); $cens_i$ - censoring indicator (0 = alive at study end or lost to follow-up, 1 = death due to AIDS or AIDS related factors) and $x_{i1}$ $\left(1 = yes, 0 = no\right)$ represents the history of drug use.

The aim of the study is to relate the survival time ($y$) with the history of drug use ($v$). We consider the following regression model

$$y_i = \beta_0 + \beta_1 v_i + \sigma z_i,$$

where $y_i$ has the LGKW density (13), for $i = 1, \dots, 100$. Table 6 represents the MLEs of the model parameters of the LGKW and LW regression models fitted to the current data and the log-likelihood and AIC statistics. These results indicate that the LGKW regression model has the lowest values of these statistics, and so LGKW model provides better fitting than LW model for current data. For the fitted regression models, note that $\beta_1$ is marginally significant at the 1% level and then there is a significant difference between the drug user and drug non-user for the survival time.

A comparison of the LGKW regression model with LW regression model using LR statistics is performed. LR test statistic is calculated as 11.066 and corresponding $p$-value is 0.011. These results indicate that the LGKW model provides better fit to these data than the LW regression model.

The plots in Figure 5(a) provide the Kaplan-Meier (KM) estimate and the estimated survival functions of the LGKW regression model. There is significant difference between drug users and drug non-users survival functions. The plots of the hrf in Figure 5(b) corresponding to the survival time variable under the LGKW regression model indicate that the hrf is larger for drug non-users than drug users. Based on these plots, we conclude that the LGKW regression model provides a good fit to these data.

## 9. CONCLUSION

We propose a new class of continuous distributions named the generalized Kumaraswamy family to extended the some classes of distributions such as Exp-G by Gupta *et al.* [7] and K-G by Cordeiro and de Castro [1]. We obtain some mathematical properties of proposed family including quantile function, moments, generating function, entropies, order statistics and probability weighted moments. The maximum

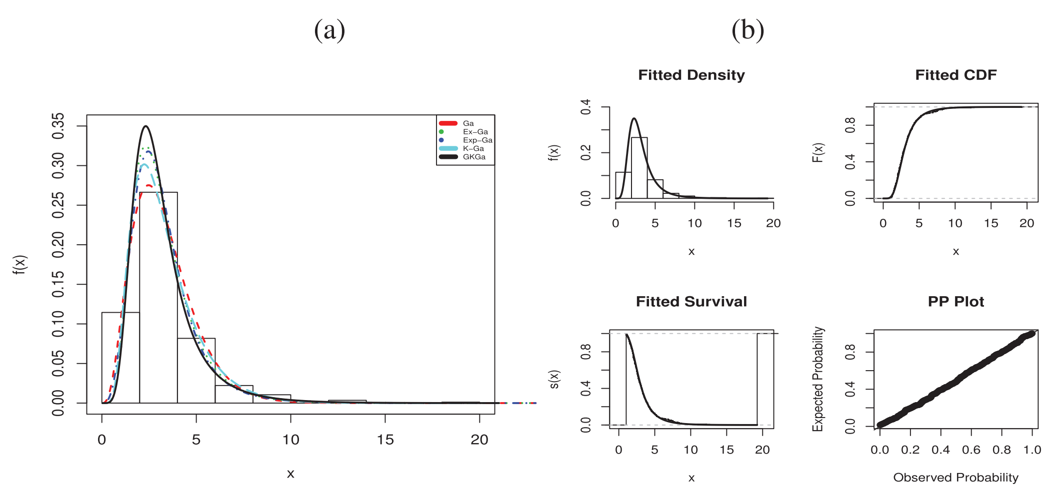


**Figure 4** (a) Fitted densities of models and (b) fitted functions of GK-gamma (GKGa) for used data set.

**Table 6** MLEs of the parameters (standard errors in parentheses and $p$-values in [ · ]) and the log-likelihood and AIC measures.

| Model | $\alpha$ | $a$ | $b$ | $\sigma$ | $\beta_0$ | $\beta_1$ | $-\ell$ | *AIC* |
|---|---|---|---|---|---|---|---|---|
| LW | 1 | 1 | 1 | 1.070 | 3.003 | −1.051 | 146.437 | 298.875 |
| | – | – | – | (0.088) | (0.166) | (0.239) | | |
| | | | | | [< 0.001] | [< 0.001] | | |
| LGKW | 4.07E-09 | 22.383 | 25.742 | 3.675 | −2.255 | −0.865 | 140.904 | 293.808 |
| | (0.0001) | (4.098) | (4.379) | (1.917) | (4.393) | (0.271) | | |
| | | | | | [0.607] | [0.001] | | |

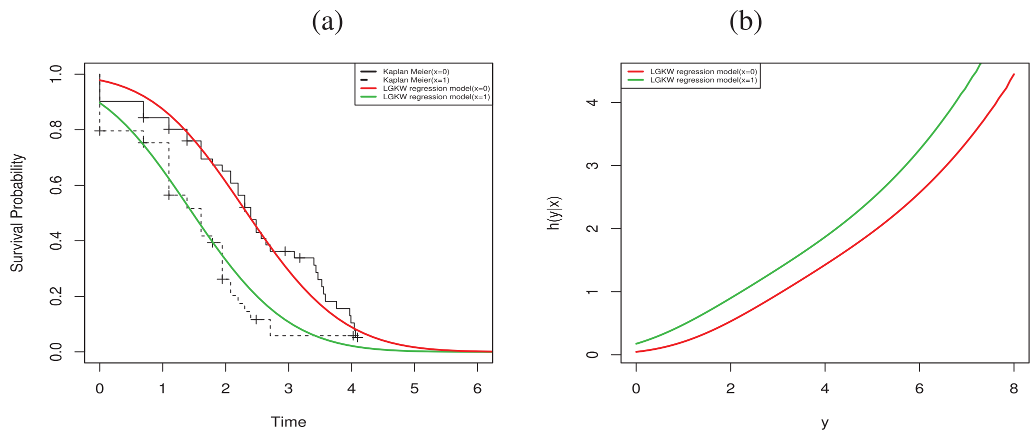


**Figure 5** | (a) Estimated survival functions and the empirical survival: :Log-generalized Kumaraswamy-Weibull (LGKW) regression model versus KM. (b) Fitted hrf using the LGKW regression model for the history of drug use.

likelihood method is used to estimate the model parameters and the performance of the maximum likelihood estimators are discussed in terms of biases, mean squared errors, coverage probability and estimated average length by means of Monte-Carlo simulation study. The usefulness of the proposed family is discussed by means of two real data applications.

## CONFLICT OF INTEREST

The authors declare no conflict of interest.

## AUTHORS' CONTRIBUTIONS

All authors contributed equally to this work.

## Funding Statement

This work has no fund.

## ACKNOWLEDGMENTS

The authors would like to thank the Editor in Chief and two reviewers for their constructive comments which improved the final version of the paper.

## REFERENCES

1. G.M. Cordeiro, M. de Castro, J. Stat. Comput. Simul. 81 (2011), 883–893.
2. A.W. Marshall, I. Olkin, Biometrika. 84 (1997), 641–652.
3. N. Eugene, C. Lee, F. Famoye, Commun. Stat. Theory Methods. 31 (2002), 497–512.
4. A.Z. Afify M., Alizadeh, H.M. Yousof, G. Aryal, M. Ahmad, Pak. J. Stat. 32 (2016), 139–160.
5. A.Z. Afify H.M. Yousof, S. Nadarajah, Stat. Interface. 10 (2017), 505–520.
6. Z.M. Nofal, A.Z. Afify, H.M. Yousof, G.M. Cordeiro, Commun. Stat. Theory Methods. 46 (2017), 4119–4136.
7. R.C. Gupta, P.L. Gupta, R.D. Gupta, Commun. Stat. Theory Methods. 27 (1998), 887–904.
8. B.O. Oluyede, F. Mutiso, S. Huang, J. Data Sci. 13 (2015), 281–309.
9. D.W. Hosmer, S. Lemeshow, Applied Survival Analysis: Regression Modeling of Time to Event Data, John Wiley and Sons Inc., New York, 1998.