Journal of Data Science,18(2). P. 218 – 237,2020
DOI:10.6339/JDS.202004_18(2).0001

A New Family of Generalized Distributions on the Unit Interval: The $T-$Kumaraswamy Family of Distributions

Patrick Osatohanmwen[1], F.O. Oyegue[2] ,F. Ewere[3] ,B. Ajibade[4].

[1,2,3] *Department of Statistics, University of Benin, Benin City, Edo State, Nigeria.*
[4] *Department of General Studies, Petroleum Training Institute, Effurun, Delta State, Nigeria.*

Abstract

The so-called Kumaraswamy distribution is a special probability distribution developed to model doubled bounded random processes for which the mode do not necessarily have to be within the bounds. In this article, a generalization of the Kumaraswamy distribution called the T-Kumaraswamy family is defined using the T-R {Y} family of distributions framework. The resulting T-Kumaraswamy family is obtained using the quantile functions of some standardized distributions. Some general mathematical properties of the new family are studied. Five new generalized Kumaraswamy distributions are proposed using the T-Kumaraswamy method. Real data sets are further used to test the applicability of the new family.

Keywords: $\mathrm{T}-\mathrm{R}\,\{\mathrm{Y}\}$ family; Quantile function; Hazard function; Kumaraswamy distribution.

1. Introduction

Kumaraswamy (1980) developed a double-bounded probability distribution to model random processes which are limited to interval of finite length for which the mode doesn't necessarily have to be within the Interval. A special case of the interval being $(0,1)$ has been studied extensively and called the Kumaraswamy distribution. The Kumaraswamy distribution which closely mimics the beta distribution has been thought of as a good alternative to the beta distribution due to the circumstance that it has both closed form cumulative distribution function (cdf), and probability density function (pdf), a characteristic which the beta distribution do not possess (For details on some important properties of the Kumaraswamy distribution, see Jones, 2009; Mitnik, 2013). Many probability distributions have been generated in the literature using the Kumaraswamy distribution as the generator (see. Cordeiro and Castro, 2011; Marcelino et al. 2011; Paranaiba et al. 2013; Cordeiro et al. 2014; Behairy et al. 2016). Many generalized families of distributions have appeared in the literature within the last two decades (see. Cordeiro et al. 2013; Bourguignon et al. 2014). Furthermore, the development of generalized distributions with support on the interval $(0,1)$ seems very rare in the literature, although the importance of such generalized distributions cannot be over-emphasized. Examples of such few generalized distribution on $(0,1)$ include the generalized beta distribution of the first kind (McDonald, 1984), the generalized Kumaraswamy distribution (Carrasco et al. 2010), the Kumaraswamy – Kumaraswamy distribution (El Sherpieny and Ahmad 2014), and the exponentiated generalized Kumaraswamy distribution (Elgarhy et al. 2018). For random processes that assume values on the interval $(0,1)$, there is a great need to develop flexible and highly adaptive distributions to model such processes. Areas of application of distributions defined on $(0,1)$ include but not limited to serving as conjugate prior to some of the classical discrete distribution in Bayesian inference and modeling of the random behavior of percentages and proportions.

The $T-X$ family of distributions was developed by Alzaatreh *et al.* (2013a). They utilized a random variable T defined on the interval $[a,b]$, $-\infty \leq a < b \leq \infty$ with cdf and pdf $R(t)$ and $r(t)$ respectively, and another random variable X with pdf and cdf $f(x)$ and $F(x)$ respectively. Using a transformation $W\big(F(x)\big)$ of the cdf of X, they defined a new class of distributions by the cdf of the form

$$G(x) = \int_{a}^{W(F(x))} r(t)dt, \tag{1}$$

where $W(.)$ satisfies the conditions

i. $W\big(F(x)\big) \in [a,b]$,
ii. W is differentiable and monotonically non-decreasing,
iii. $W\big(F(x)\big) \to a \; as \; x \to -\infty$ and $W\big(F(x)\big) \to b \; as \; x \to \infty$.

Examples of some probability distributions developed using the $T-X$ frame work include the Weibull-Pareto distribution (Alzaatreh *et al.*, 2013b), Gumbel-Weibull Distribution (Al-Aqtash *et al.*, 2014) and the Gumbel-Burr XII distribution (Osatohanmwen *et al.*, 2017). Aljarrah *et al.* (2014) later took $W\big(F(x)\big)$ to be the quantile function of a random variable Y and defined the $T-X\,\{Y\}$ family as

$$G(x) = \int_a^{Q_Y(F(x))} r(t)dt \; = R\left(Q_Y\big(F(x)\big)\right), \tag{2}$$

where $Q_Y(p)$ is the quantile function of the random variable Y. Observe that in (2), X is used as a random variable having cdf $F(x)$ and at the same time having cdf $G(x)$ which may be confusing. This made Alzaatreh *et al.* (2014) to re-define the $T-X\,\{Y\}$ as $T-R\,\{Y\}$and proposed several generalizations of the normal distribution using the $T-R\,\{Y\}$ framework.

In section 2 the $T-$Kumaraswamy family of distributions is defined. General mathematical properties of the proposed family are presented in section 3. Some members of the new family are specified in section 4 alongside their properties. In section 5 some applications to real data sets is carried out and the paper closes in section 6 with summary and conclusions.

2. The $T-$ Kumaraswamy Family of Distributions

Suppose T , R and Y are random variables with respective cdfs $F_T(x)$, $F_R(x)$ and $F_Y(x)$. Let the corresponding quantile functions be $Q_T(p), Q_R(p)$ and $Q_Y(p)$, where the quantile function is defined as $Q_W(p) = \inf\{w: F_W(w) \geq p\}, 0 < p < 1$. Suppose the corresponding densities of T, R and Y exist and denote them by $f_T(x), f_R(x)$ and $f_Y(x)$. Assume that $T\epsilon(a,b)$ and $Y\epsilon(c,d)$ for $-\infty \leq a < b \leq \infty$ and $-\infty \leq c < d \leq \infty$. Take R to be a Kumaraswamy random variable defined on $(0,1)$with cdf and pdf given by $F_R(x) = 1-(1-x^\alpha)^\beta$ and $f_R(x) = \alpha\beta x^{\alpha-1}(1-x^\alpha)^{\beta-1}, \alpha, \beta > 0$. Following Aljarrah (2014), the cdf of the random variable X following the $T-$ Kumaraswamy family of distributions is defined as

$$\begin{aligned} F_X(x) &= \int_a^{Q_Y(1-(1-x^\alpha)^\beta)} f_T(t)dt \\ &= P\big[T \leq Q_Y\big(1-(1-x^\alpha)^\beta\big)\big] \\ &= F_T\left(Q_Y\big(1-(1-x^\alpha)^\beta\big)\right). \end{aligned} \tag{3}$$

The corresponding pdf associated with (3) is

$$\begin{aligned} f_X(x) &= \alpha\beta x^{\alpha-1}(1-x^\alpha)^{\beta-1} \times Q_Y'\big(1-(1-x^\alpha)^\beta\big) \times f_T\left(Q_Y\big(1-(1-x^\alpha)^\beta\big)\right) \\ &= \alpha\beta x^{\alpha-1}(1-x^\alpha)^{\beta-1} \times \frac{f_T\left(Q_Y\big(1-(1-x^\alpha)^\beta\big)\right)}{f_Y\big(Q_Y(1-(1-x^\alpha)^\beta)\big)} \quad 0 < x < 1, \end{aligned} \tag{4}$$

where $Q_Y'(x) = \frac{\mathrm{d}}{\mathrm{d}x} Q_Y(x)$.

Remark 1. If X follows the $T-$ Kumaraswamy family of distributions then

(i) $X \overset{d}{\rightarrow} \left(1-\left(1-F_Y(T)\right)^{1/\beta}\right)^{1/\alpha}$,

(ii) $Q_X(p) = \left(1 - \left(1 - F_Y\left(Q_T(p)\right)\right)^{1/\beta}\right)^{1/\alpha}$,

(iii) If $T \xrightarrow{d} Y$ then $X \xrightarrow{d}$ Kumaraswamy distribution with parameters α and β,

(iv) If $T \xrightarrow{d}$ Kumaraswamy distribution with parameters α and β, then $X \xrightarrow{d}$ T.

The hazard function of the random variable X can be written as

$$h_X(x) = \frac{f_X(x)}{1 - F_X(x)}$$

$$= \frac{\alpha\beta x^{\alpha-1}}{1-x^{\alpha}} \times \frac{h_T\left(Q_Y\left(1-(1-x^{\alpha})^{\beta}\right)\right)}{h_Y\left(Q_Y\left(1-(1-x^{\alpha})^{\beta}\right)\right)}, \tag{5}$$

where $h_T(.)$ and $h_Y(.)$ are the hazard functions of the random variable T and Y respectively.

2.1 The $T -$ Kumaraswamy {exponential} distribution

If Y follows the standard exponential distribution with quantile function $Q_Y(p) = -\log(1-p)$, then $Q_Y\left(1-(1-x^{\alpha})^{\beta}\right) = \log(1-x^{\alpha})^{-\beta}$. Using (3), the cdf of the $T -$ Kumaraswamy {exponential} distribution is given by

$$F_X(x) = F_T\left(\log(1-x^{\alpha})^{-\beta}\right). \tag{6}$$

The corresponding pdf is

$$f_X(x) = \frac{\alpha\beta x^{\alpha-1}}{1-x^{\alpha}} f_T\left(\log(1-x^{\alpha})^{-\beta}\right). \tag{7}$$

2.2 The $T -$ Kumaraswamy {logistic} distribution

If Y follows the standard logistic distribution with quantile function $Q_Y(p) = \log(p/(1-p))$, then $Q_Y\left(1-(1-x^{\alpha})^{\beta}\right) = \log\left((1-x^{\alpha})^{-\beta} - 1\right)$. Using (3), the cdf of the $T -$ Kumaraswamy {logistic} distribution is given by

$$F_X(x) = F_T\left(\log\left((1-x^{\alpha})^{-\beta} - 1\right)\right). \tag{8}$$

The corresponding pdf is

$$f_X(x) = \frac{\alpha\beta x^{\alpha-1}}{(1-x^{\alpha})(1-(1-x^{\alpha})^{\beta})} f_T\left(\log\left((1-x^{\alpha})^{-\beta} - 1\right)\right). \tag{9}$$

2.3 The $T -$Kumaraswamy {extreme value} distribution

If Y follows the standard extreme value distribution with quantile function $Q_Y(p) = \log(-\log(1-p))$, then $Q_Y\left(1-(1-x^\alpha)^\beta\right) = \log\left(\log(1-x^\alpha)^{-\beta}\right)$. Using (3), the cdf of the $T-$ Kumaraswamy {extreme value} distribution is given by

$$F_X(x) = F_T\left(\log\left(\log(1-x^\alpha)^{-\beta}\right)\right). \tag{10}$$

The corresponding pdf is

$$f_X(x) = \frac{\alpha\beta x^{\alpha-1}}{(1-x^\alpha)(\log(1-x^\alpha)^{-\beta})} f_T\left(\log\left(\log(1-x^\alpha)^{-\beta}\right)\right). \tag{11}$$

2.4 The T −Kumaraswamy {log-logistic} distribution

If Y follows the standard log-logistic distribution with quantile function $Q_Y(p) = p/(1-p)$, then $Q_Y\left(1-(1-x^\alpha)^\beta\right) = (1-x^\alpha)^{-\beta} - 1$. Using (3), the cdf of the $T-$ Kumaraswamy {log-logistic} distribution is given by

$$F_X(x) = F_T\left((1-x^\alpha)^{-\beta} - 1\right). \tag{12}$$

The corresponding pdf is

$$f_X(x) = \frac{\alpha\beta x^{\alpha-1}}{(1-x^\alpha)^{\beta+1}} f_T\left((1-x^\alpha)^{-\beta} - 1\right). \tag{13}$$

3. General Mathematical Properties of the $T-$Kumaraswamy family of Distributions

Some mathematical properties of the T −Kumaraswamy family are presented in this section.

Lemma 1. For any random variable T with density $f_T(x)$, then the random variable

(i) $X = \left(1-(\mathrm{e}^{-T})^{1/\beta}\right)^{1/\alpha}$ follows the T −Kumaraswamy {exponential} distribution in (6).

(ii) $X = \left(1-(1/(1+\mathrm{e}^{T}))^{1/\beta}\right)^{1/\alpha}$ follows the $T-$ Kumaraswamy {logistic} distribution in (8).

(iii) $X = \left(1-\left(\mathrm{e}^{-\mathrm{e}^{T}}\right)^{1/\beta}\right)^{1/\alpha}$ follows the $T-$ Kumaraswamy {extreme value} distribution in (10).

(iv) $X = \left(1-(1/(1+\mathrm{T}))^{1/\beta}\right)^{1/\alpha}$ follows the $T-$ Kumaraswamy {log-logistic} distribution in (12).

Proof. The proof follows from Remark 1(i).

The results obtained in Lemma 1 enable one to establish a relationship between the random variable X following the T −Kumaraswamy distribution and the random variable T.

Consequently, random samples from the T −Kumaraswamy distribution can be simulated by first simulating random samples from the distribution of the random variable T and applying the transformation accordingly.

Lemma 2. The quantile functions for the $T-$ Kumaraswamy {exponential}, T −Kumaraswamy {logistic}, T −Kumaraswamy {extreme value} and T −Kumaraswamy {log-logistic} distributions are given respectively as

(i) $Q_X(p) = \left(1 - \left(e^{-Q_T(p)}\right)^{1/\beta}\right)^{1/\alpha}$,

(ii) $Q_X(p) = \left(1 - \left(1/\left(1 + e^{Q_T(p)}\right)\right)^{1/\beta}\right)^{1/\alpha}$,

(iii) $Q_X(p) = \left(1 - \left(e^{-e^{Q_T(p)}}\right)^{1/\beta}\right)^{1/\alpha}$,

(iv) $Q_X(p) = \left(1 - \left(1/\left(1 + Q_T(p)\right)\right)^{1/\beta}\right)^{1/\alpha}$.

Proof. The proof readily follows from Remark 1(ii).

Theorem 1. The mode(s) of the T −Kumaraswamy family of distributions is/are the solution(s) of the equation

$$\frac{1-\alpha\left(1-\beta x^{\alpha}+x^{\alpha}(1-x^{\alpha})^{-1}\right)}{\alpha\beta x^{\alpha-1}(1-x^{\alpha})^{\beta-1}} = \frac{Q_Y''\left(1-(1-x^{\alpha})^{\beta}\right)}{Q_Y'\left(1-(1-x^{\alpha})^{\beta}\right)} + \frac{f_T'\left(Q_Y\left(1-(1-x^{\alpha})^{\beta}\right)\right)}{f_T\left(Q_Y\left(1-(1-x^{\alpha})^{\beta}\right)\right)} Q_Y'\left(1-(1-x^{\alpha})^{\beta}\right) \tag{14}$$

for x.

Proof. The proof follows from setting the derivative of the pdf given in (4) to zero.

Corollary 1. The mode (s) of the T −Kumaraswamy {exponential}, T −Kumaraswamy {logistic}, $T-$ Kumaraswamy {extreme value} and $T-$ Kumaraswamy {log-logistic} distributions, respectively, are the solution of the equations

(i) $$1-\alpha\left(1-\beta x^{\alpha}+x^{\alpha}(1-x^{\alpha})^{-1}\right) = \frac{\alpha\beta x^{\alpha-1}}{1-x^{\alpha}}\left\{1+\frac{f_T'\left(\log(1-x^{\alpha})^{-\beta}\right)}{f_T\left(\log(1-x^{\alpha})^{-\beta}\right)}\right\},$$

(ii) $$\frac{\left(1-\alpha\left(1-\beta x^{\alpha}+x^{\alpha}(1-x^{\alpha})^{-1}\right)\right)\left((1-x^{\alpha})\left(1-(1-x^{\alpha})^{\beta}\right)\right)}{\alpha\beta x^{\alpha-1}} = 2\left(1-(1-x^{\alpha})^{\beta}\right)-1+\frac{f_T'\left(\log\left((1-x^{\alpha})^{-\beta}-1\right)\right)}{f_T\left(\log\left((1-x^{\alpha})^{-\beta}-1\right)\right)},$$

(iii) $$\frac{\left(1-\alpha\left(1-\beta x^{\alpha}+x^{\alpha}(1-x^{\alpha})^{-1}\right)\right)\left((1-x^{\alpha})\left(\log(1-x^{\alpha})^{-\beta}\right)\right)}{\alpha\beta x^{\alpha-1}} = \log(1-x^{\alpha})^{-\beta}-1+\frac{f_T'\left(\log\left(\log(1-x^{\alpha})^{-\beta}\right)\right)}{f_T\left(\log\left(\log(1-x^{\alpha})^{-\beta}\right)\right)},$$

(iv) $$1-\alpha\left(1-\beta x^{\alpha}+x^{\alpha}(1-x^{\alpha})^{-1}\right) = \frac{\alpha\beta x^{\alpha-1}}{1-x^{\alpha}}\left\{2+\frac{f_T'\left((1-x^{\alpha})^{-\beta}-1\right)}{(1-x^{\alpha})^{\beta} f_T\left((1-x^{\alpha})^{-\beta}-1\right)}\right\}.$$

Remark 2. The mode obtained using the result in Theorem 1 may not be unique. It is possible for there to exist more than one value satisfying (14).

Shannon (1948) defined the entropy of a random variable X as $E\{-\log(g(X))\}$, where $g(X)$ is the pdf of the random variable. The entropy of the random variable X measures variation of uncertainty (Rényi, 1961).

Theorem 2. The Shannon entropy of the T −Kumaraswamy family of distributions is given by

$$\eta_X = \eta_T + E\left(\log f_Y(T)\right) - \log\alpha - \log\beta - (\alpha-1)E(\log X) - (\beta-1)E\left(\log(1-X^\alpha)\right), \quad (15)$$

where η_T is the Shannon entropy of the distribution of the random variable T.

Proof. From Remark 1(i), it follows that $T = Q_Y\left(1-(1-X^\alpha)^\beta\right)$ and hence the pdf in (4) can be written as $f_X(X) = \frac{f_T(T)}{f_Y(T)} \times \alpha\beta X^{\alpha-1}(1-X^\alpha)^{\beta-1}$.Taking the expectation of the negative logarithm of the pdf gives the required result.

Corollary 2. The Shannon entropy of the $T-$ Kumaraswamy {exponential}, $T-$Kumaraswamy {logistic}, $T-$Kumaraswamy {extreme value} and $T-$Kumaraswamy {log-logistic} distributions are given respectively by

(i) $\eta_X = \eta_T - \mu_T - \log\alpha - \log\beta - (\alpha-1)E(\log X) - (\beta-1)E\left(\log(1-X^\alpha)\right)$,

(ii) $\eta_X = \eta_T + \mu_T - \log\alpha - \log\beta - 2E\left(\log(1+e^T)\right) - (\alpha-1)E(\log X) - (\beta-1)E\left(\log(1-X^\alpha)\right)$,

(iii) $\eta_X = \eta_T + \mu_T - \log\alpha - \log\beta - E(e^T) - (\alpha-1)E(\log X) - (\beta-1)E\left(\log(1-X^\alpha)\right)$,

(iv) $\eta_X = \eta_T - \log\alpha - \log\beta - 2E\left(\log(1+T)\right) - (\alpha-1)E(\log X) - (\beta-1)E\left(\log(1-X^\alpha)\right)$,

where μ_T is the mean of the random variable T. Observe that the results in Corollary 2 (i-iv) follow from the fact that $f_Y(T) = e^{-T}, e^T(1+e^T)^{-2}, e^T e^{-e^T}$ and $(1+T)^{-2}$ for the exponential, logistic, extreme value and log-logistic distribution respectively.

Theorem 3. The r^{th} non-central moments of the $T-$Kumaraswamy {exponential}, $T-$Kumaraswamy {logistic}, $T-$Kumaraswamy {extreme value} and $T-$Kumaraswamy {log-logistic} distributions are given respectively by

(i) $$E(X^r) = \sum_{k_1=0}^{\infty}\sum_{k_2=0}^{\infty}\binom{r/\alpha}{k_1}\frac{(-1)^{k_1+k_2}k_1^{k_2}\beta^{-k_2}}{k_2!}E(T^{k_2}), \quad (16)$$

(ii) $$E(X^r) = \sum_{k_1=0}^{\infty}\sum_{k_2=0}^{\infty}\binom{r/\alpha}{k_1}\binom{k_1/\beta}{k_2}(-1)^{k_1+k_2}E\left\{\frac{e^T}{1+e^T}\right\}^{k_2}, \quad (17)$$

(iii) $$E(X^r) = \sum_{k_1=0}^{\infty}\sum_{k_2=0}^{\infty}\sum_{k_3=0}^{\infty}\binom{r/\alpha}{k_1}\frac{(-1)^{k_1+k_2}\beta^{-k_2}k_1^{k_2}k_2^{k_3}}{k_2!\,k_3!}E(T^{k_3}), \quad (18)$$

(iv) $$E(X^r) = \sum_{k_1=0}^{\infty}\sum_{k_2=0}^{\infty}\binom{r/\alpha}{k_1}\binom{k_1/\beta}{k_2}(-1)^{k_1+k_2}E\left\{\frac{T}{1+T}\right\}^{k_2}, \quad (19)$$

where $\binom{s}{k} = \frac{s(s-1)(s-2)\ldots(s-k+1)}{k!}$ is the Pochhammer falling factorial.

Proof. We shall first prove (16). Considering Lemma 1, the r^{th} non-central moments of the $T-$ Kumaraswamy {exponential} distribution can be written as $E(X^r) = E\left(1-e^{-T/\beta}\right)^{r/\alpha}$. Using the generalized binomial expansion formula and taking the expectation, the result in (16) is obtained. The results of (17) – (19) can be obtained by applying the same technique.

Remark 3. The results in (16) and (19) hold if the support of the random variable T is on the positive real line, while (17) and (18) hold if T is on the entire real line. These fully validate the choice of $W(F(x))$ as opined by Alzaatreh *et al.* (2013a), for a given distribution T.

The dispersion and the spread in a population from the center are often measured by the deviation from the mean, and the deviation from the median. Denote the mean deviation from the mean (μ) by $D(\mu)$ and the mean deviation from the median (M) by $D(M)$.

Theorem 4. The $D(\mu)$ and $D(M)$ for the $T-$ Kumaraswamy {exponential}, $T-$Kumaraswamy {logistic}, $T-$Kumaraswamy {extreme value} and $T-$Kumaraswamy {log-logistic} distributions are given respectively by

(i) $$D(\mu) = 2\mu F_X(\mu) - 2\sum_{k_1=0}^{\infty}\sum_{k_2=0}^{\infty}\binom{1/\alpha}{k_1}\frac{(-1)^{k_1+k_2}k_1^{k_2}\beta^{-k_2}}{k_2!}Z_u(\mu, 0, k_2), \quad (20)$$

$$D(M) = \mu - 2\sum_{k_1=0}^{\infty}\sum_{k_2=0}^{\infty}\binom{1/\alpha}{k_1}\frac{(-1)^{k_1+k_2}k_1^{k_2}\beta^{-k_2}}{k_2!}Z_u(M, 0, k_2), \quad (21)$$

(ii) $$D(\mu) = 2\mu F_X(\mu) - 2\sum_{k_1=0}^{\infty}\sum_{k_2=0}^{\infty}\binom{1/\alpha}{k_1}\binom{k_1/\beta}{k_2}(-1)^{k_1+k_2}Z_{\frac{e^u}{1+e^u}}(\mu, -\infty, k_2), \quad (22)$$

$$D(M) = \mu - 2\sum_{k_1=0}^{\infty}\sum_{k_2=0}^{\infty}\binom{1/\alpha}{k_1}\binom{k_1/\beta}{k_2}(-1)^{k_1+k_2}Z_{\frac{e^u}{1+e^u}}(M, -\infty, k_2), \quad (23)$$

(iii) $$D(\mu) = 2\mu F_X(\mu) - 2\sum_{k_1=0}^{\infty}\sum_{k_2=0}^{\infty}\sum_{k_3=0}^{\infty}\binom{1/\alpha}{k_1}\frac{(-1)^{k_1+k_2}\beta^{-k_2}k_1^{k_2}k_2^{k_3}}{k_2!\,k_3!}Z_u(\mu, -\infty, k_3), \quad (24)$$

$$D(M) = \mu - 2\sum_{k_1=0}^{\infty}\sum_{k_2=0}^{\infty}\sum_{k_3=0}^{\infty}\binom{1/\alpha}{k_1}\frac{(-1)^{k_1+k_2}\beta^{-k_2}k_1^{k_2}k_2^{k_3}}{k_2!\,k_3!}Z_u(M, -\infty, k_3), \quad (25)$$

(iv) $$D(\mu) = 2\mu F_X(\mu) - 2\sum_{k_1=0}^{\infty}\sum_{k_2=0}^{\infty}\binom{1/\alpha}{k_1}\binom{k_1/\beta}{k_2}(-1)^{k_1+k_2}Z_{\frac{u}{1+u}}(\mu, 0, k_2), \quad (26)$$

$$D(M) = \mu - 2\sum_{k_1=0}^{\infty}\sum_{k_2=0}^{\infty}\binom{1/\alpha}{k_1}\binom{k_1/\beta}{k_2}(-1)^{k_1+k_2}Z_{\frac{u}{1+u}}(M, 0, k_2), \quad (27)$$

where $Z_k(c, a, n) = \int_a^{Q_Y(1-(1-x^\alpha)^\beta)} k^n f_T\,(u)du.$

Proof. By definition

$$D(\mu)=\int_0^{\mu}(\mu-x)f_X(x)dx+\int_{\mu}^{1}(x-\mu)f_X(x)dx=2\int_0^{\mu}(\mu-x)f_X(x)dx$$

$$=2\mu F_X(\mu)-2\int_0^{\mu}xf_X(x)dx. \quad (28)$$

$$D(M)=\int_0^{M}(M-x)f_X(x)dx+\int_{M}^{1}(x-M)f_X(x)dx=2\int_0^{M}(M-x)f_X(x)dx+\mu-M$$

$$=\mu-2\int_0^{M}xf_X(x)dx. \quad (29)$$

To prove (20) for the $T-$Kumaraswamy {exponential} distribution, define the integral

$$I_c=\int_0^{c}xf_X(x)dx=\alpha\beta\int_0^{c}\frac{x^{\alpha}}{1-x^{\alpha}}f_T\left(\log(1-x^{\alpha})^{-\beta}\right)dx, \quad (30)$$

and using the substitution $u=\log(1-x^{\alpha})^{-\beta}$, (30) can be written as

$$I_c=\int_0^{\log(1-c^{\alpha})^{-\beta}}\left(1-\mathrm{e}^{-u/\beta}\right)^{1/\alpha}f_T(u)du. \quad (31)$$

Using the result of the generalized binomial expansion in Theorem 3, (31) can be written as

$$I_c=\sum_{k_1=0}^{\infty}\sum_{k_2=0}^{\infty}\binom{1/\alpha}{k_1}\frac{(-1)^{k_1+k_2}k_1^{k_2}\beta^{-k_2}}{k_2!}Z_u\,(c,0,k_2), \quad (32)$$

where $Z_k(c,a,n)=\int_a^{Q_Y\left(1-(1-c^{\alpha})^{\beta}\right)}k^nf_T\,(u)du$ and $Q_Y\left(1-(1-x^{\alpha})^{\beta}\right)=\log(1-x^{\alpha})^{-\beta}$. Putting (32) into (28) and (29) and replacing c with μ and M gives (20) and (21). Applying the same techniques of proving (20) and (21), the results of (22) and (23) for (ii), (24) and (25) for (iii) and (26) and (27) for (iv) follow.

4. Some Members of the $T-$ Kumaraswamy Family of Distributions

In this section, five generalizations of the Kumaraswamy distribution are presented by making use of different T distributions for some standard Y distributions. These generalizations include the Weibull-Kumaraswamy {exponential}, log-logistic-Kumaraswamy {exponential}, exponential-Kumaraswamy {log-logistic}, normal-Kumaraswamy {logistic} and logistic-Kumaraswamy {extreme value} distributions. Some properties of the Weibull-Kumaraswamy {exponential} are examined. To conserve space, properties of the other distributions are not given. One can follow the same pattern to study the properties of the other generalized Kumaraswamy distributions.

4.1 The Weibull-Kumaraswamy {exponential} (WKUM) distribution

A random variable T is said to follow the Weibull distribution with parameters c and γ if it has the cdf $F_T(\mathrm{x}) = 1 - \mathrm{e}^{-\left(x/\gamma\right)^c}, x > 0, c, \gamma > 0$. Using (6) and (7), the cdf and pdf of the WKUM distribution are given respectively by

$$F_X(\mathrm{x}) = 1 - \exp\left\{-\left[\frac{\log(1 - x^{\alpha})^{-\beta}}{\gamma}\right]^c\right\}, \tag{33}$$

$$f_X(\mathrm{x}) = \frac{\alpha\beta c x^{\alpha-1}\left[\gamma^{-1}\log(1 - x^{\alpha})^{-\beta}\right]^{c-1}}{\gamma(1 - x^{\alpha})} \exp\left\{-\left[\frac{\log(1 - x^{\alpha})^{-\beta}}{\gamma}\right]^c\right\}, \tag{34}$$

$$\alpha, \beta, c, \gamma > 0, \qquad 0 < x < 1.$$

Remark 4.

(i) When $c = 1$, the WKUM distribution reduces to the exponential-Kumaraswamy {exponential} distribution.

(ii) When $c = \gamma = 1$, the WKUM distribution reduces to the Kumaraswamy distribution.

(iii) When $\beta = c = \gamma = 1$, the WKUM distribution reduces to the power function distribution.

(iv) When $\alpha = \beta = c = \gamma = 1$, the WKUM distribution reduces to the uniform distribution

(v) When $c = 1$ and $\gamma^{-1} = n \in N$, the pdf in (34) reduces to the distribution of the minimum order statistics, $x_{(1)}$, from a Kumaraswamy random sample of size n.

The graphs of the various shape of the WKUM distribution are provided in Figure 1.

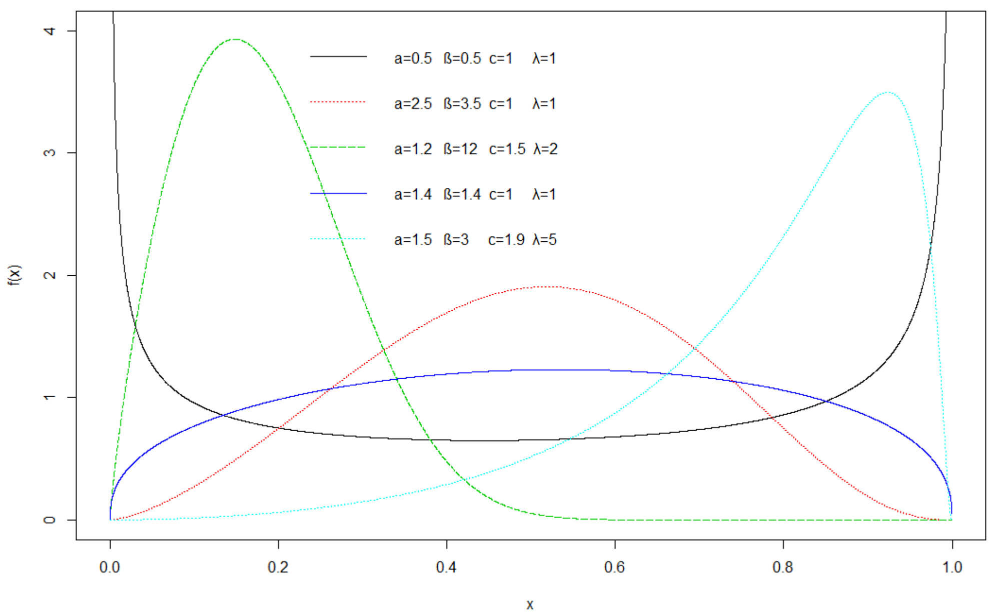

Figure 1. The pdf of the WKUM distribution.

Figure 1 clearly shows that the WKUM distribution can be right-skewed, left-skewed, symmetric, unimodal, uniantimodal.

The following are some of the properties of WKUM distribution using the general properties discussed in section 3.

(1) Quantile Function: Using Lemma 2, the quantile function of the WKUM distribution is given by

$$Q_X(\mathrm{p}) = \left(1-\left(\exp\left(-\gamma\left(-\log(1-p)\right)^{1/c}\right)\right)^{1/\beta}\right)^{1/\alpha}.$$

(2) Mode: Using Corollary 1, the mode of the WKUM distribution is the solution of the equation

$$1-\alpha(1-\beta x^{\alpha}+x^{\alpha}(1-x^{\alpha})^{-1}) = \frac{\alpha\beta x^{\alpha-1}}{1-x^{\alpha}}\left\{1+\frac{c-1}{\log(1-x^{\alpha})^{-\beta}}-c\gamma^{-c}\left(\log(1-x^{\alpha})^{-\beta}\right)^{c-1}\right\} \text{ for } x.$$

(3) Shannon entropy: Using the result in Corollary 2 and given that $\mu_T = \gamma\Gamma(1+1/c)$ and $\eta_T = \xi(1-1/c)+\log(\gamma/c)+1$ (see Song, 2001), the Shannon entropy of the WKUM distribution can be expressed as

$$\eta_X = 1+\xi\left(1-\frac{1}{c}\right)+\log\left(\frac{\gamma}{c}\right)-\gamma\Gamma\left(1+\frac{1}{c}\right)-\log\alpha-\log\beta-(\alpha-1)E(\log X)$$
$$-(\beta-1)E\left(\log(1-X^{\alpha})\right),$$

where ξ is the Euler-Mascheroni constant and $\Gamma(.)$ is the complete gamma function.

(4) Moments: Using Theorem 3, and using the fact that $E(T^{k_2}) = \gamma^{k_2}\Gamma(1+k_2/c)$, the r^{th} non-central moments of the WKUM distribution is given by

$$E(X^r) = \sum_{k_1=0}^{\infty}\sum_{k_2=0}^{\infty}\binom{r/\alpha}{k_1}\frac{(-1)^{k_1+k_2}k_1^{k_2}\beta^{-k_2}}{k_2!}\gamma^{k_2}\Gamma(1+k_2/c).$$

(5) Mean deviations: Using Theorem 4, the mean deviation from the mean and the mean deviation from the median of the WKUM distribution are given respectively by

$$D(\mu) = 2\mu F_X(\mu)-2\sum_{k_1=0}^{\infty}\sum_{k_2=0}^{\infty}\binom{1/\alpha}{k_1}\frac{(-1)^{k_1+k_2}k_1^{k_2}\beta^{-k_2}}{k_2!}\gamma^{k_2}\Gamma\left(1+\frac{k_2}{c},\left(\frac{\log(1-\mu^{\alpha})^{-\beta}}{\gamma}\right)^{c}\right),$$

$$D(M) = \mu-2\sum_{k_1=0}^{\infty}\sum_{k_2=0}^{\infty}\binom{1/\alpha}{k_1}\frac{(-1)^{k_1+k_2}k_1^{k_2}\beta^{-k_2}}{k_2!}\gamma^{k_2}\Gamma\left(1+\frac{k_2}{c},\left(\frac{\log(1-M^{\alpha})^{-\beta}}{\gamma}\right)^{c}\right),$$

where $\Gamma(a,x) = \int_0^x u^{a-1}\, e^{-u}$ is the incomplete gamma function.

4.2 The log-logistic-Kumaraswamy {exponential} (LLKUM) distribution

A random variable T is said to follow the log-logistic distribution with parameter λ if it has the cdf $F_T(x) = 1-\left(1+x^{\lambda}\right)^{-1}, x>0, \lambda>0.$ Using (6) and (7), the cdf and pdf of the LLKUM distribution are given respectively by

$$F_X(x) = 1-\left(1+(\log(1-x^{\alpha})^{-\beta})^{\lambda}\right)^{-1}, \tag{35}$$

$$f_X(x) = \frac{\alpha\beta\lambda x^{\alpha-1}}{1-x}\left(1+(\log(1-x^{\alpha})^{-\beta})^{\lambda}\right)^{-2}, \tag{36}$$

$$o<x<1, \alpha, \beta, \lambda>0.$$

The graph of the pdf of the LLKUM distribution is given in Figure 2.

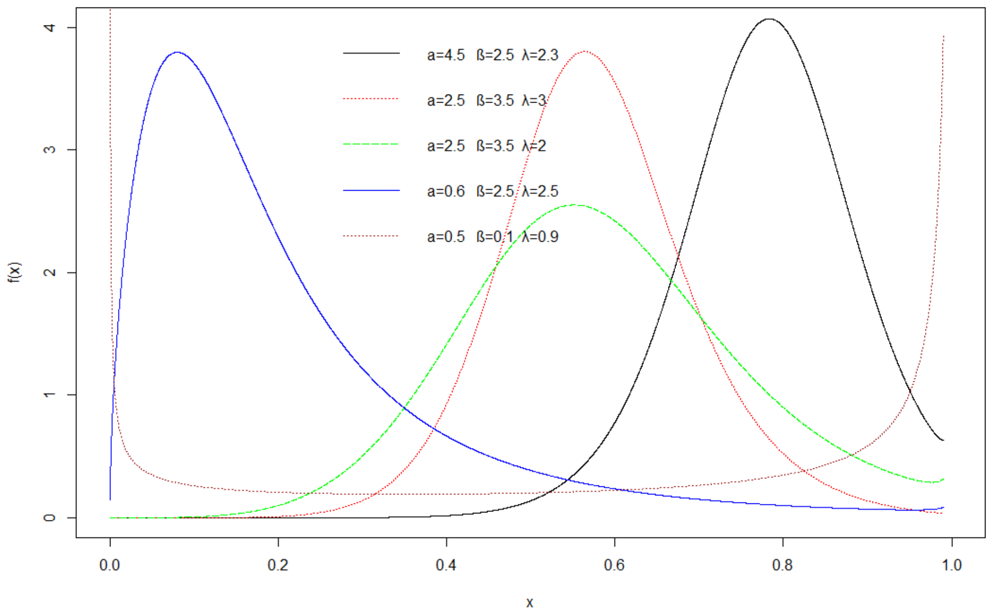

Figure 2. The pdf of the LLKUM distribution.

4.3 The exponential Kumaraswamy {log-logistic} (EKUM) Distribution

A random variable T is said to follow the exponential distribution with parameter θ if it has the cdf $F_T(x) = 1 - e^{-\theta x}, x > 0, \theta > 0$. Using (12) and (13), the cdf and pdf of the EKUM distribution are given respectively by

$$F_X(x) = 1 - \exp\{-\theta[(1 - x^{\alpha})^{-\beta} - 1]\}, \quad (37)$$

$$f_X(x) = \alpha\beta\theta x^{\alpha-1}\exp\{-\theta[(1 - x^{\alpha})^{-\beta} - 1]\}, \quad (38)$$

$$0 < x < 1, \alpha, \beta, \theta > 0.$$

The graph of the pdf of the EKUM distribution is given in Figure 3.

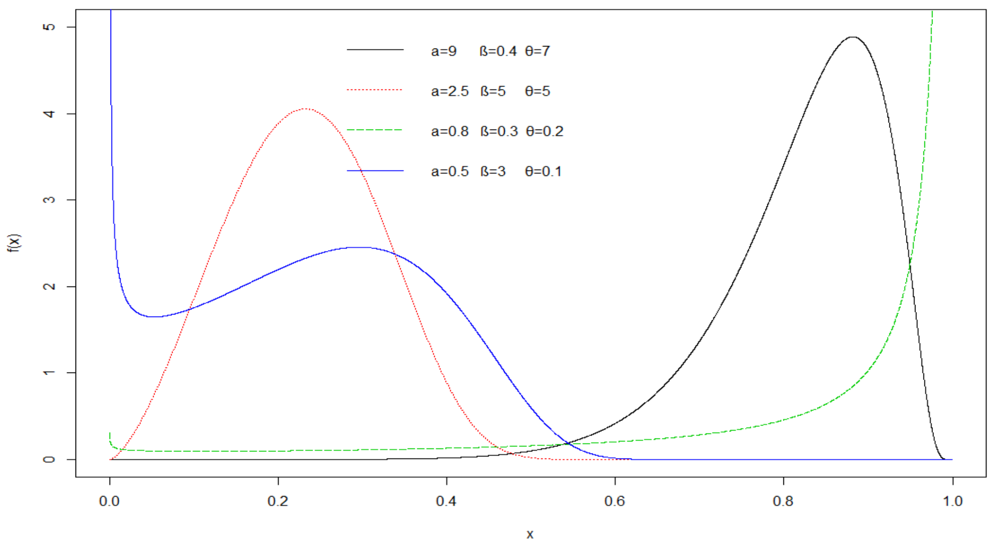

Figure 3. The pdf of the EKUM distribution.

4.3 The normal -Kumaraswamy {logistic} (NKUM) Distribution

A random variable T is said to follow the standard normal distribution if it has the cdf $F_T(x) = \Phi(x), -\infty < x < \infty$ and $\Phi(.)$ is defined in terms of the error function. Using (8) and (9) the cdf and pdf of the NKUM distribution are given respectively by

$$F_X(x) = \Phi\left(\log\left[(1-x^\alpha)^{-\beta} - 1\right]\right), \quad (39)$$

$$f_X(x) = \frac{\alpha\beta(1-x^\alpha)^{-\beta-1}x^{\alpha-1}}{(1-x^\alpha)^{-\beta} - 1}\phi\left(\log\left[(1-x^\alpha)^{-\beta} - 1\right]\right), \quad (40)$$

$$\phi(.) = \Phi'(.), 0 < x < 1, \alpha, \beta > 0.$$

The graph of the pdf of the NKUM distribution is given in Figure 4.

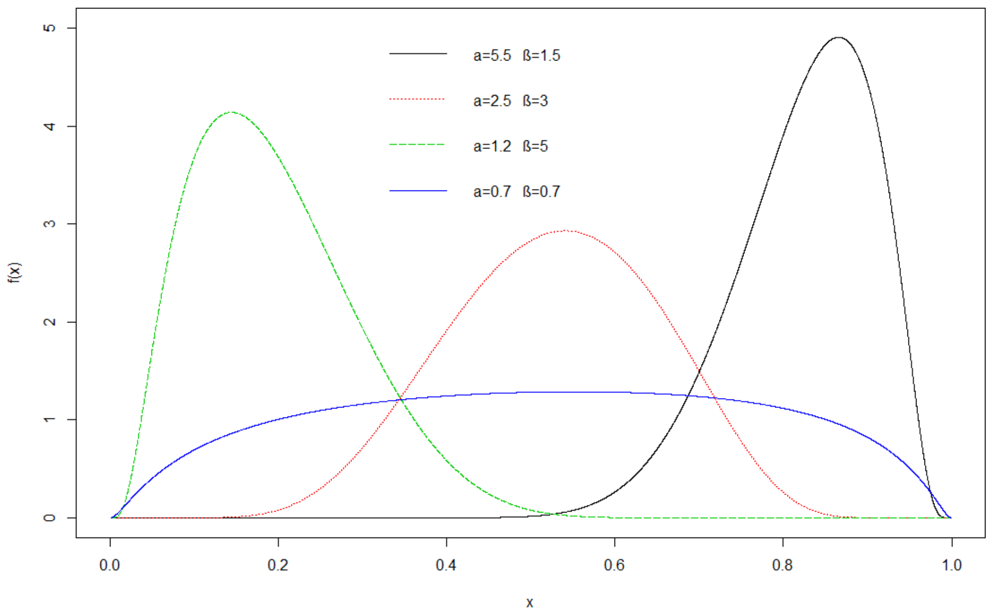

Figure 4. The pdf of the NKUM distribution.

4.4 The logistic -Kumaraswamy {extreme value} (LKUM) Distribution

A random variable T is said to follow the logistic distribution with parameter $\lambda > 0$, if it has the cdf $F_T(x) = 1 - \left(1 + e^{\lambda x}\right)^{-1}, -\infty < x < \infty$. Using (10) and (11), the cdf and pdf of the LKUM distribution are given respectively by

$$F_X(x) = \frac{\left[\log(1-x^\alpha)^{-\beta}\right]^\lambda}{1 + [\log(1-x^\alpha)^{-\beta}]^\lambda}, \quad (41)$$

$$f_X(x) = \frac{\alpha\beta\lambda x^{\alpha-1}\left[\log(1-x^\alpha)^{-\beta}\right]^{\lambda-1}}{(1-x^\alpha)\{1 + [\log(1-x^\alpha)^{-\beta}]^\lambda\}^2}, \quad (42)$$

$$0 < x < 1, \alpha, \beta, \lambda > 0.$$

The graph of the pdf of the LKUM distribution is given in Figure 5.

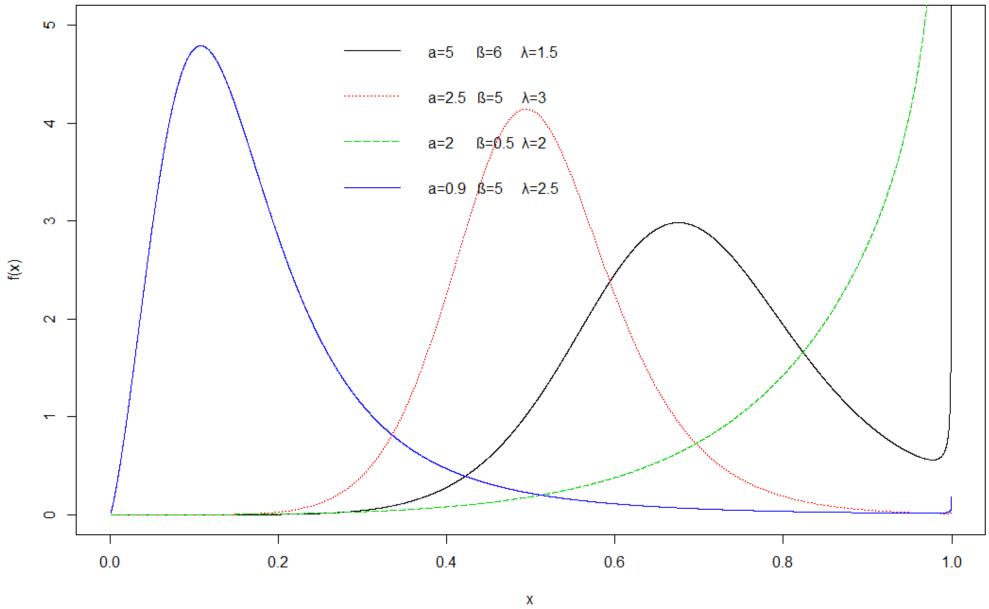

Figure 5. The pdf of the LKUM distribution.

5. Applications

In this section applications of some members of the generalized Kumaraswamy distributions will be carried out. Using the maximum likelihood estimation technique which involves the maximization of the log-likelihood function

$$L = \sum_{i=1}^{n} \log\big(f(x_i)\big),$$

for a random independent sample $x_1, x_2 \dots, x_n$ where $f(.)$ is the pdf of a distribution, we shall fit the proposed members of the $T-$Kumaraswamy family alongside the beta and Kumaraswamy distributions to two real data sets and assess the performance of all the distributions. A random variable X is said to follow the beta distribution with parameters $\alpha > 0$ and $\beta > 0$, if it has the pdf $f(x) = \frac{1}{\mathrm{B}(\alpha,\beta)} x^{a-1}(1-x)^{\beta-1}, 0 < x < 1$, where $\mathrm{B}(.,.)$ is the complete beta function.

The first data set represents the first 58 observations of the failure times of Kevlar 49/epoxy strands when the pressure is at 90% stress level, obtained from Andrews and Herzberg (1985). The data set is contained in Table 1. The WKUM, LLKUM, EKUM, NKUM, LKUM, beta and Kumaraswamy (Kumar) distributions are used to fit the data set. The results which include the parameter estimates, the log-likelihood values, and the values of the Kolmogorov-Smirnov (K-S) statistic as well as its p-value for all the distributions are contained in Table 2. Figure 6 displays the histogram and fitted densities to the data set.

The second data set represents the percentage of poor children living below and equal R$140 in 1991 in 5496 Brazilian Municipal Districts. The data were extracted from the Atlas of Brazil Human Development database available at http://www.pnud.org.br. The NKUM, beta and Kumaraswamy distributions are used to fit the data set. The results of the fit which include the parameter estimates, the log-likelihood values and the values of the

Kolmogorov-Smirnov (K-S) statistic as well as its p-value for all the distributions are contained in Table 3. Figure 7 displays the histogram and fitted densities to the data set.

Table 1: Kevlar 49/epoxy strands failure times data (pressure at 90%)

0.01, 0.01, 0.02, 0.02, 0.02, 0.03, 0.03, 0.04, 0.05, 0.06, 0.07, 0.07, 0.08, 0.09, 0.09, 0.10, 0.10, 0.11, 0.11, 0.12, 0.13, 0.18, 0.19, 0.20, 0.23, 0.24, 0.24, 0.29, 0.34, 0.35, 0.36, 0.38, 0.40, 0.42, 0.43, 0.52, 0.54, 0.56, 0.60, 0.60, 0.63, 0.65, 0.67, 0.68, 0.72, 0.72, 0.72, 0.73, 0.79, 0.79, 0.80, 0.80, 0.83, 0.85, 0.90, 0.92, 0.95, 0.99

Table 2: Maximum likelihood fit of the failure times data

Distribution	WKUM	LLKUM	EKUM	NKUM	LKUM	beta	Kumar
Parameter Estimates	$\alpha = 0.8957$ (0.0011) $\beta = 12.450$ (7.1227) $c = 0.8353$ (0.0082) $\gamma = 8.8990$ (5.2626)	$\alpha = 0.1663$ (0.2053) $\beta = 0.5707$ (0.3824) $\lambda = 2.9746$ (1.5706)	$\alpha = 0.6388$ (0.1207) $\beta = 0.0847$ (0.0847) $\theta = 10.817$ (12.9019)	$\alpha = 0.2771$ (0.0530) $\beta = 0.4950$ (0.0624)	$\alpha = 0.1662$ (0.2082) $\beta = 0.5706$ (0.3877) $\lambda = 2.9763$ (1.5934)	$\alpha = 0.6776$ (0.1107) $\beta = 1.0411$ (0.1873)	$\alpha = 0.6826$ (0.1146) $\beta = 1.0478$ (0.1795)
Log-likelihood	5.8027	4.1832	5.5132	6.0852	4.1832	5.6714	5.6824
AIC	−3.6053	−2.3664	−5.0265	−8.1705	−2.3664	−7.3427	−7.3648
BIC	4.6365	3.8150	1.1548	−4.0496	3.8149	−3.2218	−3.2439
K – S p-value	0.0980 0.5570	0.0960 0.5522	0.1071 0.5447	0.1084 0.5310	0.0972 0.5563	0.1038 0.5255	0.1034 0.5304

(Standard error of estimates in parenthesis)

From Table 2, it can be observed that all the generalized Kumaraswamy distributions as well as the beta and Kumaraswamy distributions provided adequate fit for the data by virtue of the reported p-value of the K – S statistic values with the WKUM distribution providing the best fit by possessing the highest p-value.

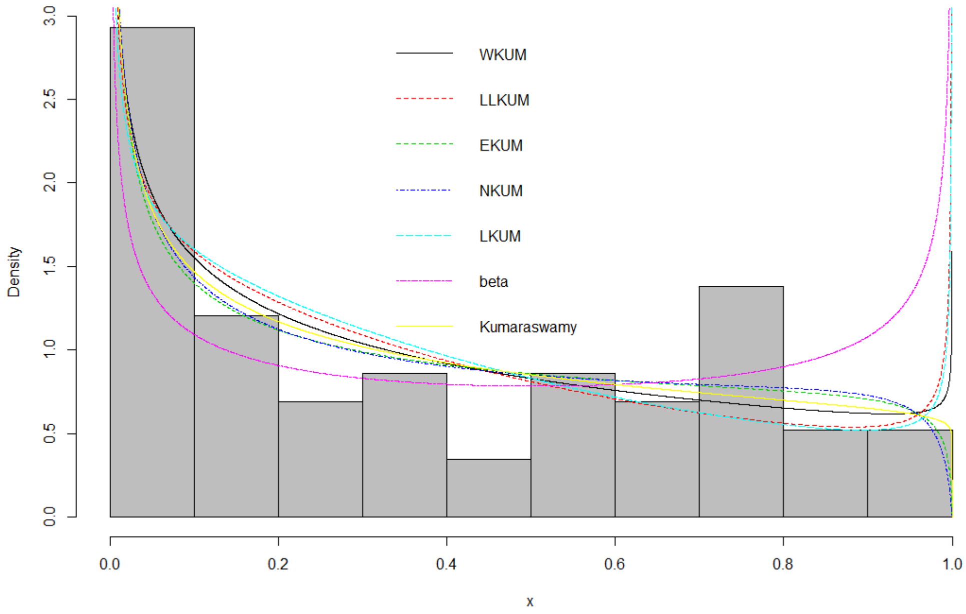

Figure 6. Histogram and fitted densities of the failure times data.

Table 3: Maximum likelihood fit of the percentage data

Distribution	NKUM	beta	Kumar
Parameter Estimates	$\alpha = 0.4349$ (0.083) $\beta = 0.5089$ (0.0066)	$\alpha = 1.1092$ (0.0197) $\beta = 1.1100$ (0.0197)	$\alpha = 1.1071$ (0.0190) $\beta = 1.1111$ (0.0200)
Log-likelihood	31.6852	20.0866	20.0460
AIC	−59.3714	−36.1732	−36.0921
BIC	−46.1479	−22.9497	−22.8685
K – S p-value	0.0119 0.3614	0.0128 0.3264	0.0127 0.3327

(Standard error of estimates in parenthesis)

Results in Table 3 clearly indicate the superiority of the NKUM distribution over the beta and Kumaraswamy distributions in fitting the data set since it reported the highest p-value value. This application clearly suggests that the 2-parameter NKUM distribution can be more flexible than the 2-parameter beta and Kumaraswamy distributions.

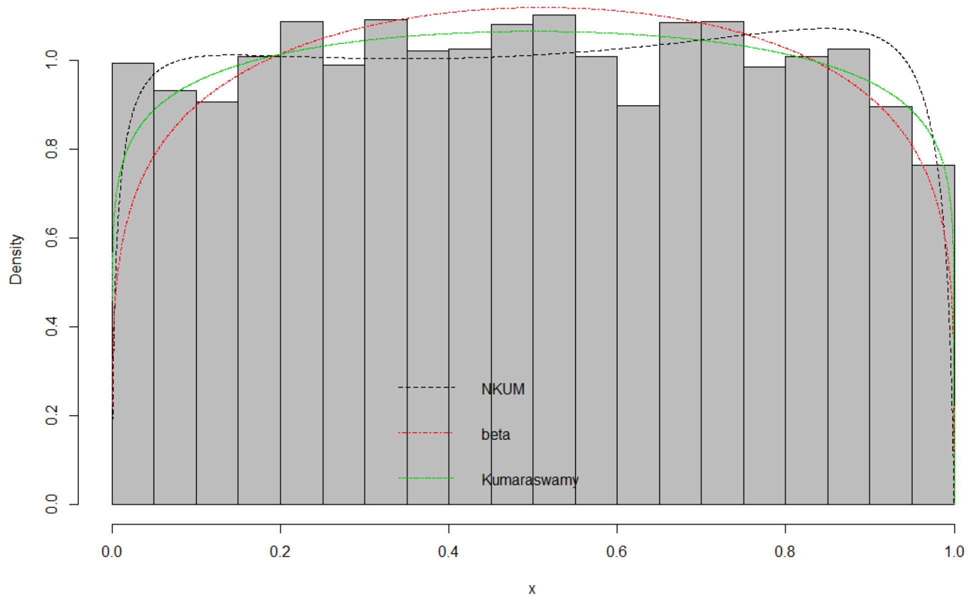

Figure 7. Histogram and Fitted densities of the percentage data.

6. Summary and Conclusion

A new family of generalized univariate distributions on the unit interval called the $T-$ Kumaraswamy distributions, which generalizes the Kumaraswamy distribution has been introduced in this paper. General expression for the quantile function, mode, moments, entropy and mean deviations of the generalized family have been given. Five members of the new family have been defined and applied to real data sets to demonstrate their applicability. Results obtained indicate that the members of the new family can be used as good alternatives to the beta and Kumaraswamy distributions. In particular, the normal-Kumaraswamy {logistics} distribution proved to be more flexible than the beta and Kumaraswamy distributions. We hope that the proposed family of distribution will attract wider applications in the analysis of proportion and percentage data.

References

[1] Al-Aqtash, R., Lee, C., and Famoye, F. (2014). Gumbel-Weibull distribution: Properties and application. *Journal of Modern Applied Statistical Method*, 13(2), 201 – 225.

[2] Aljarrah, M. A., Lee, C., and Famoye, F. (2014). A new method of generating T – X family of distributions using quantile functions, *Journal of Statistical Distributions and Applications*, 1(2), 17.

[3] Alzaatreh, A., Lee, C., and Famoye, F. (2013a). A new method for generating families of continuous distributions, *Metron*, 71(1), 63 - 79.

[4] Alzaatreh, A., Lee, C., and Famoye, F. (2013b). Weibull-Pareto distribution and its Applications. *Communications in Statistics-Theory & Method*, 42, 1673-1691.

[5] Alzaatreh, A., Lee, C., and Famoye, F. (2014). T – normal family of distributions: a new approach to generalize the normal distribution, *Journal of Statistical Distributions and Applications*, 1:16.

[6] Andrews, D.F., and Herzberg, A.M. (1985). Data: A Collection of Problems from Many Fields for the Student and Research Worker (Springer Series in Statistics). *New York: Springer.*

[7] Behairy, S.M., AL-Dayian, G.R., and EL-Helbawy, A.A. (2016). The Kumaraswamy-Burr Type III distribution: properties and estimation. *British Journal of Mathematics and Computer Science*, 14, 1 – 26.

[8] Bourguignon, M., Silva, R.B., and Cordeiro, G.M. (2014). The Weibull –G family of probability distributions, *Journal of Data Science, 12, 53 - 68.*

[9] *Carrasco, J., Ferari, S.L.P., and Cordeiro, G.M. (2010). A new generalized Kumaraswamy distribution. arXiv:1004.0911v1[stat.ME].*

[10] Cordeiro, G.M. and de Castro, M. (2011). A new family of generalized distributions, *Journal of Statistical Computation and Simulation*, 81(7), 883 - 898.

[11] Cordeiro, G.M., Ortega, G.M., and da Cunha, D.C.C. (2013). The exponentiated generalized class of distributions, *Journal of Data Science*, 11, 1 - 27.

[12] Cordeiro, G.M., Antonio, C.R., Braga, J.A.C., Demetrio, C.G., Ortega, E.M.M., and Pescim, R.R. (2014). Some new results for the Kumaraswamy modified Weibull distribution. *Journal of Statistical Theory and Applications*, 13, 86 - 104.

[13] Elgarhy, M., ul Haq, M.A., and ul Ain, Q. (2018). Exponentiated generalized Kumaraswamy distribution with applications. *Annals of Data Science*, 5(2), 273 – 292.

[14] El-Sherpieny, E.A. and Ahmad, M.A. (2014). On the Kumaraswamy-Kumaraswamy distribution. *International Journal of Basic and Applied Sciences*, 3(4), 372 – 381.

[15] Jones, M.C. (2009). Kumaraswamy's distributions: A beta-type distribution with some tractability advantages, *Statistical Methodology*, 6, 70 – 81.

[16] Kumaraswamy, P. (1980). A generalized probability density function for double bounded random processes, *Journal of Hydrology*, 46, 79 – 81.

[17] Marcelino, A.R., Ortega, E., and Cordeiro, G.M. (2011). The Kumaraswamy generalized gamma distribution with application in survival analysis. *Statistical Methodology,* 8, 411 – 433.

[18] McDonald, J.B. (1984). Some generalized function for the size distribution of income. *Econometrica*, 52, 647-663.

[19] Mitnik, P. (2013). New Properties of the Kumaraswamy distributions, *Communications in Statistics – Theory and Method*, 42(5), 741 – 755.

[20] Osatohanmwen, P., Oyegue, F.O. and Ogbonmwan, S.M. (2019). A New Member from the T-X Family of Distributions: The Gumbel-Burr XII distribution and its Properties. *Sankhya A*. 81(2), 298 - 322.

[21] Paranaiba, P.F., Ortega, E.M.M., Cordeiro, G.M., and de Pascoa, M.AR. (2013). The Kumaraswamy Burr XII distribution: theory and practice. *Journal of Statistical Computation and Simulation.* iFirst. 1 – 27.

[22] Rényi, A. (1961). On measures of entropy and information, In: *Proceedings of the fourth Berkeley Symposium on Mathematical Statistics and Probability,* 547 – 561, University of California Press, Berkeley, CA.

[23] Shannon, C.E. (1948). A mathematical theory of communication, *Bell System Technical Journal*, 27, 379 -432.

[24] Song, K.S. (2001). Rényi information, log likelihood and intrinsic distribution measure, *Journal of Statistical Planning and Inference*, 93, 51 – 69.