Generalized Inverted Kumaraswamy Distribution Properties and Estimation

Mahmoud, A. A. and Refaey, R. M.

Statistics Department, Faculty of Commerce
AL-Azhar University (Girls Branch), Cairo, Egypt

Abstract.

The modeling and analysis of lifetimes is an important aspect of statistical work in a wide variety of scientific and technological fields. In recent years it is observed that inverted Kumaraswamy distribution has been used quite effectively to model many lifetime data. The main objective of this research is to construct a generalized inverted Kumaraswamy distribution based on M mixture representation. Also, this research is to develop a general form of inverted Kumaraswamy distribution which is flexible more than the inverted Kumaraswamy distribution and all of its related and submodules. Some properties of the generalized inverted Kumaraswamy distribution such as probability density function and cumulative distribution function are presented. The method of maximum likelihood is used for estimating the model parameters and the observed information matrix is derived. Also, the Bayesian method is used to obtain the estimators of the parameters. A simulation study is carried out to illustrate the theoretical results of the maximum likelihood estimation and Bayesian estimation. Finally, the importance and flexibility of the new model of real data set are proved empirically.

Keywords

Generalized Inverted Kumaraswamy Distribution; M Mixture; Maximum Likelihood Estimation; Bayesian estimation.

Introduction

In the past years, several ways of generating inverted distributions from classic ones were developed and discussed. In the recent times, there has been an increased interest in applying some inverted distributions to data applications in the areas of medical, economic and engineering sciences (See Calabria and Pulcini (1990), Al-Dayian (1999), Abd El-Kader *et al.* (2003), Prakash (2012)). Such distributions include the inverse Weibull, inverted Burr Type XII, the Pareto Type I and the exponentiated inverted Weibull models. However, to further improve the goodness of fit especially in exploring tail properties, researchers have also considered to derive new generators for univariate continuous families of distributions by introducing one or more additional shape parameter(s) to the baseline distribution. Kumaraswamy (1980) obtained a distribution, which is derived from beta distribution after fixing some parameters in beta distribution. But it has a closed form cumulative distribution function which is invertible and for which the moments do exist. The distribution is appropriate to natural phenomena whose outcomes are bounded from both sides, such as the individual's heights, test scores, temperatures and hydrological daily data of rain fall.

Gupta *et al*. (1998) introduced two-parameter distribution as generalization of the standard Pareto of second kind, called the *Exponentiated Pareto* (EP) distribution. Also, proved that distribution is effective in analyzing many lifetime data. The EP distribution has failure rates that take decreasing and upside-down bathtub shapes depending on the value of the shape parameter as was done for the *exponentited Weibull* (EW) distribution by Mudholkar *et al.* (1995). They observed that exponential distribution, generalized exponential distribution, Weibull distribution, beta distribution, Gamma distribution, uniform distribution, exponentiated exponential distribution, exponentiated Gamma distribution and other distributions can be obtained as special cases of the EP distribution.

Abd AL-Fattah *et al.* (2017) derived and studied in details the *inverted Kumaraswamy* (IKum) distribution using special transformation $Y = \frac{1}{x} - 1$, which has the same pdf of EP distribution. They introdPuced IKum distribution; some of its properties are presented through, some models of stress strength, measures of central tendency and dispersion and order statistics. Also, they obtained the *maximum likelihood* (ML) estimation and Bayesian estimation for the unknown parameters. Furthermore, there are some authors who were interested in IKum distribution. For more details [see Iqbal *et al.* (2017), Usman and Haq (2018) and Mohie El-Din and Abu-Moussa (2018)]. This distribution is important in a wide range of applications; for example engineering, medical research and lifetime problems.

The layout of the paper contains the following sections. In Section 2, Construction of the *generalized inverted Kumaraswamy* (GIKum) distribution based on the M mixture representation is introduced. ML estimation and asymptotic fisher information matrix are considered in Sections 3. In Section 4, Bayesian estimation of the unknown parameters is presented.

2. Construction of the Generalized Inverted Kumaraswamy Distribution Based on the M

Mixture

This section is devoted to illustrate construction of the GIKum distribution based on the M mixture representation and some statistical properties of GIKum distribution. Let a continuous non-negative random variable T follows an IKum distribution, if its *probability density function* (pdf) and *cumulative density function* (cdf) are given, respectively by

$$f(t;\beta,\alpha) = \alpha\beta(1+t)^{-(\alpha+1)}[1-(1+t)^{-\alpha}]^{\beta-1},\ \ t>0,\ \ \alpha,\beta>0, \quad (1)$$

$$F(t;\beta,\alpha) = [1-(1+t)^{-\alpha}]^{\beta}, \quad t>0, \quad \alpha,\beta>0, \quad (2)$$

The *survival function* (sf) and *hazard rate function* (hrf) of the random variable T are given, respectively, by

$$S(t;\beta,\alpha) = 1-(1-(1+t)^{-\alpha})^{\beta}, \quad (3)$$

and

$$h(t;\beta,\alpha) = \frac{\alpha\beta(1+t)^{-(\alpha+1)}(1-(1+t)^{-\alpha})^{\beta-1}}{1-(1-(1+t)^{-\alpha})^{\beta}}, 0<t, \quad \alpha,\beta>0. \quad (4)$$

The *cumulative hazard rate function* (chrf) is given by

$$H(t;\beta,\alpha) = \int_0^t h(t)dt = -\ln[1-(1-(1+t)^{-\alpha})^{\beta}]. \quad (5)$$

There are different methods to obtain generalization for distributions, in this chapter study; the generalization is constructed by using a kernel $f(t|u)$ and mixing distribution Gamma$(\theta,1)$.

If T is a continuous non-negative random variable which follows IKum distribution and U is a latent variable follows Gamma$(\theta,1)$ distribution. Then, the pdf of GIKum can be characterized as a mixture of a kernel $f(t|u)$ and mixing distribution Gamma$(\theta,1)$.

Where

$$f(t|u) = \frac{h(t)}{u} I(u > H(t))$$

$$= \frac{\alpha\beta(1+t)^{-(\alpha+1)}(1-(1+t)^{-\alpha})^{\beta-1}}{1-(1-(1+t)^{-\alpha})^{\beta}u} I(u > -\ln[1-(1-(1+t)^{-\alpha})^{\beta}]),$$

(6)

and the pdf of distribution Gamma$(2,1)$ is as follows

$$f(u) = ue^{-u},\ u > 0 \qquad (7)$$

Then, the pdf of distribution Gamma$(\theta,1)$ is as follows

$$f(u) = \frac{1}{\Gamma(\theta)} u^{\theta-1}e^{-u},\ u > 0, (\theta > 0). \qquad (8)$$

The GIKum distribution with parameters α, β and θ can be written as GIKum(α,β,θ) and non-negative random variable $T \sim$GIKum(α,β,θ) with pdf is given by

$$f_T(t) = \frac{\alpha\beta(1+t)^{-(\alpha+1)}(1-(1+t)^{-\alpha})^{\beta-1}}{[1-(1-(1+t)^{-\alpha})^{\beta}](\theta-1)}\left(1-\frac{IG(H(t),(\theta-1))}{\Gamma(\theta-1)}\right), t, \alpha > 0, \theta > 1, \qquad (9)$$

where $H(t)$ is given by (5) and $IG(x,\alpha) = \int_0^x v^{\alpha-1} e^{-v} dv$, is the incomplete gamma function.

The plot of the pdf, f(x), is provided for different values of parameters

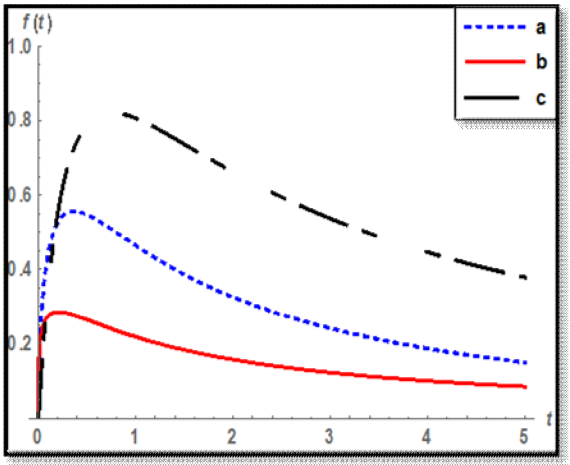

Figure 1

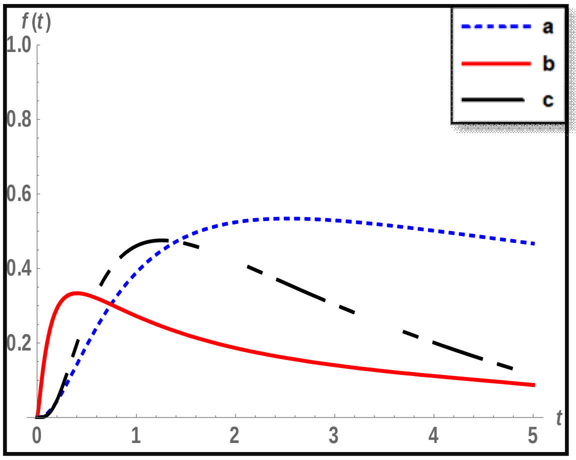

Figure 2

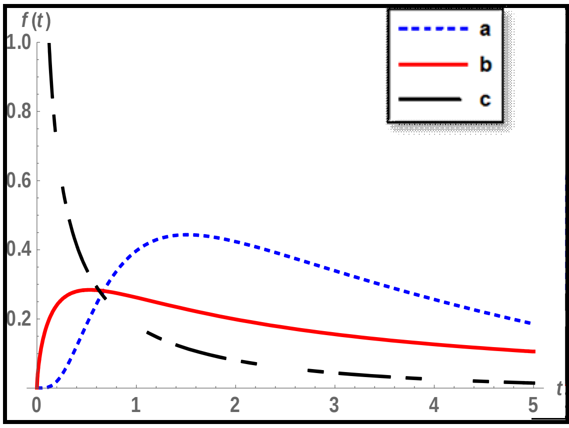

Figure 3

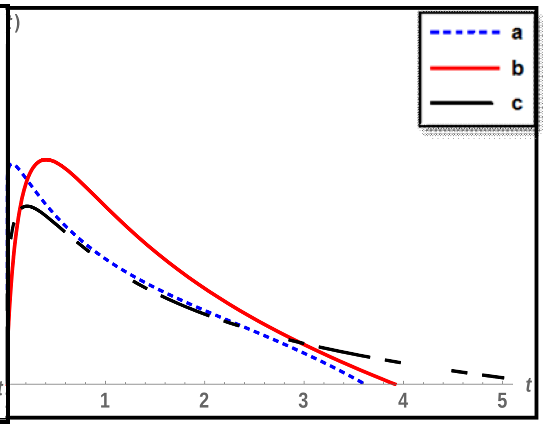

Figure 4

Figure 1, when the parameters α, β and θ has values (a where $\alpha = 1.0$, $\beta = 1.5, \theta = 1.6$), ($b$ where $\alpha = 0.16$, $\beta = 1.2, \theta = 1.2$) and ($c$ where $\alpha = 1.0$, $\beta = 2.0, \theta = 1.3$).

Figure 2, when the parameters α, β and θ has values (a where $\alpha = 0.6$, $\beta = 3.5, \theta = 1.1$), ($b$ where $\alpha = 4.5$, $\beta = 2.4, \theta = 9$) and (c where $\alpha = 2.0$, $\beta = 5.0, \theta = 1.8$).

Figure 3, when the parameters α, β and θ has values (a where $\alpha = 1.89$ $\beta = 5.69, \theta = 1.71$), ($b$ where $\alpha = 0.7$, $\beta = 1.66, \theta = 1.63$) and ($c$ where $\alpha = 0.5$ $\beta = 0.22, \theta = 2.48$).

Figure 4, when the parameters α, β and θ has values (a where $\alpha = 2.76$, $\beta = 1.08, \theta = 4.64$), ($b$ where $\alpha = 1.85$, $\beta = 1.85, \theta = 2.04$) and ($c$ where $\alpha = 1.45$ $\beta = 1.33, \theta = 2.8$).

From Figures 1,2,3 and 4, can see that the pdf can have decreasing and skewed to right (positive skewed).

The distribution function can be obtained from (9) as follows:

$$F(t) = \frac{H(t)}{(\theta - 1)} - \frac{1}{\Gamma(\theta)}\left[\int_0^t h(t)\left\{\int_0^{H(t)} u^{\theta-2}e^{-u}du\right\}dt\right], \quad (10)$$

where $H(t)$ is given by (5) and $h(t)$ is given by (4).

In order to see the behavior of the rf, the following form of the distribution function is used

$$F(t) = \int_0^{\infty} F(t|u)f(u)du, \quad (11)$$

and

$$F(t|u) = \frac{H(t)}{u}$$

The distribution function can be obtained from (9) as follows:

$$F(t) = \frac{H(t)}{(\theta-1)} - \frac{1}{\Gamma(\theta)}\left[\int_0^t h(t)\left\{\int_0^{H(t)} u^{\theta-2}e^{-u}\,du\right\}dt\right], \quad (10)$$

where $H(t)$ is given by (5) and $h(t)$ is given by (4).

In order to see the behavior of the rf, the following form of the distribution function is used

$$F(t) = \int_0^\infty F(t|u)f(u)du, \quad (11)$$

and

$$F(t|u) = \frac{H(t)}{u} = \begin{cases} \frac{-\ln[1-(1-(1+t)^{-\alpha})^{\beta}]}{u}, & 0 < t < \left[1-\{1-e^{-u}\}^{\frac{1}{\beta}}\right]^{-\frac{1}{\alpha}} - 1 \\ 1 & otherwise \end{cases}.$$

Since $F(t)$ is not in a closed form, the integral of (10) can be approximated using Markov chain integral approximation. Hence,

$$F(t) \simeq \frac{1}{m}\sum_{i=1}^{m} F(t|u^{`}),$$

where, for $i = 1, \dots, m$, u^{*} is the i^{th} variate sampled from $f(u)$, given by (7) and m is the sample size that is taken to be large. The approximate sf and hrf can be obtained.

3. Estimation of the Parameters of Generalized Inverted Kumaraswamy Distribution

In this section, the ML and Bayesian methods are used to estimate (points and intervals) the unknown parameters of the GIKum distribution based on mixture representation.

3.1 Maximum likelihood estimation

In this subsection, the ML estimation of the unknown parameters of the GIKum distribution is discussed. First, assume that $\underline{T} = ((T_1, \dots, T_n)$ is a random sample from GIKum distribution given by (1) and $\underline{U} = (U, \dots, U_n)$ is a random sample of size n from Gamma (2, 1) distribution given by (7). Then the likelihood function is given by

$$L(\alpha, \beta | \underline{t}, \underline{u}) = (\alpha\beta)^n \exp\left\{\sum_{i=1}^{n} ln\left[\frac{(1+t_i)^{-(\alpha+1)}(1-(1+t_i)^{-\alpha})^{\beta-1}}{1-(1-(1+t_i)^{-\alpha})^{\beta}}\right]\right\} exp[-\sum_i^n u_i], \quad (12)$$

Therefore, the log likelihood function of (12) given by

$$\ell = n \ln \alpha + n \ln \beta - (\alpha + 1) \sum_{i=1}^{n} ln(1 + t_i) + (\beta - 1) \sum_{i=1}^{n} ln(1 - (1 + t_i)^{-\alpha})$$

$$-\sum_{i=1}^{n}\ln[1-(1-(1+t_i)^{-\alpha})^{\beta}]-\sum_{i=1}^{n}u_i\,. \qquad (13)$$

The ML estimators can be obtained by differentiating ℓ in (13) with respect to α and β the following equations are obtained

$$\frac{\partial\ell}{\partial\beta}=\frac{n}{\beta}+\sum_{i=1}^{n}ln(1-(1+t_i)^{-\alpha})\left[1+\sum_{i=1}^{n}\frac{(1-(1+t_i)^{-\alpha})^{\beta}}{1-(1-(1+t_i)^{-\alpha})^{\beta}}\right], \qquad (14)$$

and

$$\frac{\partial\ell}{\partial\alpha}=\frac{n}{\alpha}+(\beta-1)\sum_{i=1}^{n}\frac{ln(1+t_i)}{(1+t_i)^{\alpha}-1}-\sum_{i=1}^{n}ln(1+t_i)$$

$$+\sum_{i=1}^{n}\frac{\beta[1-(1+t_i)^{-\alpha}]^{\beta-1}(1+t_i)^{-\alpha}ln(1+t_i)}{1-(1-(1+t_i)^{-\alpha})^{\beta}}\,. \qquad (15)$$

The ML estimates $\hat{\alpha}$ and $\hat{\beta}$ of the parameters α and β can be obtained by equating Equations (14) and (15) to zero and using the numerical methods such as Newton-Raphson to find the solutions of the non-linear system.

3.2 Asymptotic Fisher information matrix and approximated confidence intervals

The approximate confidence interval of the parameters $\underline{\omega_1}=(\beta,\alpha)$ can be obtained based on the asymptotic distribution of ML estimators of the parameters. Using the large sample and under appropriate regularity conditions, the ML estimators of the parameters have approximately normal distribution with mean $\underline{\omega_1}$ and asymptotic variance-covariance matrix $I^{-1}(\underline{\omega_1})$, which can be obtained as follows.

$$I^{-1}(\underline{\omega_1}) = \begin{bmatrix} -\frac{\partial^2 \ell}{\partial \beta^2} & -\frac{\partial^2 \ell}{\partial \beta \partial \alpha} \\ -\frac{\partial^2 \ell}{\partial \beta \partial \alpha} & -\frac{\partial^2 \ell}{\partial \alpha^2} \end{bmatrix}^{-1} \Bigg|_{\beta=\hat{\beta}_{ML}, \alpha=\hat{\alpha}_{ML}}$$

$$= \begin{bmatrix} var(\hat{\beta}_{ML}) & cov(\hat{\beta}_{ML}, \hat{\alpha}_{ML}) \\ cov(\hat{\alpha}_{ML}, \hat{\beta}_{ML}) & var(\hat{\alpha}_{ML}) \end{bmatrix},$$

where I is the asymptotic Fisher information matrix. The second partial derivatives will be simplified as follows:

$$\frac{\partial^2 \ell}{\partial \beta^2} = -\frac{n}{\beta^2} + \sum_{i=1}^{n} \frac{[ln(1-(1+t_i)^{-\alpha})]^2 (1-(1+t_i)^{-\alpha})^{\beta}}{[1-(1-(1+t_i)^{-\alpha})^{\beta}]^2}$$

$$\frac{\partial^2 \ell}{\partial \alpha^2} = -\frac{n}{\alpha^2} - (\beta-1) \sum_{i=1}^{n} \frac{[ln(1+t_i)]^2 (1+t_i)^{\alpha}}{[(1+t_i)^{\alpha}-1]^2} +$$

$$\beta(ln(1+t_i))^2 \left[\frac{\left((\beta-1)(1-(1+t_i)^{-\alpha})^{\beta-2}\right)}{1-(1-(1+t_i)^{-\alpha})^{\beta}} - \frac{(1+t_i)^{-\alpha}[[1-(1+t_i)^{-\alpha}]^{\beta-1}]}{1-(1-(1+t_i)^{-\alpha})^{\beta}} \right.$$

$$\left. + \frac{\beta(1-(1+t_i)^{-\alpha})^{2\beta-2}}{[1-(1-(1+t_i)^{-\alpha})^{\beta}]^2} \right],$$

$$\frac{\partial^2 \ell}{\partial \alpha \partial \beta} = \sum_{i=1}^{n} (1+t_i)^{-\alpha} ln(1+t_i) \times$$

$$\left[\frac{1}{(1-(1+t_i)^{-\alpha})}+\frac{[[1-(1+t_i)^{-\alpha}]^{\beta-1}]}{1-(1-(1+t_i)^{-\alpha})^{\beta}}+\frac{\beta[[1-(1+t_i)^{-\alpha}]^{\beta-1}]ln(1+t_i)^{-\alpha}}{1-(1-(1+t_i)^{-\alpha})^{\beta}}\right.$$

$$\left.+\frac{\beta(1-(1+t_i)^{-\alpha})^{2\beta-1}\ln[1-(1+t_i)^{-\alpha}]}{[1-(1-(1+t_i)^{-\alpha})^{\beta}]^2}\right].$$

Then the $100(1-\gamma)\%$ approximate confidence intervals for the parameters are:

$$\hat{\beta}_{ML}\pm Z_{\gamma/2}\sqrt{var(\hat{\beta}_{ML})}, \qquad \text{and} \qquad \hat{\alpha}_{ML}\pm Z_{\gamma/2}\sqrt{var(\hat{\alpha}_{ML})}.$$

where $Z_{\gamma/2}$ is the percentile of the standard normal distribution with right-tail probability $\gamma/2$.

4. Bayesian Estimation

The Bayesian method is used to obtain the estimators of the parameters of the GIKum distribution. Gibbs sampler algorithm is provided to obtain Bayesian estimates of the parameters of the GIKum distribution. It is used, for obtaining random numbers from posterior distribution of the GIKum distribution.

Suppose that $\underline{T}=(T_1,\dots,T_n)$ is a random sample of size n from a GIKum distribution and $\underline{U}=(U_1,\dots,U_n)$ is a random sample from Gamma (2, 1) distribution. The likelihood function is given by

$$L(\alpha,\beta|\underline{t},\underline{u}) = \prod_{i=1}^{n} f(t_i,u_i), \tag{16}$$

where

$$f(t_i,u_i) = f_{T|U}(t_i|u_i)f(u_i). \tag{17}$$

Using (6) and (7)

$$f_{T,U}(t_i,u_i) = \frac{\alpha\beta(1+t_i)^{-(\alpha+1)}(1-(1+t_i)^{-\alpha})^{\beta-1}}{1-(1-(1+t_i)^{-\alpha})^{\beta}} e^{-u_i} I(u_i > -\ln[1-(1-(1+t_i)^{-\alpha})^{\beta}],$$

$$i = 1, \dots, n. \tag{18}$$

The likelihood function can be rewritten as

$$L(\alpha,\beta|\underline{t},\underline{u}) = (\alpha\beta)^n \exp\left\{\sum_{i=1}^{n} ln\left[\frac{(1+t_i)^{-(\alpha+1)}(1-(1+t_i)^{-\alpha})^{\beta-1}}{1-(1-(1+t_i)^{-\alpha})^{\beta}}\right]\right\} exp\left[-\sum_{i}^{n} u_i\right]$$

$$\times I(u_i > -\ln[1-(1-(1+t_i)^{-\alpha})^{\beta}]. \tag{19}$$

Suppose that, before sampling, little or no information about the parameters α and β are available. In this case, the improper non-informative uniform distribution is used. Let the priors be of the form

$$\pi(\alpha) \propto \frac{1}{\alpha} \quad \text{and}$$

$$\pi(\beta) \propto \frac{1}{\beta},$$

The posterior distribution is

$$\pi(\alpha,\beta|\underline{t},\underline{u}) \propto \pi(\alpha)\pi(\beta)L(\alpha,\beta|\underline{t},\underline{u}),$$

where $\pi(\alpha)$ and $\pi(\beta)$ represent prior distributions. By assuming that the parameters α and β are positive. An improper joint non-informative prior distribution to α and β can be set as follows:

$$\pi(\alpha,\beta) \propto \pi(\alpha)\pi(\beta) \qquad (20)$$

Combining both (19) and (20), the posterior distribution is

$$\pi(\alpha,\beta|\underline{t},\underline{u}) \propto \pi(\alpha,\beta)L(\alpha,\beta|\underline{t},\underline{u})$$

$$\propto (\alpha\beta)^{n-1}\exp\left[\sum_{i=1}^{n} ln\left\{\frac{(1+t_i)^{-(\alpha+1)}(1-(1+t_i)^{-\alpha})^{\beta-1}}{1-\{1-(1+t_i)^{-\alpha}\}^{\beta}}\right\}\right]$$

$$\times I(u_i > -\ln[1-(1-(1+t_i)^{-\alpha})^{\beta}], \qquad (21)$$

The conditional distribution of the parameters α and β are not in closed form. The full conditional distribution of u_i is sampled,

$$\pi(u_i|\underline{t},u_{-i}) \propto e^{-u_i} \times I(u_i > -\ln[1-(1-(1+t_i)^{-\alpha})^{\beta}], \qquad (22)$$

where, $u_{-i} = (u_1, u_{i-1}, u_{i+1}, \dots, u_n)$, for $i = 1, \dots, n$.

The conditional distribution of u_i is exponential restricted to $\{-\ln[1-(1-(1+t_i)^{-\alpha})^{\beta}], \infty\}$. Also, the conditional distribution of α,

$$\pi(\alpha|\underline{t},\underline{u},\beta)$$
$$\propto \alpha^{n-1}\exp\left[\sum_{i=1}^{n}\ln\left\{\frac{(1+t_i)^{-(\alpha+1)}(1-(1+t_i)^{-\alpha})^{\beta-1}}{1-\{1-(1+t_i)^{-\alpha}\}^{\beta}}\right\}\right]$$
$$\times I\left(\alpha < min\left\{\frac{-\ln\left[1-(1-e^{-u})^{\frac{1}{\beta}}\right]}{\ln(1+t)}\right\}\right). \quad (23)$$

The conditional distribution of the parameter α is not in closed form. The following steps are suggested to deal with this case.

1. Given a random sample of $\underline{T} = (T_1, \dots, T_n)$ from IKum distribution, and a random sample of $\underline{U} = (U_1, \dots, U_n)$ from Gamma (2, 1) distribution.
2. Add a non-negative latent variable τ. the joint pdf of τ and α is given by

$$\pi(\alpha,\tau) \propto \alpha^{n-2} I(\tau < \alpha d_1) I(\alpha < G_1), \quad (24)$$

where

$$G_1 = min\left\{\frac{-\ln\left[1-(1-e^{-u})^{\frac{1}{\beta}}\right]}{\ln(1+t)}\right\}, \quad (25)$$

$$d_1 = \exp\left[\sum_{i=1}^{n} ln\left\{\frac{(1+t_i)^{-(\alpha+1)}(1-(1+t_i)^{-\alpha})^{\beta-1}}{1-\{1-(1+t_i)^{-\alpha}\}^{\beta}}\right\}\right]. \quad (26)$$

a. Given a value of the parameter α, τ is sampled from the uniform density on $(0,\alpha d_1)$ on and is denoted by uniform $(0,\alpha d_1)$.

b. Finally, using the distribution function inverse method to sample α

$$\pi(\alpha|\tau) \propto \alpha^{n-2} I(H_1 < \alpha < G_1), \qquad H_1 = \frac{\tau}{d_1}. \quad (27)$$

Then the full conditional distribution of α is as follows:

$$\pi(\alpha|\tau) = k_1 \alpha^{n-2} I(H_1 < \alpha < G_1),$$

and

$$k_1^{-1} = \int_{H_1}^{G_1} \alpha^{n-2} d\,\alpha,$$

hence

$$\pi(\alpha|\tau) = \frac{n-1}{H_1^{n-1} - G_1^{n-1}} \alpha^{n-2}, \qquad H_1 < \alpha < G_1.$$

The cdf is given by

$$F(\alpha|\tau) = \int_{H_1}^{\alpha} \frac{n-1}{H_1^{n-1} - G_1^{n-1}} \alpha^{n-2} d\,\alpha = \Psi,$$

that is,

$$\alpha = \{(G_1^{n-1} - H_1^{n-1})\Psi + H_1^{n-1}\}^{\frac{1}{n-1}} \quad , \qquad (28)$$

also,

$$\pi(\beta|\underline{t}, \underline{u}, \alpha) \propto \beta^{n-1} \exp\left[\sum_{i=1}^{n} ln\left\{\frac{(1-(1+t_i)^{-\alpha})^{\beta-1}}{1-\{1-(1+t_i)^{-\alpha}\}^{\beta}}\right\}\right]$$

$$\times I\left[\beta < max\left\{\frac{\ln(1-e^{-u})}{\ln[1-(1+t)^{-\alpha}]}\right\}\right]. \qquad (29)$$

The conditional distribution of the parameter β is not in closed form. The following steps are suggested to deal with this case.

1. Given a random sample of $\underline{T} = (T_1, \dots, T_n)$ from IKum distribution, and a random sample of $\underline{U} = (U_1, \dots, U_n)$ from Gamma (2, 1) distribution.
2. Adding a non-negative latent variable τ, the joint pdf of τ and β is given by

$$\pi(\beta, \tau) \propto \beta^{n-2} I(\beta < G_2) I(\tau < \beta d_2), \qquad (30)$$

where

$$G_2 = max\left\{\frac{\ln(1-e^{-u})}{\ln[1-(1+t)^{-\alpha}]}\right\}, \qquad d_2 = \exp\left[\sum_{i=1}^{n} ln\left\{\frac{(1-(1+t_i)^{-\alpha})^{\beta-1}}{1-\{1-(1+t_i)^{-\alpha}\}^{\beta}}\right\}\right]. \qquad (31)$$

a. Given a value of the parameter β, τ is sampled from the uniform density on $(0, \beta d_2)$ and is denoted by uniform $(0, \beta d_2)$.

b. Finally, using the distribution function inverse method to sample β

$$\pi(\beta|\tau) \propto \beta^{n-2} I(H_2 < \beta < G_2), \qquad H_2 = \frac{\tau}{d_2} . \tag{32}$$

Then the full conditional distribution of β is given by

$$\pi(\beta|\tau) = k_2 \beta^{n-2} I(H_2 < \beta < G_2),$$

and

$$k_2^{-1} = \int_{H_2}^{G_2} \beta^{n-2} d\beta,$$

hence

$$\pi(\beta|\tau) = \frac{n-1}{G_2^{n-1} - H_2^{n-1}} \beta^{n-2}, \; H_2 < \beta < G_2 .$$

The cdf is given by

$$F(\beta|\tau) = \int_{H_2}^{\alpha} \frac{n-1}{G_2^{n-1} - H_2^{n-1}} \beta^{n-2} d\beta = \Psi,$$

that is

$$\beta = \{(G_2^{n-1} - H_2^{n-1})\Psi + H_2^{n-1}\}^{\frac{1}{n-1}} . \tag{33}$$

Bayesian estimation of α and β based on the quadratic, absolute, LINEX, cannot be obtained in closed forms and numerical approximation methods are needed. See Table 4

5. Numerical Illustration

This section aims to investigate the precision of the theoretical results of estimation on basis of simulated and real data.

5.1 Simulation study

a. In this subsection, a simulation study is presented to illustrate the application of the various theoretical results developed in the previous section on basis of generated data from GIKum (α, β) distribution, for different sample sizes (n=30, 50 and 100) using number of replications N=10000. The computations are performed using Mathematica 9.

b. The *relative absolute biases* (RAB), *relative mean square errors* (RMSE), variances and *estimated risks* (ER) of ML and Bayes estimates of the shape parameters, rf and hrf are computed as follows:

1. RAB = $\left[\frac{\text{bias(estimates)}}{\text{true value}}\right]$,

2. RMSE = $\left[\frac{\text{ER}}{\text{true value}}\right]$,

3. Variances (estimates) = $\text{ER} - \text{bias}^2(\text{estimates})$,

4. ER = $\frac{\sum_{i=1}^{N}(\text{estimates} - \text{true value})^2}{N}$.

c. Table 1 displays the RAB, RMSE and variances of ML estimates and 95% *confidence intervals* (CI) where the population parameter values are $(\alpha = 2, \beta = 1.5)$. Table 2 displays the same computational results but for different population parameter values $(\alpha = 0.8, \beta = 1.5)$.

d. Table 6 shows the Bayes averages of the parameters, CI based on complete sample using the joint non-informative prior. The computations are performed using *Markov Chain Monte Carlo* (MCMC) method.

5.2 Concluding remarks

i) From Tables 1 and 2 one can observe that the RAB, variances and RMSE of the ML estimates of the shape parameters α and β decrease when the sample size n increases. The lengths of the CI becomes narrower as the sample size increases.

ii) It is clear from Table 4 that the ER of the Bayes averages of the parameters performs better and the lengths of the CI get shorter when the sample size increases.

5.3 Applications

In this subsection, the application of real data set is provided to illustrate the importance of the GIKum distribution. To check the validity of the fitted model, Kolmogorov- Smirnov goodness of fit test is performed for the data set and the p values in this case indicates that the model fits the data very well. Table 3 shows ML averages of the parameters and their ER, for the real data based on complete sample. Table 6 displays the Bayes averages of the parameters and their ER based on complete sample using the joint non-informative prior.

The application is the vinyl chloride data obtained from clean upgrading, monitoring wells in mg/L; this data set was used by Bhaumik *et al*. (2009). The data is 5.1, 1.2, 1.3, 0.6, 0.5, 2.4, 0.5, 1.1, 8.0, 0.8, 0.4, 0.6, 0.9, 0.4, 2.0, 0.5, 5.3, 3.2, 2.7, 2.9, 2.5, 2.3, 1.0, 0.2, 0.1, 0.1, 1.8, 0.9, 2.0, 4.0, 6.8, 1.2, 0.4, 0.2.

6 General Conclusion

In this paper, a new distribution called GIKum distribution is introduced. Some properties of GIKum distribution are derived. In addition, different two methods (ML and Bayesian) estimation of the unknown parameters are discussed and their behavior are compared through numerical simulations. Finally, real data set is used and the result of the analysis showed that the proposed distribution provide satisfactory performance compared to some other very well-known distributions.

Table 1

ML averages, estimated risks, and 95% credible intervals of the parameters of GIKum $(N = 10000, \alpha = 2, \beta = 1.5)$

n	Parameters	Averages	Var	ER	UL	LL	Length
30	α	1.3282	0.1544	0.6056	2.0985	0.5580	1.5404
	β	1.8224	0.4357	0.5397	3.1162	0.5286	2.5876
50	α	1.0805	0.0277	0.8731	1.4068	0.7541	0.6527
	β	1.4424	0.0398	0.0431	1.8336	1.0511	0.7824
100	α	1.0935	0.0260	0.8477	1.4101	0.7769	0.6331
	β	1.3770	0.0347	0.0498	1.7424	1.0117	0.7307

Table 2

ML averages, estimated risks, and 95% credible intervals of the parameters of GIKum $(N = 10000, \alpha = 0.8, \beta = 1.5)$

N	Parameters	Averages	Var	ER	UL	LL	Length
30	α	0.4278	0.0287	0.1672	0.7599	0.0956	0.6642
	β	1.3939	0.0556	0.0669	1.8564	0.9314	0.9249
50	α	0.4135	0.0090	0.1583	0.5997	0.2273	0.3724
	β	1.4482	0.0444	0.0470	1.8612	1.0352	0.8260
100	α	0.4193	0.0030	0.1479	0.5274	0.3112	0.2161
	β	1.44093	0.0291	0.0373	1.7440	1.0746	0.6694

Table 3

ML estimates, estimated risks, upper and lower bound of the parameters for the real data set $(\alpha = 1.1, \beta = 0.9)$

N	Parameters	Estimate	ER	Var	UL	LL	Length
34	α	0.9876	0.0126	0.0	0.9876	0.9876	0.0
	β	1.8667	0.9345	0.0	1.8667	1.8667	0.0

Table 4

Bayes averages, estimated risks, relative error and 95% credible intervals of the parameters, for GIKum under squared error loss function

$$(N = 10000, \alpha = 0.25, \beta = 0.11)$$

N	parameters	Average	RAB	Bias	ER	UL	LL	Length
30	α	0.2525	0.0099	0.0025	7.326e-06	0.2540	0.2499	0.0041
	β	0.1078	0.0099	0.0022	6.8676e-06	0.1102	0.1053	0.0049
50	α	0.2483	0.0068	0.0017	3.6919e-06	0.2501	0.2469	0.0032
	β	0.1088	0.0068	0.0004	4.9302e-07	0.1112	0.1088	0.0024
100	α	0.2505	0.0020	0.0005	7.4902e-07	0.2519	0.2493	0.0026
	β	0.1101	0.0020	0.0001	4.5291e-07	0.1109	0.1083	0.0026

Table 5

Bayes averages, estimated risks, relative error and 95% credible intervals of the parameters, for GIKum under squared error loss function $(N = 10000, \alpha = 0.68, \beta = 0.37)$

n	parameters	Average	RAB	Bias	ER	UL	LL	Length
30	α	0.6783	0.0024	0.0017	3.2881e-06	0.6797	0.6771	0.0026
	β	0.3709	0.0024	0.0009	1.1489e-06	0.3718	0.3693	0.0025
50	α	0.6803	0.0004	0.0003	7.0893e-07	0.6817	0.6789	0.0028
	β	0.3701	0.0004	0.0001	1.1599e-06	0.3718	0.3686	0.0032
100	α	0.6799	0.0002	0.0001	4.5121e-07	0.6811	0.6786	0.0025
	β	0.3707	0.0002	0.0007	7.0759e-07	0.3715	0.3694	0.0021

Table 6

Bayes, estimates , estimated risks of the parameters for the real data set

$(\alpha = 1.1, \beta = 0.9)$

n	Parameters	Estimate	ER
34	A	1.0992	3.3093e-07
	B	0.8977	8.6039e-07

REFERENCES

Abd AL-Fattah, A. M., EL-Helbawy, A. A. and AL-Dayian, G. R. (2017). Inverted Kumaraswamy distribution: properties and estimation. *Pakistan Journal of Statistics*, 33(1), 37-61.

Abd EL-Kader, R.E, AL-Dayian, G.R. and AL-Gendy, S.A. (2003). Inverted Pareto Type I distribution: properties and estimation. *Journal of Faculty of Commerce AL-Azhar University, Girls' Branch*, 21, 19-40.

AL-Dayian, G.R. (1999). Burr Type III distribution: properties and estimation. *The Egyptian Statistical Journal*, 43, 102-116.

Bhaumik, D. K., Kapur, K. and Gibbons, R. D. (2009). Testing parameters of a gamma distribution for small samples. *Technometrics.* 51(3), 326-334. doi.org/10.1198/tech.2009.07038

Calabria, R. and Pulcini, G. (1990). On the maximum likelihood and least squares sstimation in the inverse Weibull distribution. *Journal of Statistica Applicate*, 2, 3-66.

Gupta, R. C., Gupta, P. L., and Gupta, R. D. (1998). Modeling failure time data by lehman alternatives. *Communications in Statistics-Theory and Methods*, 27(4), 887-904.

Iqbal, Z., Tahir, M., M., Riaz, N., Ali, S., A. and Ahmad, M. (2017). Generalize inverted Kumaraswamy distribution properties and applications. *Open Journal of Statistics*, Vol. 7, pp. 645-662.

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

Mohie El-Din, M. and Abu-Moussa, M. (2018). On estimation and prediction for the inverted Kumaraswamy distribution based on general progressive censored samples. *Pakistan Journal of Statistics* and Operation Research**,** Vol. 14, pp. 717-736.

Mudholkar, G. S., Sriastava, D. K. and Freimer, M. (1995). The exponentiated Weibull family: a reanalysis of the bus motor failure data. *Technimetrics*, 37, 436-445.

Prakash, G. (2012). Inverted exponential distribution under a Bayesian view point. *Journal of Modern Applied Statistical Methods*, 11, 190-202. https://doi.org/10.22237/jmasm/1335845700.

Usman, R., M. and Haq, M., A. (2018). The Marshall-Olkin extended inverted Kumaraswamy distribution theory and applications. *Journal of King Saud University – Science,* https://doi.org/10.1016/j.jksus.2018.05.0.