Journal of Biostatistics and Epidemiology

J Biostat Epidemiol. 2020;6(1):40-51

Original Article

The Topp Leone Kumaraswamy-G Family of Distributions with Applications to Cancer Disease Data

Sule Ibrahim[1*], Sani Ibrahim Doguwa[1], Audu Isah[2], Jibril M. Haruna[3]

[1]Department of Statistics, Faculty of Physical Sciences, Ahmadu Bello University, Zaria, Nigeria.
[2]Department of Statistics, School of Physical Sciences, Federal University of Technology, Minna, Nigeria.
[3]Department of Mathematics, Faculty of Physical Sciences, Ahmadu Bello University, Zaria, Nigeria.

ARTICLE INFO

Received 12.12.2019
Revised 19.01.2020
Accepted 15.02.2020
Published 10.03.2020

Key words:
Hazard function;
Topp Leone Kumaraswamy-G family;
Modi distribution;
T-X family;
Topp Leone Kumaraswamy exponential;
Topp Leone Kumaraswamy log-logistic.

ABSTRACT

Background: In the last few years, statisticians have introduced new generated families of univariate distributions. These new generators are obtained by adding one or more extra shape parameters to the underlying distribution to get more flexibility in fitting data in different areas such as medical sciences, economics, finance and environmental sciences. The addition of parameter(s) has been proven useful in exploring tail properties and also for improving the goodness-of-fit of the family of distributions under study.
Methods: A new three parameter family of distributions was introduced by using the idea of T-X methodology. Some statistical properties of the new family were derived and studied.
Results: A new Topp Leone Kumaraswamy-G family of distributions was introduced. Two special sub-models, that is, the Topp Leone Kumaraswamy exponential distribution and Topp Leone Kumaraswamy log-logistic distribution were investigated. Two real data sets were used to assess the flexibility of the sub-models.
Conclusion: The results suggest that the two sub-models performed better than their competitors.

Introduction

In the past few years, several ways of generating new distributions from classical ones were developed and discussed. This has span the interest of researchers in defining new generators or generalized classes of univariate continuous distributions by introducing additional shape parameter(s) to a baseline distribution. The addition of parameters has been proven useful in exploring skewness and tail properties, and also for improving the goodness-of-fit of the generated family. Some of the recent families of distributions appearing in literature are: The Kumaraswamy-G family of distributions in [1], Topp–Leone Family of distributions in [2], Exponentiated Weibul-H family of distributions in [3], The Topp Leone odd Lindley-G family of distributions in [4], A Topp-Leone generated family of distributions in [5], The Marshall-Olkin Odd Lindley-G Family of distributions in [6], The Exponentiated Kumaraswamy-G family of distributions in [7], The Fréchet Topp Leone G Family of distributions in [8], The Topp Leone Generalized Inverted Kumaraswamy distribution

*. Corresponding Author Email: ibrahimsule76@yahoo.com

in [9], Power Lindley-G Family of distributions in [10], Modi family of continuous probability distributions in [11], The Topp Leone Exponentiated-G family of distributions in [12].
Let r(t) be the probability density function (pdf) of a random variable $T \in [a, b]$ for $-\infty \leq a < b < \infty$ and let W[G(x)] be a function of the cumulative distribution function (cdf) of a random variable X such that W[G(x)] satisfies the following conditions:

(i) W [G(x)] $\in [a, b]$;
(ii) W [G(x)] is differentiable and monotonically non decreasing, and
(iii) W [G(x)] → a as $x \to -\infty$ and W [G(x)] $\to b$ as $x \to \infty$ (1)

[13] defined the T-X family of distributions given by

$$F(x) = \int_a^{W[G(x)]} r(t)dt \tag{2}$$

Where $W[G(x)]$ satisfies (i), (ii) and (iii). The pdf corresponding to (2) is given by

$$f(x) = \left\{\frac{dW[G(x)]}{dx}\right\} r\{W[G(X)]\} \tag{3}$$

If $H(x; \varphi)$ denotes the cumulative distribution function (cdf) of a random variable x, with corresponding pdf $h(x; \varphi)$, then the Kumaraswamy-G family defined in [1] is given as

$$G(x; \alpha, \lambda, \varphi) = 1 - [1 - H(x; \varphi)^{\alpha}]^{\lambda} \tag{4}$$

and pdf corresponding to (4) is

$$g(x; \alpha, \lambda, \varphi) = \alpha\lambda h(x; \varphi) H(x; \varphi)^{\alpha-1}[1 - H(x; \varphi)^{\alpha}]^{\lambda-1} \tag{5}$$

Let $g(x; \xi)$ and $G(x; \xi)$ denote the pdf and cdf of a baseline distribution with parameter vector ξ, then [2] defined the Topp Leone-G family of distributions with cdf given as

$$F(x; \theta, \xi) = \{1 - [1 - G(x; \xi)]^2\}^{\theta} \tag{6}$$

and pdf corresponding to (6) is

$$f(x; \theta, \xi) = 2\theta g(x; \xi)[1 - G(x; \xi)]\{1 - [1 - G(x; \xi)]^2\}^{\theta-1} \tag{7}$$

Realizing the need for more flexible life time distributions, we proposed the Topp Leone Kumaraswamy-G family of distributions in section 2. Its mathematical properties are discussed in section 3. Two sub-models of the TLK-G family are presented in section 4. Estimation of parameters and application to real data sets are discussed in section 5 while section 6 concludes the paper.

2. The Topp Leone Kumaraswamy-G family (TLK-G) of distributions.

Based on the idea of T-X family pioneered by [13], we introduced a new family of continuous distributions called the Topp Leone Kumaraswamy-G family of distributions with cdf given by

$$F(x;\alpha,\lambda,\theta,\varphi)=\int_0^{G(x;\alpha,\lambda,\varphi)} 2\theta g(t;\alpha,\lambda,\varphi)[1-G(t;\alpha,\lambda,\varphi)]\{1-[1-G(t;\alpha,\lambda,\varphi)]^2\}^{\theta-1}\,dt$$

where $G(t;\alpha,\lambda,\varphi)$ is given in (4) and $g(t;\alpha,\lambda,\varphi)$ is defined in (5).

Let $y=1-[1-G(t;\alpha,\lambda,\varphi)]^2$

$dy=2[1-G(t;\alpha,\lambda,\varphi)]g(t;\alpha,\lambda,\varphi)dt$

if $t=0$, then $y_0=0$ and if $t=x$, then

$y_x=1-[1-G(x;\alpha,\lambda,\varphi)]^2$ and

$$F(x;\alpha,\lambda,\theta,\varphi)=\int_0^{y_x}\theta y^{\theta-1}dy$$

$$F(x;\alpha,\lambda,\theta,\varphi)=\frac{\theta y^{\theta}}{\theta}\bigg|_0^{y_x}$$

$$F(x;\alpha,\lambda,\theta,\varphi)=y^{\theta}\big|_0^{y_x}$$

$$F(x;\alpha,\lambda,\theta,\varphi)=\{1-[1-G(x;\varphi)]^2\}^{\theta}$$

but from (4),

$G(x;\alpha,\lambda,\varphi)=1-[1-H(x;\varphi)^{\alpha}]^{\lambda}$

and

$[1-G(x;\alpha,\lambda,\varphi)]^2=[1-H(x;\varphi)^{\alpha}]^{2\lambda}$

Therefore,

$$F(x;\alpha,\lambda,\theta,\varphi)=\left\{1-[1-H(x;\varphi)^{\alpha}]^{2\lambda}\right\}^{\theta} \tag{8}$$

where $H(x;\varphi)$ is the cdf of the baseline distribution with parameter vector φ. The corresponding pdf to (8) is

$$f(x;\alpha,\lambda,\theta,\varphi)=2\alpha\lambda\theta h(x;\varphi)H(x;\varphi)^{\alpha-1}[1-H(x;\varphi)^{\alpha}]^{2\lambda-1}\left\{1-[1-H(x;\varphi)^{\alpha}]^{2\lambda}\right\}^{\theta-1} \tag{9}$$

which is the pdf of the proposed TLK-G family of distributions. Also, $x\geq 0$, $\alpha,\lambda,\theta,\varphi>0$.

3.Mathematical properties

3.1.Survival function

The survival function, which is the probability of an item not failing prior to some time, can be defined as

$$R(x;\alpha,\theta,\lambda,\varphi)=1-F(x;\alpha,\theta,\lambda,\varphi) \tag{10}$$

The survival function of the TLK-G family of distributions is given as

$$R(x;\alpha,\theta,\lambda,\varphi)=1-\left\{1-[1-H(x;\varphi)^{\alpha}]^{2\lambda}\right\}^{\theta} \tag{11}$$

3.2. Hazard rate function

The hazard rate function is an important measure use to characterize a life phenomenon. It is given as

$$\tau(x;\alpha,\theta,\lambda,\varphi)=\frac{f(x;\alpha,\theta,\lambda,\varphi)}{R(x;\alpha,\theta,\lambda,\varphi)} \quad (12)$$

The hazard rate function of the TLK-G family of distributions is given as

$$\tau(x;\alpha,\theta,\lambda,\varphi)=\frac{2\alpha\lambda\theta h(x;\varphi)H(x;\varphi)^{\alpha-1}[1-H(x;\varphi)^{\alpha}]^{2\lambda-1}\left\{1-[1-H(x;\varphi)^{\alpha}]^{2\lambda}\right\}^{\theta-1}}{1-\left\{1-[1-H(x;\varphi)^{\alpha}]^{2\lambda}\right\}^{\theta}} \quad (13)$$

3.3.Quantile Function

The TLK-G family is easily simulated by inverting (8) as follows: if u has a uniform U(0,1) distribution, then the solution of the nonlinear equation is given by

$$x=Q(u)=H^{-1}([1-(1-U^{\frac{1}{\theta}})^{\frac{1}{2\lambda}}]^{\frac{1}{\alpha}}) \quad (14)$$

That is, $x\sim TLK-H(x;\varphi)$.

Where H^{-1} is the quantile function of the baseline distribution, $H(x;\varphi)$

In particular, the median of the TLK-G family of distributions can be derived by substituting u = 0.5 into (14) as follows:

$$Q(0.5)=H^{-1}([1-(1-0.5^{\frac{1}{\theta}})^{\frac{1}{2\lambda}}]^{\frac{1}{\alpha}}) \quad (15)$$

4.Sub-models of the TLK-G family of distributions

In this section, we provide some sub-models of the TLK-G family. The pdf of the TLK-G family will be most tractable when f(x) and F(x) have simple analytic expressions. These special models generalize some well-known distributions reported in the literature. Here, we provide two sub-models of this family corresponding to the baseline Exponential (Ex) and Log-logistic (LL) distributions to show the flexibility of the new family.

4.1. The TLK exponential (TLKEx) distribution

The parent exponential distribution has cdf and pdf given as:

$$H(x;\beta)=1-e^{-\beta x} \quad (16)$$

$$h(x;\beta)=\beta e^{-\beta x} \quad (17)$$

The cdf and pdf of TLKEx distribution are given by

$$F(x;\alpha,\theta,\lambda,\beta)=\{1-(1-\left(1-e^{-\beta x}\right)^{\alpha})^{2\lambda}\}^{\theta} \quad (18)$$

And

$$f(x;\alpha,\theta,\lambda,\beta)=2\alpha\theta\beta\lambda e^{-\beta x}[1-e^{-\beta x}]^{\alpha-1}[1-\left(1-e^{-\beta x}\right)^{\alpha}]^{2\lambda-1}\{1-(1-\left(1-e^{-\beta x}\right)^{\alpha})^{2\lambda}\}^{\theta-1} \quad (19)$$

respectively.

$x\geq 0$, $\alpha,\lambda,\theta>0$ are the additional shape parameters and $\beta>0$ is the scale parameter.

4.1.1 Quantile function of TLKEx distribution

Lemma I: If u is uniformly distributed on (0,1), then the random variable

$$x=\frac{-1}{\beta}\log\left\{1-\left(1-\left(1-u^{\frac{1}{\theta}}\right)^{\frac{1}{2\lambda}}\right)^{\frac{1}{\alpha}}\right\}$$ has a TLKEx distribution with parameter $\alpha,\beta,\theta,\lambda$.

Lemma II: If u is uniformly distributed on (0,1) and defined the random variable

$$y=\left(1-\left(1-u^{\frac{1}{\theta}}\right)^{\frac{1}{2\lambda}}\right)^{\frac{1}{\alpha}},$$ then the random variable x defined as $x=\frac{-1}{\beta}\log(1-y)$ has a TLKEx distribution with parameters $\alpha,\beta,\theta,\lambda$.

4.1.2. The TLK log-logistic (TLKLL) distribution

The parent log-logistic distribution has cdf and pdf given as

$$H(x;\beta)=\frac{x^{\beta}}{1+x^{\beta}} \quad (20)$$

$$h(x;\beta)=\frac{\beta x^{\beta-1}}{(1+x^{\beta})^{2}} \quad (21)$$

The cdf and pdf of TLExLL distribution are given by

$$F(x;\alpha,\theta,\lambda,\beta)=\left[1-(1-(\frac{x^{\beta}}{1+x^{\beta}})^{\alpha})^{2\lambda}\right]^{\theta} \quad (22)$$

and

$$f(x;\alpha,\theta,\lambda,\beta)=2\alpha\theta\frac{\beta x^{\beta-1}}{(1+x^{\beta})^{2}}\left[\frac{x^{\beta}}{1+x^{\beta}}\right]^{\alpha-1}[1-(\frac{x^{\beta}}{1+x^{\beta}})^{\alpha}]^{2\lambda-1}\{1-(1-(\frac{x^{\beta}}{1+x^{\beta}})^{\alpha})^{2\lambda}\}^{\theta-1} \quad (23)$$

respectively.

The proposed models are more flexible than the existing models because the extra parameters allow us to obtain all possible shapes of the density as well as the hazard functions. The hazard functions of this models may be monotone, non-monotone, increasing, decreasing, unimodal and bathtub shaped. Hence, TLK-G family distributions is suitable to model and analyze a wide variety of real-life data applications.

4.2.1 Quantile function of TLKLL distribution

Lemma III: If u is uniformly distributed on (0,1), then the random variable x defined as

$$x = \left\{ \left[1 - \left[1 - u^{\frac{1}{\theta}} \right]^{\frac{1}{2\lambda}} \right]^{-\frac{1}{\alpha}} - 1 \right\}^{-\frac{1}{\beta}}$$

has a TLKLL distribution with parameters $\alpha, \beta, \theta, \lambda$.

Lemma IV: If u is uniformly distributed on (0,1), and defined the random variable

$$y = \left[1 - \left[1 - u^{\frac{1}{\theta}} \right]^{\frac{1}{2\lambda}} \right]^{\frac{1}{\alpha}},$$

then the random variable $x = \{y^{-1} - 1\}^{-\frac{1}{\beta}}$ has a TLKLL distribution with parameters $\alpha, \beta, \theta, \lambda$.

5. Estimation and Data Analysis

5.1 Maximum Likelihood Estimation

Let $X_1, X_2, \dots, X_n$ be an iid observed random sample of size n from the TLK-G family. Then, the log-likelihood function based on observed sample for the vector of parameter $\emptyset = (\alpha, \theta, \lambda, \varphi)^T$ is given by:

$$L(\emptyset) = nlog2 + nlog\alpha + nlog\theta + nlog\lambda + \sum_{i=1}^{n} \log[h(x_i;\varphi)] + (\alpha - 1)\sum_{i=1}^{n} \log[H(x_i;\varphi)] + (2\lambda - 1)\sum_{i=1}^{n} \log(1 - H(x_i;\varphi)^{\alpha}) + (\theta - 1)\sum_{i=1}^{n} \log(1 - (1 - H(x_i;\varphi)^{\alpha})^{2\lambda}) \quad (24)$$

The components of score vector $U = (U_\alpha, U_\theta, U_\lambda, U_\varphi)^T$ are given as

$$U_\alpha = \frac{n}{\alpha} + \sum_{i=1}^{n} \log[H(x_i;\varphi)] + (2\lambda - 1)\sum_{i=1}^{n} \left[\frac{H(x_i;\varphi)^{\alpha} \log[H(x_i;\varphi)]}{1 - H(x_i;\varphi)^{\alpha}} \right] + (\theta - 1)\sum_{i=1}^{n} \left[\frac{2\lambda[1 - H(x_i;\varphi)^{\alpha}]^{2\lambda - 1} H(x_i;\varphi)^{\alpha} \log[H(x_i;\varphi)]}{1 - [1 - H(x_i;\varphi)^{\alpha}]^{2\lambda}} \right] \quad (25)$$

$$U_\lambda = \frac{n}{\lambda} + \sum_{i=1}^{n} log[1 - H(x_i;\varphi)^{\alpha}] + (\theta - 1)\sum_{i=1}^{n} \frac{[1 - H(x_i;\varphi)^{\alpha}]^{2\lambda} log[1 - H(x_i;\varphi)^{\alpha}]}{1 - [1 - H(x_i;\varphi)^{\alpha}]^{2\lambda}} = 0 \quad (26)$$

$$U_\theta = \frac{n}{\theta} + \sum_{i=1}^{n} \log[1 - (1 - H(x_i;\varphi)^{\alpha})^{2\lambda}] = 0 \quad (27)$$

$$U_\varphi = \sum_{i=1}^{n} \left[\frac{h(x_i;\varphi)^{\varphi}}{h(x_i;\varphi)} \right] + (\alpha - 1)\sum_{i=1}^{n} \left[\frac{H(x_i;\varphi)^{\varphi}}{H(x_i;\varphi)} \right] - (2\lambda - 1)\sum_{i=1}^{n} \left[\frac{2\lambda\alpha[1 - H(x_i;\varphi)^{\varphi}]^{2\lambda - 1} H(x_i;\varphi)^{\alpha - 1} H(x_i;\varphi)^{\varphi}}{1 - H(x_i;\varphi)^{\alpha}} \right] - (\theta - 1)\sum_{i=1}^{n} \left[\frac{2\lambda H(x_i;\varphi)^{\alpha - 1} H(x_i;\varphi)^{\varphi}}{1 - [1 - H(x_i;\varphi)^{\alpha}]^{2}} \right] = 0 \quad (28)$$

The solution of the non-linear equations (25), (26), (27) and (28) cannot be obtained analytically so, we resolved to use iterative methods.

5.2.Application to real-life data

This section presents the real-life application of the TLK-G family of distributions to a real-life data set by using the TLKEx distribution as one of the sub-models and for illustrative purposes also present a comparative study with the fits of Topp Leone Exponential (TLEx) distribution by [2], Kumaraswamy Exponential (KEx) distribution by [15], Exponentiated Exponential (ExEx) distribution by [16] and Exponential (Ex) distribution. This application proved empirically the flexibility of the new distribution in modeling real-life data. All the computations are performed using the AdequacyModel package in R software. The pdfs of the competing distributions are:

TLEx distribution

$$f(x;\beta,\theta) = 2\beta\theta e^{-2\beta x}\left(1 - e^{-2\beta x}\right)^{\theta - 1}$$

KEx distribution

$$f(x;\alpha,\beta,\lambda)=\alpha\beta\lambda e^{-\beta x}\left(1-e^{-\beta x}\right)^{\alpha-1}\left[1-\left(1-e^{-\beta x}\right)^{\alpha}\right]^{\lambda-1}$$

ExEx distribution

$$f(x;\beta,\theta)=\beta\theta e^{-\beta x}\left(1-e^{-\beta x}\right)^{\theta-1}$$

Ex distribution

$$f(x;\beta)=\beta e^{-\beta x}$$

Data set 1:

The first data set represents the survival times of a group of patients suffering from Head and Neck cancer diseases and treated using a combination of radiotherapy and chemotherapy (RT+CT). The data set has been previously used by [17] and [18].

The data set has forty-four (44) observations and they are as follows:

12.20, 23.56, 23.74, 25.87, 31.98, 37, 41.35, 47.38, 55.46, 58.36, 63.47, 68.46, 78.26, 74.47, 81.43, 84, 92, 94, 110, 112, 119, 127, 130, 133, 140, 146, 155, 159, 173, 179, 194,195, 209, 249, 281, 319, 339, 432, 469, 519, 633, 725, 817, 1776.

Data set 2

The second data set was taken from [19] and it represents the survival times of one hundred and twenty-one (121) patients with breast cancer obtained from a large hospital in a period from 1929 to 1938. It has also been applied by [20]. The data set is as follows:

0.3, 0.3, 4.0, 5.0, 5.6, 6.2, 6.3, 6.6, 6.8, 7.4, 7.5, 8.4, 8.4, 10.3, 11.0, 11.8, 12.2, 12.3, 13.5, 14.4, 14.4, 14.8, 15.5, 15.7, 16.2, 16.3, 16.5, 16.8, 17.2, 17.3, 17.5, 17.9, 19.8, 20.4, 20.9, 21.0, 21.0, 21.1, 23.0, 23.4, 23.6, 24.0, 24.0, 27.9, 28.2, 29.1, 30.0, 31.0, 31.0, 32.0, 35.0, 35.0, 37.0, 37.0, 37.0, 38.0, 38.0, 38.0, 39.0, 39.0, 40.0, 40.0, 40.0, 41.0, 41.0, 41.0, 42.0, 43.0, 43.0, 43.0, 44.0, 45.0, 45.0, 46.0, 46.0, 47.0, 48.0, 49.0, 51.0, 51.0, 51.0, 52.0, 54.0, 55.0, 56.0, 57.0, 58.0, 59.0, 60.0, 60.0, 60.0, 61.0, 62.0, 65.0, 65.0, 67.0, 67.0, 68.0, 69.0, 78.0, 80.0, 83.0, 88.0, 89.0, 90.0, 93.0, 96.0, 103.0, 105.0, 109.0, 109.0, 111.0, 115.0, 117.0, 125.0, 126.0, 127.0, 129.0, 129.0, 139.0, 154.0

5.2.1.Results

Two sub-models are proposed in this study which corresponds to the TLKEx and TLKLL distributions. TLKEx distribution is used to test the goodness-of-fit of the new family. The results of the analysis are presented in Table 1 and Table 2. From the Table 1 and Table 2, it can be seen that the TLKEx distribution provides a better fit compared to the distributions considered and thus the TLKEx distribution should be preferred to the other distributions mentioned in this work for modeling real-life data-sets. Figure 1 shows the plots of the pdf of the TLKEx distribution with different parameter values and Figure 2 shows the plots of the survival function with different parameter values estimated from the data sets. Figure 2(a) is the survival time of patients suffering from Head and Neck cancer diseases and treated using a combination of radiotherapy and chemotherapy (RT+CT). It can be seen from Figure 2(a) that at time 500 we have only 10% survival rate, at time 1000 we have only 6% survival rate and at time 1500 we have only 4% survival rate. Also, Figure 2(b) is the survival time of patients suffering from breast cancer disease. From Figure 2(b), it can be seen that at time 50 we have 35% survival rate, at time 100 we have 10% survival rate and at time 150 we have 6% survival rate. The histogram of the data set 1 and data set 2 and the plots of the fitted pdfs for the TLKEx distribution are shown in Figure 3 and Figure 4 respectively.

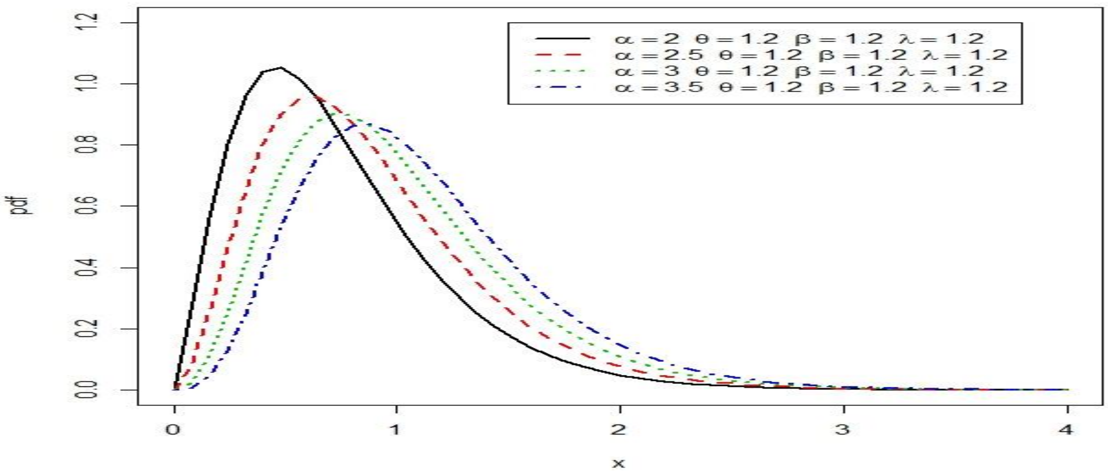

Figure 1. Pdf plots of TLKEx distribution for parameter α varies, other parameters fixed at 1.2

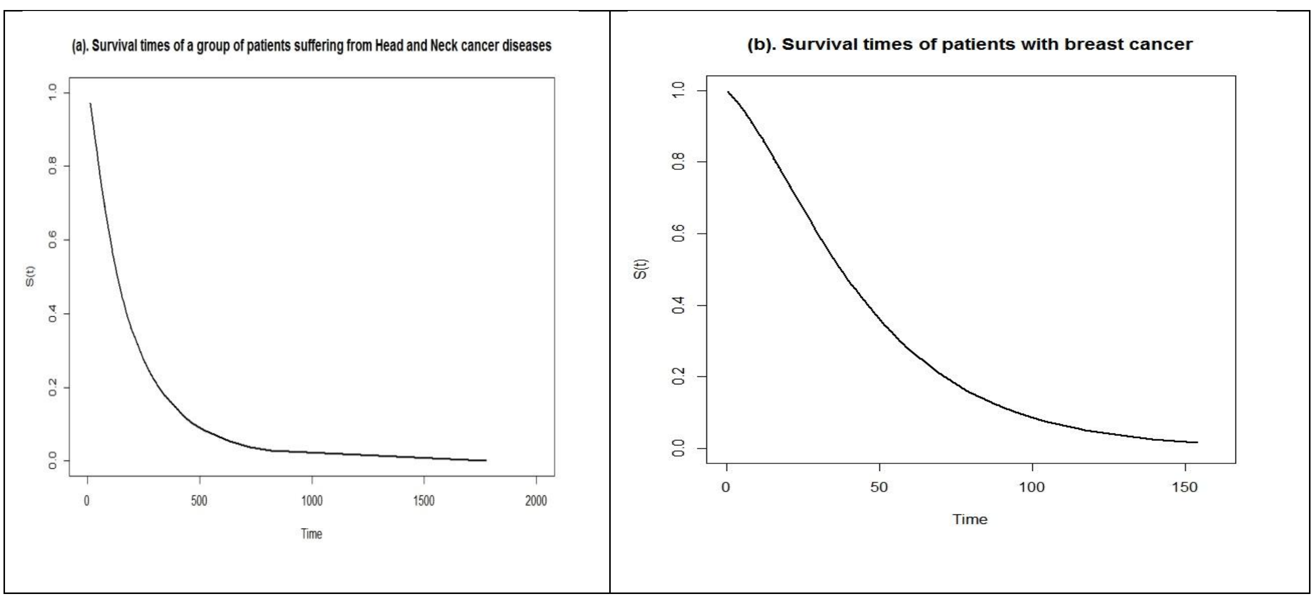

Figure 2. Plots of survival function of the TLKEx distribution with estimated parameters from the data set

Table 1. The ML estimate of the TLKEx distribution parameters and AIC value for Data Set 1

Distributions	Estimates	-LL	AIC
TLKEx	$\hat{\alpha} = 0.3492$ $\hat{\beta} = 0.0005$ $\hat{\theta} = 6.6545$ $\hat{\lambda} = 2.3834$	278.7118	565.4235
KEx	$\hat{\alpha} = 1.5250$ $\hat{\beta} = 0.0202$ $\hat{\lambda} = 0.2385$	281.1617	568.3235
TLEx	$\hat{\beta} = 0.0023$ $\hat{\theta} = 1.0721$	281.9550	567.9101
ExEx	$\hat{\beta} = 0.0025$ $\hat{\theta} = 1.1791$	282.0628	568.1256
Ex	$\hat{\beta} = 0.0045$	282.0135	566.0271

Table 2. The ML estimate of the TLKEx distribution parameters and AIC value for Data Set 2

Distributions	Estimates	-LL	AIC
TLKEx	$\hat{\alpha} = 0.3513$ $\hat{\beta} = 0.2288$ $\hat{\theta} = 2.1179$ $\hat{\lambda} = 0.0663$	578.0122	1164.0830
KEx	$\hat{\alpha} = 1.3859$ $\hat{\beta} = 0.2361$ $\hat{\lambda} = 0.0945$	583.1573	1172.3150
TLEx	$\hat{\beta} = 0.0147$ $\hat{\theta} = 1.5426$	580.7893	1165.5790
ExEx	$\hat{\beta} = 0.0275$ $\hat{\theta} = 1.4836$	581.0982	1166.1120
Ex	$\hat{\beta} = 0.0216$	585.1278	1172.2560

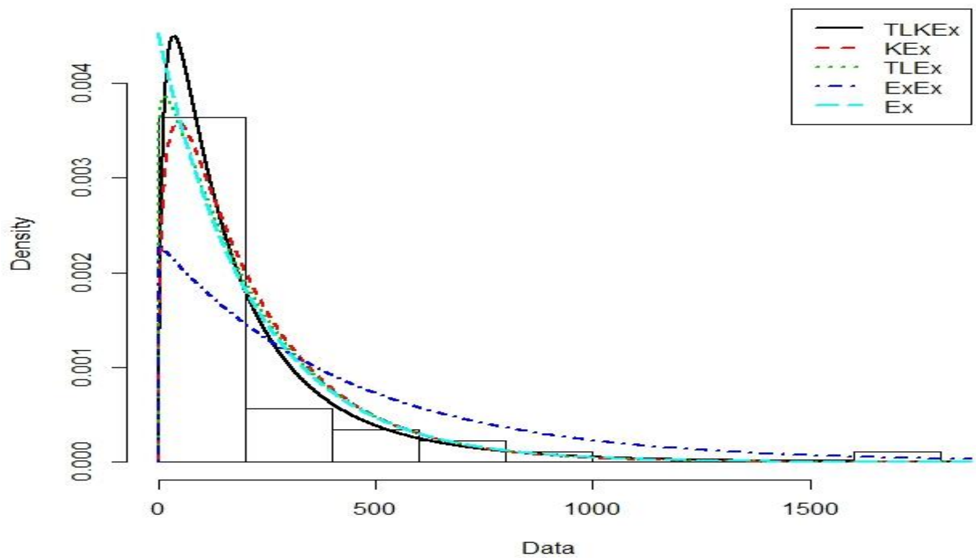

Figure 3. Fitted pdfs for the TLKEx, KEx, TLEx, ExEx and Ex models to the data set 1

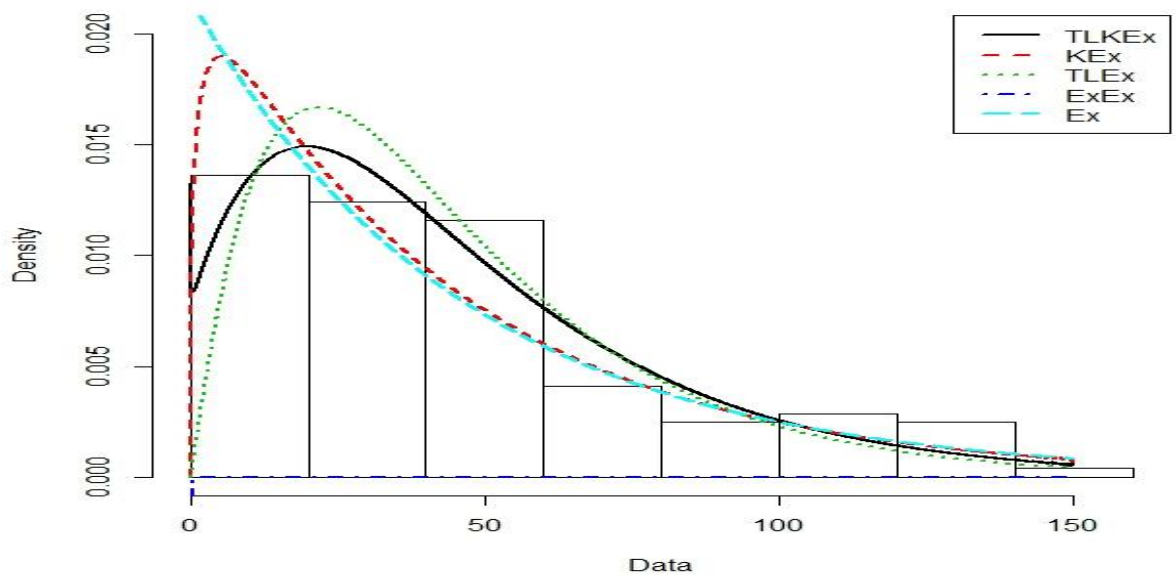

Figure 4. Fitted pdfs for the TLKEx, KEx, TLEx, ExEx and Ex models to the data set 2

Conclusion

In this paper, we have proposed a new family of continuous distributions, called the Topp Leone Kumaraswamy-G family following the idea of the T-X family with the Topp Leone-G class of distributions as the transformer. Two sub-models of the new family are presented and the TLKEx distribution is used to analyze the data. The results of the analysis indicate that the TLKEx distribution has the lowest value of AIC. Therefore, the values of minimum AIC of the distributions shows the superiority of the TLKEx distribution over the other distributions. Hence, the TLKEx distribution can be seen as an important distribution in modeling real-life data.

Acknowledgments
The authors are grateful to God for making this work a success. We would also like to thank the Editor-in-Chief, the Associate Editor and the referees for careful reading and for their comments which greatly improved the paper.

Conflict of Interests

Authors have declared that no conflict of interests exists.

References

1. Cordeiro, G. M., and de Castro, M. (2011). A new family of generalized distributions, Journal of Statistical Computation and Simulation, 81, 7, 883–898.

2.Al-Shomrani, A., Arif, O., Shawky, A., Hanif, S. and Shahbaz, M. Q. (2016). Topp-Leone family of distributions: Some properties and application. Pakistan Journal of Statistics and Operation Research. XII, 3, 443-451.

3.Cordeiro, G. M, Afify, A. Z., Yousof, H. M, Pessim, R. R. and Aryal, G. (2017). The exponentiated Weibull-H family of distributions: Theory and applications. Mediterranean Journal of Mathematics, 14:1-22.

4.Reyad, H., Alizadeh, M., Jamal, F. and Othman, S. (2018). The Topp Leone odd Lindley-G family of distributions: Properties and applications, Journal of Statistics and Management Systems, 21:7, 1273-1297.

5.Aryuyuen, S. (2018). A Topp-Leone Generator of Exponentiated Power Lindley Distribution and Its Application, Applied Mathematical Sciences, 12, 12, 567 – 579.

6.Jamal, F. (2019). The Marshall-Olkin Odd Lindley-G Family of Distributions: Theory and Applications, Punjab University Journal of Mathematics, 51, 7, 111-125.

7.Silva, R., Silva, G. F., Ramos, M., Cordeiro, G., Marinho, P., and De Andrade, T. A. N. (2019). The Exponentiated Kumaraswamy-G Class: General Properties and Application. Revista Colombiana de Estadística, 42, 1, 1-33.

8.Reyad, H., Korkmaz, M. C., Afify, A. Z., Hamedan, G. G. and Othman, S. (2019). The Fréchet Topp Leone-G Family of Distributions: Properties, Characterizations and Applications, Annals of Data Science, doi.org/10.1007/s40745-019-00212-9.

9.Reyad, H., Jamal, F., Othman, S. and Yahia, N. (2019). The Topp Leone Generalized Inverted Kumaraswamy Distribution: Properties and Applications, Asian Research Journal of Mathematics, 13, 3: 1-15, DOI: 10.9734/ARJOM/2019/v13i330107.

10. Hassan S, Nassr SG. (2019). Power Lindley-G Family of Distributions. Annals of Data Science, 6: 189-210.

11. Modi, K., Kumar, D. and Singh, Y. (2020). A New Family of Distribution with Application on Two Real Data sets on Survival Problem, Science and Technology Asia, 25, 1, 1-10.

12.Ibrahim S, Doguwa SI, Audu I and Jibril HM. (2020). On the Topp Leone exponentiated-G Family of Distributions: Properties and Applications. Asian Journal of Probability and Statistics; 7, 1, 1-15.

13. Alzaatreh, A., Lee, C. and Famoye, F. (2013). A new method for generating families of continuous distributions. Metron, 71:63–79

14.Rényi, A. (1961). On measures of entropy and information. In: Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, Berkeley

15.Adepoju, K. A. and Chukwu, O. I. (2015). Maximum Likelihood Estimation of the Kumaraswamy Exponential Distribution with Applications, Journal of Modern Applied Statistical Methods, 14, 1, 208-214. DOI:10.22237/jmasm/1430453820

16.Gupta, R. D. and Kundu, D. (1999). Generalized exponential distributions, Australian and New Zealand Journal of Statistics, 41, 173-188.

17.Efron, B. (1988). Logistic regression, survival analysis and the kaplan-meier curve, Journal of the American Statistical Association, 83, 402: 414-425.

18.Shanker, R., Fesshaye, H. and Selvaraj, S. (2015). On modeling lifetimes data using exponential and lindley distributions, Biometric and Biostatistics International Journal, 2, 5: 1-9.

19.Lee, E. T. (1992). Statistical methods for survival data analysis (2nd Edition), John Wiley and Sons Inc., New York, USA, 156 Pages.

20.Ramos, M. A., Cordeiro, G. M., Marinho, P. D., Dias, C. B. and Hamadani, G. G. (2013). The zografos-balakrishman log-logistic distribution: properties and applications, Journal of Statistical Theory and Applications, 12(3): 225-244.