Türkiye Klinikleri Biyoistatistik Dergisi
Turkiye Klinikleri Journal of Biostatistics Turkiye Klinikleri J Biostat. 2024;16(1):1-15

ORIGINAL RESEARCH ORİJİNAL ARAŞTIRMA DOI: 10.5336/biostatic.2023-99766

A Study on Topp-Leone Kumaraswamy Fréchet Distribution with Applications: Methodological Study

Uygulamalarla Topp-Leone Kumaraswamy Fréchet Dağılımı Üzerine Bir Araştırma: Metodolojik Çalışma

Sule Omeiza BASHIRU[a]

[a]Department of Mathematics and Statistics, Confluence University of Science and Technology, Osara, Kogi State, Nigeria

ABSTRACT Objective: The aim of this study is to introduce and analyze the Topp-Leone Kumaraswamy Fréchet (TLKFr) distribution a new probability distribution model. This model extends the classical Fréchet distribution by introducing three additional positive shape parameters to enhance flexibility and robustness. **Material and Methods:** The novel distribution is based on the Topp-Leone Kumaraswamy generalized distribution family, incorporating three shape parameters. This foundation allows for the derivation of both the cumulative distribution function and probability density function for the new distribution. The study also extensively explores the mathematical properties of this distribution, including the survival, hazard, and quantile functions. Parameter estimation is conducted using the maximum likelihood estimation (MLE) technique. **Results:** The study introduces the TLKFr distribution, providing mathematical expressions for key statistical measures. Parameters were estimated using MLE technique. Plots illustrate the distribution's characteristics, showing positive skewness and reliable survival analysis predictions. Simulation analysis indicates decreasing bias and root mean square error (RMSE) with larger sample sizes, emphasizing reliability. Application to two real-life datasets, demonstrates the TLKFr distribution consistently outperforming rival models, supported by goodness-of-fit statistics. **Conclusion:** The study introduces and analyzes the TLKFr distribution, enhancing the traditional Fréchet distribution with three additional shape parameters. Mathematical characteristics are explored, employing MLE for parameter estimation. A simulation study underscores the distribution's strong reliability, evident in decreasing bias and RMSE with increasing sample size. When tested with real datasets, the TLKFr distribution outperforms alternative models, highlighting its significance in representing real-world phenomena and validating its applicability to the two datasets examined in this study.

Keywords: Maximum likelihood estimation; COVID-19; goodness of fit; Frechet distribution; order statistic

ÖZET Amaç: Bu çalışmanın amacı, yeni bir olasılık dağılım modeli olan Topp-Leone Kumaraswamy Fréchet (TLKFr) dağılımını tanıtmak ve analiz etmektir. Bu model, esnekliği ve sağlamlığı artırmak için üç ek pozitif şekil parametresi uygulayarak klasik Fréchet dağılımını genişletir. **Gereç ve Yöntemler:** Yeni dağılım, üç şekil parametresini içeren Topp-Leone Kumaraswamy genelleştirilmiş dağılım ailesine dayanmaktadır. Bu temel yeni dağılım için hem kümülatif dağılım fonksiyonunun hem de olasılık yoğunluk fonksiyonunun türetilmesine olanak sağlar. Çalışma aynı zamanda sağkalım, tehlike ve niceliksel işlevler de dâhil olmak üzere bu dağılımın matematiksel özelliklerini de kapsamlı bir şekilde araştırmaktadır. Parametre tahmini, maksimum olabilirlik tahmini [the maximum likelihood estimation (MLE)] tekniği kullanılarak gerçekleştirilir. **Bulgular:** Çalışma, temel istatistiksel ölçümler için matematiksel ifadeler sağlayan TLKFr dağılımını tanıtmaktadır. Parametreler, MLE tekniğinin kullanılmasıyla tahmin edilmiştir. Grafikler, pozitif çarpıklığı ve güvenilir sağkalım analizi tahminlerini gösteren dağılımın özelliklerini göstermektedir. Simülasyon analizi, daha büyük örneklem boyutlarıyla azalan sapma ve kök ortalama kare hatasına [root mean square error (RMSE)] işaret ederek güvenilirliği vurgulamaktadır. Gerçek hayattan iki veri setine yapılan uygulama, TLKFr dağılımının uyum iyiliği istatistikleriyle desteklenen rakip modellerden sürekli olarak daha iyi performans gösterdiğini ortaya koymaktadır. **Sonuç:** Bu çalışma, geleneksel Fréchet dağılımını üç ek şekil parametresi ile geliştiren TLKFr dağılımını tanıtmakta ve analiz etmektedir. Parametre tahmini için MLE kullanılarak matematiksel özellikler araştırılmıştır. Bir simülasyon çalışması, artan örneklem büyüklüğü ile azalan sapma ve RMSE ile dağılımın güçlü güvenilirliğinin altını çizmektedir. TLKFr dağılımı, gerçek veri kümeleriyle test edildiğinde alternatif modellerden daha iyi performans göstererek gerçek dünya olaylarını temsil etmedeki önemini vurgulamakta ve bu çalışmada incelenen iki veri kümesine uygulanabilirliğini doğrulamaktadır.

Anahtar kelimeler: Maksimum olabilirlik tahmini; COVID-19; uyumun iyiliği; Fréchet dağılımı; sıra istatistiği

TO CITE THIS ARTICLE:
Bashiru SO. A study on Topp-Leone Kumaraswamy Fréchet distribution with applications: Methodological study. Turkiye Klinikleri J Dental Sci. 2024;16(1):1-15.

Correspondence: Sule Omeiza BASHIRU
Department of Mathematics and Statistics, Confluence University of Science and Technology, Osara, Kogi State, Nigeria
E-mail: bash0140@gmail.com

Peer review under responsibility of Turkiye Klinikleri Journal of Biostatistics.

Received: 02 Oct 2023 ***Received in revised form:*** 05 Feb 2024 ***Accepted:*** 05 Feb 2024 ***Available online:*** 28 Feb 2024

In the realm of distribution theory, researchers have increasingly shifted their focus towards a comprehensive exploration of the tail behavior of various distributions. This evolving research direction is driven by the inherent significance and wide-ranging applicability of such distribution models in portraying real-world phenomena spanning diverse fields and domains.

The pivotal Fréchet (Fr) distribution, originally introduced by, holds a central position within the realm of extreme value theory.[1] As underscored by and further elaborated by, the Fr distribution serves as a versatile tool with applications spanning accelerated life testing, the analysis of rainfall patterns, seismic events, flood modeling, horse racing predictions, wind speed assessments, and the study of sea wave dynamics.[2,3]

Over time, the Fr distribution has undergone various generalizations and extensions within the statistical literature. These extensions have been put forth by a multitude of authors, including but not limited to, and numerous others.[4-19]

The rich tapestry of these diverse extensions and generalizations has significantly broadened the Fr distribution's scope of application, extending its relevance across a wide spectrum of disciplines and problem domains. This collective effort has played a crucial role in advancing both the theoretical and practical aspects of statistical science. Researchers remain committed to further exploration and refinement of these models, aiming to more effectively capture the intricate tail behavior of distributions and their enduring relevance in modeling real-world phenomena.

In recent years, the Fr distribution's widespread adoption has been fueled by advancements in computational techniques and data availability. With the advent of the proliferation of computational tools, researchers can now analyze real-world datasets with unprecedented depth and accuracy. This has opened up new avenues for the application of the Fr distribution in fields such as finance, environmental science, insurance, and reliability engineering.

Furthermore, the development of hybrid models that combine the Fr distribution with other probability distributions has allowed researchers to capture complex phenomena more accurately. These hybrid models, often referred to as mixture distributions or copula-based models, have gained popularity for their ability to handle dependencies and non-standard behavior in data.

The Fr distribution and its myriad extensions continue to be at the forefront of distribution theory, shaping how we understand extreme events and their impact across various domains. As research in this area evolves, it promises to provide valuable insights and practical tools for addressing the challenges posed by extreme events in an increasingly interconnected and data-driven world.

The Fr distribution is characterized by its cumulative distribution function (CDF) and probability density function (PDF) defined as follows:

$$G(x) = e^{-\left(\frac{\alpha}{x}\right)^{b}} \tag{1}$$

and

$$g(x) = \alpha^{b} b x^{-b-1} e^{-\left(\frac{\alpha}{x}\right)^{b}} \tag{2}$$

where $\alpha > 0$ is the scale parameter and $b > 0$ is the shape parameter.

The aim of this paper is to derive and study a new model called Topp-Leone Kumaraswamy Fr (TLKFr) distribution. The new model introduces three additional parameters to (1) in order to enhance flexibility and improve robustness. Utilizing the TLK-G (TLK-G) family introduced by, the new five-parameter TLKFr model is constructed. The TLK-G family of distributions, proposed by, is a powerful framework that extends the flexibility of the traditional Kumaraswamy distribution through the Topp-Leone transformation.[20] This transformation enriches the model's ability to capture various data characteristics, making it a valuable tool for statisticians and researchers. The TLK-G family of distributions can be defined as follows:

$$F(x)=\left[1-\left[1-\left[G(x)\right]^{\lambda}\right]^{2\gamma}\right]^{\theta} \tag{3}$$

where $\gamma>0$, $\lambda>0$ and $\theta>0$ are the three additional shape parameters whose role is to govern the skewness and tail weight. The PDF corresponding to (3) is given as

$$f(x)=2\gamma\lambda\theta g(x)\left[G(x)\right]^{\lambda-1}\left[1-\left[G(x)\right]^{\lambda}\right]^{2\gamma-1}\left[1-\left[1-\left[G(x)\right]^{\lambda}\right]^{2\gamma}\right]^{\theta-1} \tag{4}$$

MATERIAL AND METHODS

TLKFR DISTRIBUTION

By substituting the expression from (1) into (3), the CDF of the TLKFr distribution is given by

$$F(x)=\left[1-\left[1-\left[e^{-\left(\frac{\alpha}{x}\right)^{b}}\right]^{\lambda}\right]^{2\gamma}\right]^{\theta} \tag{5}$$

where $\alpha>0$ is the scale parameter, $\gamma>0$, $\lambda>0$, $b>0$ and $\theta>0$ are the shape parameters respectively. The PDF of the TLKFr distribution is obtained by substituting the expressions from (1) and (2) into (4), and it is given as

$$f(x)=2\gamma\lambda\theta\alpha^{b}bx^{-b-1}\left[e^{-\left(\frac{\alpha}{x}\right)^{b}}\right]^{\lambda}\left[1-\left[e^{-\left(\frac{\alpha}{x}\right)^{b}}\right]^{\lambda}\right]^{2\gamma-1}\left[1-\left[1-\left[e^{-\left(\frac{\alpha}{x}\right)^{b}}\right]^{\lambda}\right]^{2\gamma}\right]^{\theta-1} \tag{6}$$

IMPORTANT REPRESENTATON OF THE PDF

The PDF of the TLKFr distribution in (6) can be expanded as follows

$$\left[1-\left[1-\left[e^{-\left(\frac{\alpha}{x}\right)^{b}}\right]^{\lambda}\right]^{2\gamma}\right]^{\theta-1}=\sum_{i=1}^{\infty}(-1)^{i}\binom{\theta-1}{i}\left[1-\left[e^{-\left(\frac{\alpha}{x}\right)^{b}}\right]^{\lambda}\right]^{2\gamma i}$$

$$\left[1-\left[e^{-\left(\frac{\alpha}{x}\right)^{b}}\right]^{\lambda}\right]^{2\gamma(i+1)-1}=\sum_{j=1}^{\infty}(-1)^{j}\binom{2\gamma(i+1)-1}{j}\left[e^{-\left(\frac{\alpha}{x}\right)^{b}}\right]^{\lambda j}$$

$$f(x)=2\gamma\lambda\theta\alpha^{b}bx^{-b-1}\sum_{i,j=0}^{\infty}(-1)^{i+j}\binom{\theta-1}{i}\binom{2\gamma(i+1)-1}{j}\left[e^{-\left(\frac{\alpha}{x}\right)^{b}}\right]^{\lambda(j+1)} \tag{7}$$

In (7), it is evident that the TLKFr density can be expressed as an infinite linear combination of Fr density functions. Numerous mathematical characteristics of the TLKFr distribution can be readily deduced from the analogous properties of the Fr distribution.

IMPORTANT REPRESENTATON OF THE CDF

The CDF of the TLKFr distribution in (5) can be expanded as follows

$$[F(x)]^{h} = \left[1-\left[1-\left[e^{-\left(\frac{\alpha}{x}\right)^{b}}\right]^{\lambda}\right]^{2\gamma}\right]^{\theta h}$$

$$\left[1-\left[1-\left[e^{-\left(\frac{\alpha}{x}\right)^{b}}\right]^{\lambda}\right]^{2\gamma}\right]^{\theta h} = \sum_{k=0}^{\infty}(-1)^{k}\binom{\theta h}{k}\left[1-\left[e^{-\left(\frac{\alpha}{x}\right)^{b}}\right]^{\lambda}\right]^{2\gamma k}$$

$$\left[1-\left[e^{-\left(\frac{\alpha}{x}\right)^{b}}\right]^{\lambda}\right]^{2\gamma k} = \sum_{p=0}^{\infty}(-1)^{p}\binom{2\gamma k}{p}\left[e^{-\left(\frac{\alpha}{x}\right)^{b}}\right]^{\lambda p}$$

$$F(x) = \sum_{k,p=0}^{\infty}(-1)^{k+p}\binom{\theta h}{k}\binom{2\gamma k}{p}\left[e^{-\left(\frac{\alpha}{x}\right)^{b}}\right]^{\lambda p} \tag{8}$$

The PDF and CDF plot of the TLKFr distribution are shown in Figure 1 below.

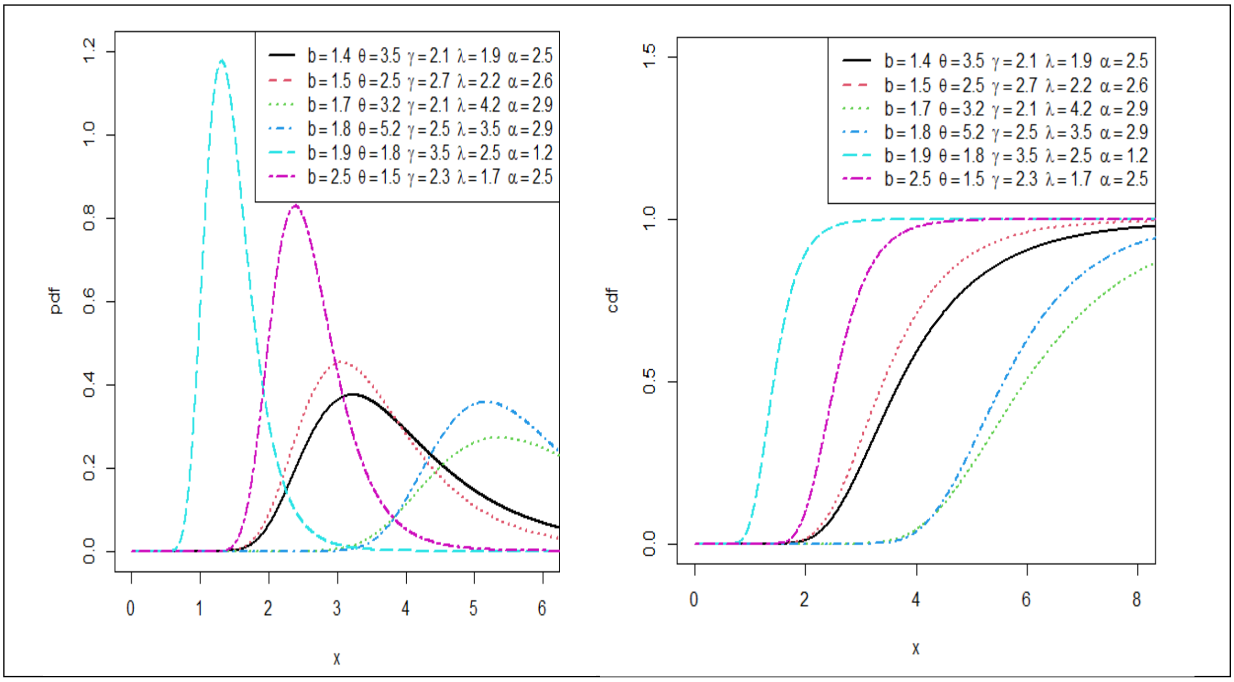

FIGURE 1: Graphs illustrating the distribution's shape through probability density function and cumulative distribution function plots for the Topp-Leone Kumaraswamy Fréchet distribution.

The PDF plot of the TLKFr distribution indicates that the distribution is positively skewed. This skewness is evident from the shape of the PDF curve, which is skewed to the right. The CDF of the proposed distribution indicates a satisfactory level of CDF not exceeding 1 on the y-axis. In other words, the CDF confirms that the TLKFr distribution covers the entire range of possible values and is properly normalized, with the total probability equaling 1.

PROPERTIES OF TLKFR DISTRIBUTION"

In this section, various statistical properties of the TLKFr distribution are deduced. These include the survival function (SF), hazard function (HF), quantile function (QF), moments, moment generating functions (MGF), and order statistics.

SF

The SF represents the probability that a unit, component, or individual will not fail at a given time.[21] It is typically expressed as:

$$sf = 1 - F(x) \tag{9}$$

Therefore, the SF of the TLKFr distribution is derived by substituting (5) into (9), resulting in the following equation:

$$sf = 1 - \left[1 - \left[1 - \left[e^{-\left(\frac{\alpha}{x}\right)^{b}}\right]^{\lambda}\right]^{2\gamma-1}\right]^{\theta} \tag{10}$$

HF

The HF is the instantaneous rate at which a system fails, and it is defined as the ratio of the PDF to the SF. The typical HF as defined by is given as:[22]

$$hf = \frac{f(x)}{sf} \tag{11}$$

By substituting (6) and (10) into (11), the HF of the TLKFr distribution is given as

$$hf = \frac{2\gamma\lambda\theta\alpha^{b}x^{-b-1}\left[e^{-\left(\frac{\alpha}{x}\right)^{b}}\right]^{\lambda}\left[1-\left[e^{-\left(\frac{\alpha}{x}\right)^{b}}\right]^{\lambda}\right]^{2\gamma-1}\left[1-\left[1-\left[e^{-\left(\frac{\alpha}{x}\right)^{b}}\right]^{\lambda}\right]^{2\gamma-1}\right]^{\theta-1}}{1-\left[1-\left[1-\left[e^{-\left(\frac{\alpha}{x}\right)^{b}}\right]^{\lambda}\right]^{2\gamma-1}\right]^{\theta}} \tag{12}$$

The SF and HF plots of the TLKFr distribution are shown in Figure 2 below.

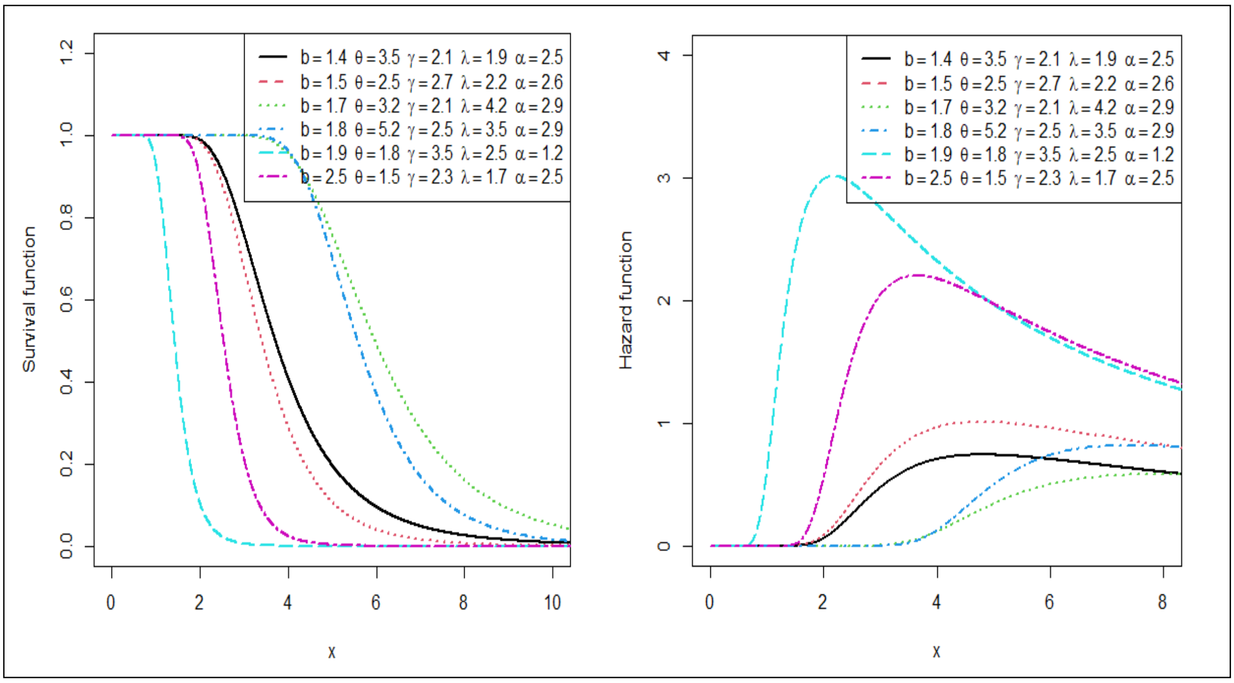

FIGURE 2: Graphs illustrating the shapes of the Topp-Leone Kumaraswamy Fréchet distribution's SF and HF.

SF: Survival function; HF: Hazard function.

The SF of the TLKFr distribution suggests that it is well-suited for modeling survival data sets. The graph of the survival function starts with an initial constant value of 1, indicating the probability that the event has not occurred at the beginning. As time progresses, the survival function decreases, reflecting a reduction in the probability of survival. The hazard function plot provides additional insights into the distribution. The hazard function is monotonically increasing and decreasing based on the values assigned to its parameters. The hazard plot exhibits a unimodal failure rate.

QF

The QF according to can be obtained using the mathematical expression:[23]

$$Q(u) = F^{-1}(u)$$

The QF for the TLKFr distribution is derived as follows:

Set $F(x) = u$,

Where $F(x) = \left[1 - \left[1 - \left[e^{-\left(\frac{\alpha}{x}\right)^{b}}\right]^{\lambda}\right]^{2\gamma}\right]^{\theta}$, then

$$F(x) = \left[1 - \left[1 - \left[e^{-\left(\frac{\alpha}{x}\right)^{b}}\right]^{\lambda}\right]^{2\gamma}\right]^{\theta} = u$$

$$u^{1/\theta} = 1 - \left[1 - \left[e^{-\left(\frac{\alpha}{x}\right)^b}\right]^{\lambda}\right]^{2\gamma}$$

$$\left(1 - u^{1/\theta}\right)^{1/2\gamma} = 1 - \left[e^{-\left(\frac{\alpha}{x}\right)^b}\right]^{\lambda}$$

$$\left[1 - \left(1 - u^{1/\theta}\right)^{1/2\gamma}\right]^{1/\lambda} = e^{-\left(\frac{\alpha}{x}\right)^b}$$

$$\log\left[1 - \left(1 - u^{1/\theta}\right)^{1/2\gamma}\right]^{1/\lambda} = -\left(\frac{\alpha}{x}\right)^b$$

$$-\log\left[1 - \left(1 - u^{1/\theta}\right)^{1/2\gamma}\right]^{1/\lambda} = \left(\frac{\alpha}{x}\right)^b$$

$$x = \alpha\left[-\log\left[1 - \left(1 - u^{1/\theta}\right)^{1/2\gamma}\right]^{1/\lambda}\right]^{-\left(1/b\right)}$$

Therefore, the QF for the TLKFr distribution is given as

$$Q(u) = \alpha\left\{-\log\left(1 - \left(1 - u^{1/\theta}\right)^{1/2\gamma}\right)^{1/\lambda}\right\}^{-\left(1/b\right)} \tag{13}$$

MEDIAN

When u is set to 0.5 in (13), the TLKFr distribution's median is determined as follows:

$$Q(u) = \alpha\left\{-\log\left(1 - \left(1 - 0.5^{1/\theta}\right)^{1/2\gamma}\right)^{1/\lambda}\right\}^{-\left(1/b\right)} \tag{14}$$

MOMENT

Moment is used to study many important properties of distribution such as dispersion, tendency, skewness and kurtosis.[24] The rth ordinary moment of the TLKFr distribution can be derived from (7) as follows:

$$\mu_r^1 = \int_0^\infty x^r f\left(x\right) dx$$

By using the useful expansion of the PDF in (7),

$$= 2\gamma\lambda\theta\alpha^b b x^{-b-1} \sum_{i,j=0}^{\infty} \left(-1\right)^{i+j} \binom{\theta-1}{i} \binom{2\gamma\left(i+1\right)-1}{j} \int_0^\infty x^r \left[e^{-\left(\frac{\alpha}{x}\right)^b}\right]^{\lambda\left(j+1\right)} dx$$

Consider the integral part

$$\int_0^\infty x^r \left[e^{-\left(\frac{\alpha}{x}\right)^b} \right]^{\lambda(j+1)} dx$$

Let, $y = \lambda(j+1)\left(\frac{\alpha}{x}\right)^b \Rightarrow x = \left[\frac{\lambda(j+1)\alpha^b}{y}\right]^{1/b}$

$$dx = \frac{dyx^{b-1}}{\lambda b(j+1)\alpha^b}$$

$$\int_0^\infty \left[\frac{\lambda(j+1)\alpha^b}{y}\right]^{r/b} e^{-y} \frac{dyx^{b-1}}{\lambda b(j+1)\alpha^b}$$

$$= 2\gamma\lambda\theta\alpha^b bx^{-b-1} \sum_{i,j=0}^\infty (-1)^{i+j} \binom{\theta-1}{i}\binom{2\gamma(i+1)-1}{j} \int_0^\infty \left[\frac{\lambda(j+1)\alpha^b}{y}\right]^{r/b} e^{-y} \frac{dyx^{b-1}}{\lambda b(j+1)\alpha^b}$$

$$= 2\gamma\theta\lambda^{r/b}\alpha^r (j+1)^{r/b-1} \sum_{i,j=0}^\infty (-1)^{i+j} \binom{\theta-1}{i}\binom{2\gamma(i+1)-1}{j} \int_0^\infty y^{-r/b} e^{-y} dy$$

Where $\int_0^\infty y^{-r/b} e^{-y} dy = \Gamma\left[1-\frac{r}{b}\right]$

$$\mu_r^1 = 2\gamma\theta\lambda^{r/b}\alpha^r (j+1)^{r/b-1} \sum_{i,j=0}^\infty (-1)^{i+j} \binom{\theta-1}{i}\binom{2\gamma(i+1)-1}{j} \Gamma\left[1-\frac{r}{b}\right] \quad (15)$$

MEAN

By substituting r with 1 in (15), the result is the mean value, which is expressed as follows:

$$\mu_1^1 = 2\gamma\theta\lambda^{1/b}\alpha^1 (j+1)^{1/b-1} \sum_{i,j=0}^\infty (-1)^{i+j} \binom{\theta-1}{i}\binom{2\gamma(i+1)-1}{j} \Gamma\left[1-\frac{1}{b}\right] \quad (16)$$

"MGF"

The mathematical expression of the MGF is given as

$$M_x(t) = \int_0^\infty e^{tx} f(x) dx$$

Consider the expansion $e^{tx} = \sum_{m=0}^\infty \frac{(tx)^m}{m!}$

The MGF of TLKFr distribution is given as

$$M_x\left(t\right)=\frac{\sum_{i,j,m=0}^{\infty}\left(-1\right)^{i+j}\begin{pmatrix}\theta-1\\ i\end{pmatrix}\begin{pmatrix}2\gamma\left(i+1\right)-1\\ j\end{pmatrix}2\gamma\theta\lambda^{r/b}\alpha^{r}\left(j+1\right)^{r/b-1}\Gamma\left[1-\frac{r}{b}\right]}{m!} \tag{17}$$

ORDER STATISTICS

Let $X_1, X_2, ...X_n$ be n "independent random variable from the TLKFr distribution" and let $X_{(1)}, X_{(2)}, ...X_{(n)}$ be their corresponding order statistic. Let $F_{r:n}(x)$ and $f_{r:n}(x)$, $r=1,2,3,...,n$ denote the CDF and PDF of the r^{th} order statistics $X_{r:n}$ respectively. "The PDF of the r^{th} order statistics of $X_{r:n}$" is given as

$$f_{r:n}(x)=\frac{1}{B(r,n-r+1)}\sum_{i=0}^{n-r}(-1)^{i}F(x)^{r+i-1}f(x) \tag{18}$$

Substituting "the CDF and PDF of TLKFr distribution" into (18), the "r^{th} order statistics for the TLKFr distribution is obtained as"

$$f_{r:n}(x)=\frac{2\gamma\lambda\theta\alpha^{b}b}{B(r,n-r+1)}\sum_{i=0}^{n-r}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\frac{(-1)^{i+j+k}\Gamma(\theta(r+i))\Gamma(2\gamma(j+1))}{j!k!\Gamma(\theta(r+i)-j)\Gamma(2\gamma(j+1)-k)}x^{-b-1}\left[e^{-\left(\frac{\alpha}{x}\right)^{b}}\right]^{\lambda k+1}$$

(19)

MAXIMUM ORDER STATISTICS

To derive the PDF for the maximum order statistics of the TLKFr distribution, you can set r equal to n in (19) as follows:

$$f_{n:n}(x)=2n\gamma\lambda\theta\alpha^{b}b\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\frac{(-1)^{j+k}\Gamma(\theta(n+i))\Gamma(2\gamma(j+1))}{j!k!\Gamma(\theta(n+i)-j)\Gamma(2\gamma(j+1)-k)}x^{-b-1}\left[e^{-\left(\frac{\alpha}{x}\right)^{b}}\right]^{\lambda k+1} \tag{20}$$

MINIMUM ORDER STATISTICS

To derive the PDF for the minimum order statistics of the TLKFr distribution, you can set r equal to 1 in (19) as follows:

$$f_{1:n}(x)=2n\gamma\lambda\theta\alpha^{b}b\sum_{i=0}^{n-1}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\frac{(-1)^{i+j+k}\Gamma(\theta(1+i))\Gamma(2\gamma(j+1))}{j!k!\Gamma(\theta(1+i)-j)\Gamma(2\gamma(j+1)-k)}x^{-b-1}\left[e^{-\left(\frac{\alpha}{x}\right)^{b}}\right]^{\lambda k+1} \tag{21}$$

ESTIMATION METHOD

Within this section, the unknown parameters of the TLKFr distribution using the maximum likelihood estimation (MLE) method are estimated. Consider a set of random variables, denoted as X_1, X_2, ..., X_n, each following the TLKFr distribution with a sample size of n. Consequently, the sample log-likelihood function for the TLKFr distribution is derived as follows:

$$\begin{aligned} logL &= n\log(2)+n\log(\gamma)+n\log(\lambda)+n\log(\theta)+nb\log(\alpha)+n\log(b)-(b-1)\sum_{i=1}^{n}\log(x_i) \\ &+\lambda\sum_{i=1}^{n}\log\left[e^{-\left(\frac{\alpha}{x}\right)^b}\right]+(2\gamma-1)\sum_{i=1}^{n}\log\left[1-\left[e^{-\left(\frac{\alpha}{x_i}\right)^b}\right]^{\lambda}\right]+(\theta-1)\sum_{i=1}^{n}\log\left[1-\left[1-\left[e^{-\left(\frac{\alpha}{x_i}\right)^b}\right]^{\lambda}\right]^{2\gamma}\right] \end{aligned} \tag{22}$$

$$\frac{\partial L}{\partial \gamma}=\frac{n}{\gamma}+2\sum_{i=1}^{n}\log\left[1-\left[e^{-\left(\frac{\alpha}{x_i}\right)^b}\right]^{\lambda}\right]-(\theta-1)\sum_{i=1}^{n}\frac{\left[1-\left[e^{-\left(\frac{\alpha}{x_i}\right)^b}\right]^{\lambda}\right]^{2\gamma}\log\left[1-\left[e^{-\left(\frac{\alpha}{x_i}\right)^b}\right]^{\lambda}\right]^{2}}{1-\left[1-\left[e^{-\left(\frac{\alpha}{x_i}\right)^b}\right]^{\lambda}\right]^{2\gamma}} \tag{23}$$

$$\frac{\partial L}{\partial \lambda}=\frac{n}{\lambda}+\sum_{i=1}^{n}\log\left[e^{-\left(\frac{\alpha}{x_i}\right)^b}\right]-(2\gamma-1)\sum_{i=1}^{n}\frac{\left[e^{-\left(\frac{\alpha}{x_i}\right)^b}\right]^{\lambda}\log\left[e^{-\left(\frac{\alpha}{x_i}\right)^b}\right]}{1-\left[e^{-\left(\frac{\alpha}{x_i}\right)^b}\right]^{\lambda}}+2\gamma(\theta-1)\sum_{i=1}^{n}\frac{\left[1-\left[e^{-\left(\frac{\alpha}{x_i}\right)^b}\right]^{\lambda}\right]^{2\gamma-1}\left[e^{-\left(\frac{\alpha}{x_i}\right)^b}\right]^{\lambda}\log\left[e^{-\left(\frac{\alpha}{x_i}\right)^b}\right]}{1-\left[1-\left[e^{-\left(\frac{\alpha}{x_i}\right)^b}\right]^{\lambda}\right]^{2\gamma}} \tag{24}$$

$$\frac{\partial L}{\partial \theta}=\frac{n}{\theta}+\sum_{i=1}^{n}\log\left[1-\left[1-\left[e^{-\left(\frac{\alpha}{x_i}\right)^b}\right]^{\lambda}\right]^{2\gamma}\right] \tag{25}$$

$$\frac{\partial L}{\partial \alpha}=\frac{nb}{\alpha}+\lambda\sum_{i=1}^{n}\frac{e^{-\left(\frac{\alpha}{x_i}\right)^b} b\left(\frac{\alpha}{x_i}\right)^{b-1}\frac{1}{x_i}}{e^{-\left(\frac{\alpha}{x_i}\right)^b}}+(2\gamma-1)\sum_{i=1}^{n}\frac{\lambda\left[e^{-\left(\frac{\alpha}{x_i}\right)^b}\right]^{\lambda-1} e^{-\left(\frac{\alpha}{x_i}\right)^b} b\left(\frac{\alpha}{x_i}\right)^{b-1}\frac{1}{x_i}}{1-\left[e^{-\left(\frac{\alpha}{x_i}\right)^b}\right]^{\lambda}}$$

$$+(\theta-1)\sum_{i=1}^{n}\frac{2\gamma\left[1-\left[e^{-\left(\frac{\alpha}{x_i}\right)^b}\right]^{\lambda}\right]^{2\gamma-1}\lambda\left[e^{-\left(\frac{\alpha}{x_i}\right)^b}\right]^{\lambda-1} e^{-\left(\frac{\alpha}{x_i}\right)^b} b\left(\frac{\alpha}{x_i}\right)^{b-1}\frac{1}{x_i}}{1-\left[1-\left[e^{-\left(\frac{\alpha}{x_i}\right)^b}\right]^{\lambda}\right]^{2\gamma}} \tag{26}$$

$$\frac{\partial L}{\partial b}=\frac{n}{b}+n\log(\alpha)-\sum_{i=1}^{n}\log(x_i)+\lambda\sum_{i=1}^{n}\frac{e^{-\left(\frac{\alpha}{x_i}\right)^b}\left(\frac{\alpha}{x_i}\right)^b\log\left(\frac{\alpha}{x_i}\right)}{e^{-\left(\frac{\alpha}{x_i}\right)^b}}-(2\gamma-1)\sum_{i=1}^{n}\frac{\lambda\left[e^{-\left(\frac{\alpha}{x_i}\right)^b}\right]^{\lambda-1} e^{-\left(\frac{\alpha}{x_i}\right)^b}\left(\frac{\alpha}{x_i}\right)^b\log\left(\frac{\alpha}{x_i}\right)}{1-\left[e^{-\left(\frac{\alpha}{x_i}\right)^b}\right]^{\lambda}}$$

$$+(\theta-1)\sum_{i=1}^{n}\frac{2\gamma\left[1-\left[e^{-\left(\frac{\alpha}{x_i}\right)^b}\right]^{\lambda}\right]^{2\gamma-1}\lambda\left[e^{-\left(\frac{\alpha}{x_i}\right)^b}\right]^{\lambda} e^{-\left(\frac{\alpha}{x_i}\right)^b}\left(\frac{\alpha}{x_i}\right)^b\log\left(\frac{\alpha}{x_i}\right)}{1-\left[1-\left[e^{-\left(\frac{\alpha}{x_i}\right)^b}\right]^{\lambda}\right]^{2\gamma}} \tag{27}$$

By equating expressions (23) through (27) to zero and solving them concurrently, the MLE are obtained. It's worth noting that these equations pose a non-linear challenge and do not lend themselves to straightforward analytical solutions. To address this, statistical software can be effectively employed to numerically solve these equations using iterative methods, such as the Newton-Raphson algorithm in R.

APPLICATION

Simulation Study

To evaluate the performance of the newly proposed TLKFr distribution, a simulation study is carried out. The simulation generates synthetic data by utilizing the QF defined in (13) for various sample sizes, including n=20, 50, 100, 200, 500, and 1000. Table 1 presents the estimation results, bias, and root mean square error (RMSE) obtained from the new distribution for the following combinations of parameters: (α, γ, λ, b , θ)=1.4, 0.4, 1, 0.9, 1.1). The results are replicated 10,000 times and the average result are presented in Table 1 below.

TABLE 1: Simulation results for varying sample sizes and parameters of the Topp-Leone Kumaraswamy Fréchet distribution.

N	Properties	α=1.4	γ=0.4	λ=1	b=0.9	θ=1.1
20	Est	1.4474	0.4277	1.0896	1.0358	1.0871
	Bias	0.0474	0.0277	0.0896	0.1358	-0.0128
	RMSE	0.2530	0.2044	0.3065	0.3704	0.2991
50	Est	1.4396	0.4166	1.0514	0.9576	1.0860
	Bias	0.0396	0.0166	0.0514	0.0576	-0.0139
	RMSE	0.2088	0.1469	0.2490	0.2368	0.2541
100	Est	1.4283	0.4055	1.0320	0.9313	1.0788
	Bias	0.0383	0.0055	0.0320	0.0313	-0.0212
	RMSE	0.1747	0.1042	0.1836	0.1612	0.2044
200	Est	1.4172	0.4043	1.0173	0.9100	1.0756
	Bias	0.0376	0.0047	0.0274	0.0100	-0.0214
	RMSE	0.1517	0.0790	0.1321	0.1146	0.1525
500	Est	1.4068	0.4035	1.0025	0.9038	1.0683
	Bias	0.0368	0.0035	0.0025	0.0038	-0.0117
	RMSE	0.1184	0.0532	0.0925	0.0784	0.1062
1000	Est	1.4013	0.4027	1.0009	0.8977	1.0634
	Bias	0.0318	0.0027	0.0006	-0.0023	-0.0616
	RMSE	0.0980	0.0390	0.0680	0.0021	0.0819

RMSE: Root mean square error.

The simulation results in Table 1 indicate that as the sample size (n) increases, both bias and RMSE decrease, demonstrating improved accuracy in the estimation process. This suggests that the TLKFr model is a consistent distribution.

REAL LIFE APPLICATIONS

In this section, the versatility of the TLKFr distribution is illustrated through two practical applications to real-world datasets. The goodness-of-fit statistics of the model are evaluated, and comparisons are made with other competing models, specifically the KFr model introduced by, the exponentiated Fréchet model proposed by, as well as the Fr and inverse Rayleigh models by and.[1,5,9,25]

Dataset 1:

The first dataset, sourced from a prior study by, represents the daily count of confirmed coronavirus disease-2019 (COVID-19) cases in Pakistan from March 24 to April 28, 2020. The values denote the number of cases reported each day during this period (unit: count).[26]

[Dataset 1: Daily Count of Confirmed COVID-19 Cases in Pakistan]

108, 102, 133, 170, 121, 99, 236, 178, 250, 161, 258, 172, 407, 577, 210, 243, 281, 186, 254, 336, 342, 269, 520, 414, 463, 514, 427, 796, 555, 742, 642, 785, 783, 605, 751, 806.

Dataset 2:

The second dataset, employed in previous studies by and, pertains to the fatigue fracture duration of Kevlar 373/epoxy specimens. These durations were observed under sustained pressure at the 90% stress threshold until the point of failure (unit: hours).[27,28]

[Dataset 2: Fatigue Fracture Duration of Kevlar 373/Epoxy Specimens]

0.0251, 0.0886, 0.0891, 0.2501, 0.3113, 0.3451, 0.4763, 0.5650, 0.5671, 0.6566, 0.6748, 0.6751, 0.6753, 0.7696, 0.8375, 0.8391, 0.8425, 0.8645, 0.8851, 0.9113, 0.9120, 0.9836, 1.0483, 1.0596, 1.0773, 1.1733, 1.2570, 1.2766, 1.2985, 1.3211, 1.3503, 1.3551, 1.4595, 1.4880, 1.5728, 1.5733, 1.7083, 1.7263, 1.7460, 1.7630, 1.7746, 1.8275, 1.8375, 1.8503, 1.8808, 1.8878, 1.8881, 1.9316, 1.9558, 2.0048, 2.0408, 2.0903, 2.1093, 2.1330, 2.2100, 2.2460, 2.2878, 2.3203, 2.3470, 2.3513, 2.4951, 2.5260, 2.9911, 3.0256, 3.2678, 3.4045, 3.4846, 3.7433, 3.7455, 3.9143, 4.8073, 5.4005, 5.4435, 5.5295, 6.5541, 9.0960.

The goodness-of-fit tables, presented as Table 2 and Table 3 below, provide a comparison of the TLKFr distribution with other distributions for the first and second data sets. Similarly, Figure 1 and Figure 2 below offer visual comparisons.

TABLE 2: Goodness-of-fit metrics for the first dataset.

Models	$\hat{\alpha}$	$\hat{b}$	$\hat{\theta}$	$\hat{\lambda}$	$\hat{\gamma}$	ll	AIC	BIC
TLKFr	6.3355	0.5718	3.7898	12.1010	2.8980	-243.6300	497.2689	505.1865
KFr	0.0193	0.1716	0.2201	0.0237	-	-351.6338	711.2676	717.6017
EFr	59.5055	-	14.1287	0.0404	-	-261.1443	528.2886	533.0391
Fr	228.9160	1.6367	-	-	-	-247.8928	499.7856	505.5712
IRa	-	91.4086	-	-	-	-278.5113	559.0226	560.6062

TLKFr: Topp-Leone Kumaraswamy Fréchet; EFr: Exponentiated Fréchet; IRa: Inverse Rayleigh; AIC: Akaike Information Criterion; BIC: Bayesian Information Criterion.

TABLE 3: Goodness-of-fit metrics for the second dataset.

Models	$\hat{\alpha}$	$\hat{b}$	$\hat{\theta}$	$\hat{\lambda}$	$\hat{\gamma}$	ll	AIC	BIC
TLKFr	1.9789	0.2142	7.2732	1.6798	6.7791	-136.9194	283.8389	291.1618
KFr	0.2384	2.6052	0.4666	3.0083	-	-201.1086	410.2172	415.0472
EFr	1.1463	-	0.7180	1.4409	-	-148.6271	303.2465	310.2465
Fr	0.8209	0.7588	-	-	-	-153.5392	311.0784	315.7399
IRa	-	0.1988	-	-	-	-345.9147	693.8293	696.1601

TLKFr: Topp-Leone Kumaraswamy Fréchet; EFr: Exponentiated Fréchet; IRa: Inverse Rayleigh; AIC: Akaike Information Criterion; BIC: Bayesian Information Criterion.

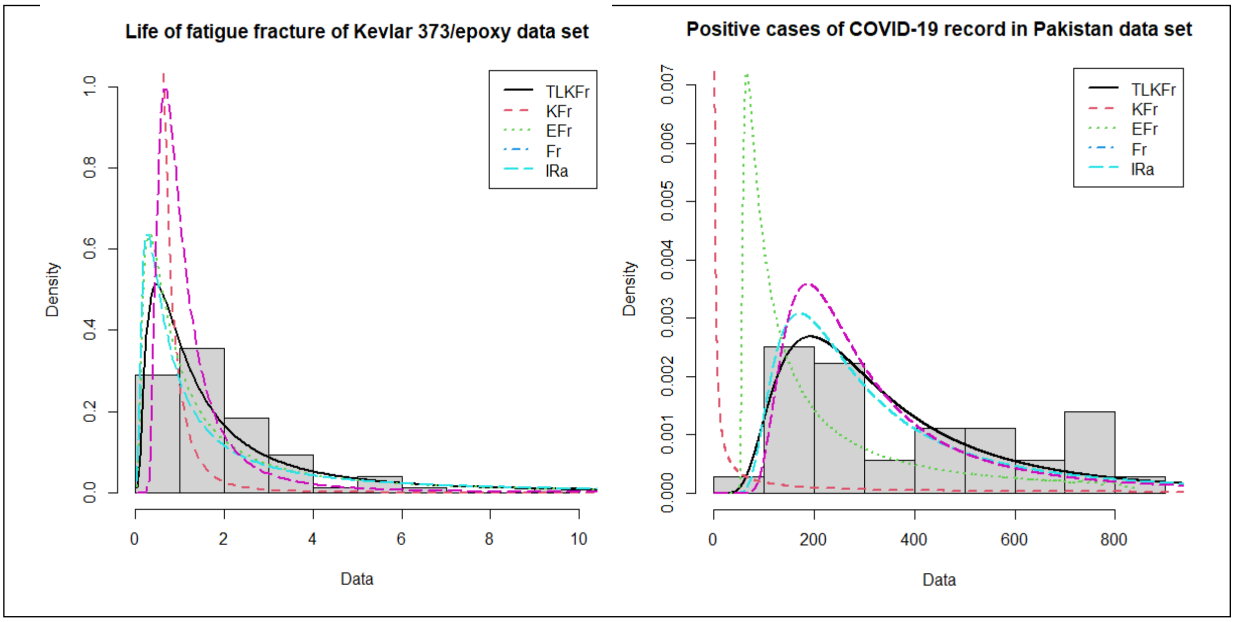

FIGURE 3: Histograms and density plots depicting "Data Set 1 and Data Set 2".

It is observed that the TLKFr model consistently yields the lowest goodness-of-fit scores, as indicated by the negative log-likelihood, Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC) metrics, across the two datasets, as shown in Table 2 and Table 3. This suggests that the TLKFr distribution performs better than the other competing distributions in terms of fitting the data. Furthermore, the histogram plots displayed in Figure 3 provide additional support for this conclusion, indicating that the TLKFr distribution exhibits greater flexibility compared to its competitors.

RESULTS

The TLKFr distribution is introduced in this study, accompanied by the derivation of its cumulative distribution function (CDF) and probability density function (PDF) in Equations 5 and 6. Various statistical properties of the TLKFr distribution, such as the survival function, hazard function, quantile function, median, moments, mean, and linear representation, are established. Parameters for the TLKFr distribution are estimated through the maximum likelihood estimation method. Figure 1 presents the PDF and CDF plots for the TLKFr distribution, demonstrating positive skewness in the PDF and a CDF that maintains a satisfactory level, not exceeding 1 on the y-axis. The hazard graph of the TLKFr distribution displays a monotonically increasing and decreasing hazard function based on assigned parameter values. Furthermore, the survival function graph starts with an initial constant value of 1, representing the probability that the event has not occurred, and gradually decreases over time. A simulation analysis is conducted to assess the performance of the TLKFr distribution. The analysis involves simulating random variables with different parameter values and sample sizes (20, 50, 100, 200, 500, and 1000) using the expression in (13). Table 1 presents the maximum likelihood estimates along with their respective Bias and Root Mean Squared Error (RMSE). Notably, the maximum likelihood estimates for the TLKFr distribution consistently exhibit decreasing bias and RMSE with increasing sample sizes, indicating their reliability.

DISCUSSION

The results of the study highlight the effectiveness of the TLKFr distribution in various aspects. The distribution's positive skewness, demonstrated in the PDF plot, suggests a valuable characteristic for modelling datasets that are positively skewed. The satisfactory behavior of the CDF further supports the applicability of the TLKFr distribution. The hazard function's pattern, showing both increasing and decreasing phases, adds nuance to the understanding of how events unfold over time under the TLKFr distribution. Similarly, the survival function's initial constant value and subsequent decrease provide insights into the probability of event occurrence over time. The simulation analysis reinforces the reliability of the maximum likelihood estimates for the TLKFr distribution, with decreasing bias and RMSE as sample sizes increase. This indicates the robustness of the distribution's parameter estimation, enhancing its practical utility. Application of the TLKFr distribution to real-life datasets, specifically those related to COVID-19 cases in Pakistan and fatigue fracture duration of Kevlar 373/epoxy specimens, showcases the distribution's superior performance compared to rival models. The consistently lower values of information criteria, such as the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and negative log-likelihood function (NLL), further support the TLKFr distribution's efficacy.

CONCLUSION

The introduction of the TLKFr distribution in this study represents a noteworthy contribution to the field of probability distribution modeling. This novel distribution extends the classical Fr distribution by introducing three positive shape parameters, a transformation made possible through the innovative application of the Topp-Leone-G family of distributions, as originally proposed by.[20] This augmentation not only breathes new life into the foundational Fr distribution but also augments its flexibility, rendering it even more adaptable to a wide range of statistical applications. Throughout the investigation, fundamental properties of the novel distribution were derived, spanning ordinary moments, generating functions, SFs, HFs, QFs, and medians. These analytical examinations have revealed that the TLKFr distribution possesses tractable mathematical properties, which greatly enhance its utility in real-world scenarios. A simulation study was carried out to assess the reliability of the distribution. The results of the simulation study further underscore the distribution's high reliability, as evidenced by the reduction in bias and RMSE with increasing sample size Furthermore, this research has demonstrated the practical viability of this model through parameter estimation using the MLE technique. This methodological approach facilitates the accurate estimation of distribution parameters, ensuring that the model aligns closely with empirical data. To underscore the practical relevance of the TLKFr distribution, it was applied to two real-world datasets, a crucial step in gauging its performance. The outcomes of these empirical applications have affirmed the superiority of the new model when compared to existing alternatives. Importantly, the TLKFr distribution demonstrated a superior fit to the two real-world datasets considered in this study, validating its effectiveness and appropriateness for representing complex phenomena. This study highlights that the TLKFr distribution stands as a robust and valuable addition to the family of probability distribution models. Its flexibility, mathematical tractability, and superior performance make it a promising tool for researchers and practitioners seeking to model and analyze real-world phenomena with precision and effectiveness. The distribution's validity for the two datasets considered in this study further emphasizes its potential and value. The journey of exploring and harnessing the potential of this distribution is ongoing, promising exciting developments and applications in the field of statistical science.

Source of Finance

During this study, no financial or spiritual support was received neither from any pharmaceutical company that has a direct connection with the research subject, nor from a company that provides or produces medical instruments and materials which may negatively affect the evaluation process of this study.

Conflict of Interest

No conflicts of interest between the authors and/or family members of the scientific and medical committee members or members of the potential conflicts of interest, counseling, expertise, working conditions, share holding and similar situations in any firm.

Authorship Contributions

This study is entirely author's own work and no other author contribution.

REFERENCES

1. Fréchet M. Sur la loi des erreurs dlobservation. Bulletin de la Société Mathématique de Moscou. 1924;33:5-8. [Link]
2. Kotz S, Nadarajah S. Extreme Value Distributions: Theory and Applications. 1st ed. London: Imperial College Press; 2000. [Crossref]
3. Mubarak M. Parameter estimation based on the Fréchet progressive type II censored data with binomial removals. Journal of Quality, Statistics and Reliability. 2012:1-5. [Crossref]
4. Nadarajah S, Gupta AK. The beta Fréchet distribution. Far East Journal of Theoretical Statistics. 2004;1(4):15-24. [Link]
5. Nadarajah S, Kotz S. The exponentiated Fréchet distribution. Interstat Electronic Journal. 2003:1-7. [Link]
6. Mahmoud MR, Mandouh RM. On the transmuted Fréchet distribution. Journal of Applied Sciences Research. 2013;9(10):5553-61. [Link]
7. Krishna E, Jose KK, Alice T, Ristic MM. The Marshall-Olkin Fréchet distribution. Communications in Statistics-Theory and Methods. 2013;42:4091-107. [Crossref]
8. da Silva RV, de Andrade TA, Maciel DBM, Campos RPS, Cordeiro GM. A new lifetime model: the gamma extended Fréchet distribution. Journal of Statistical Theory and Applications. 2013;12:39-54. [Crossref]
9. Mead ME, Abd-Eltawab AR. A note on Kumaraswamy Fréchet distribution. Australian Journal of Basic and Applied Sciences. 2014;8(15):294-300. [Link]
10. Afify AZ, Hamedani GG, Ghosh I, Mead ME. The transmuted Marshall-Olkin Fréchet distribution: properties and applications. International Journal of Statistics and Probability. 2015;4(4):132-48. [Crossref]
11. Afify AZ, Yousof HM, Cordeiro GM, Ortega EMM, Nofal ZM. The Weibull Fréchet distribution and its applications. Journal of Applied Statistics. 2016;43(14):2608-26. [Crossref]
12. Mead ME, Afify AZ, Hamedani GG, Ghosh I. The beta exponential Fréchet distribution with applications. Austrian Journal of Statistics. 2017;46(1):41-63. [Crossref]
13. Ramos PL, Nascimento D, Louzada F. The long term fréchet distribution: estimation, properties and its application. Biom Biostat Int J. 2017;6(3):357-62. [Crossref]
14. Tablada CJ, Cordeiro GM. The modified Fréchet distribution and its properties. Communications in Statistics-Theory and Methods. 2017;46(21):10617-39. [Crossref]
15. Mansour MM, Abd Elrazik EM, Altun E, Afify AZ, Iqbal Z. A new three-parameter Fréchet distribution: properties and applications. Pak J Statist. 2018;34(6):441-58. [Link]
16. Mansour MM, Aryal G, Afify AZ, Ahmad M. The kumaraswamy exponentiated Fréchet distribution. Pak J Statist. 2018;34(3):177-93. [Link]
17. Abouelmagd THM, Hamed MS, Afify AZ, Al-Mofleh H, Iqbal Z. The Burr X Fréchet distribution with its properties and applications. Journal of Applied Probability and Statistics. 2018;13(1):23-51. [Link]
18. Ramos PL, Louzada F, Ramos E, Dey S. The Fréchet distribution: estimation and application-an overview. Journal of Statistics and Management Systems. 2019:1-24. [Link]
19. Hamed MS, Aldossary F, Afify AZ. The four-parameter Fréchet distribution: properties and applications. Pakistan Journal of Statistics and Operation Research. 2020;16(2):249-64. [Crossref]
20. Ibrahim S, Doguwa SI, Audu I, Jibril HM. The Topp leone kumaraswamy-g family of distributions with applications to cancer disease data. Journal of Biostatistics and Epidemiology. 2020;6(1):37-48. [Link]
21. Sule BO, Halid OY. On gompertz exponentiated inverse rayleigh distribution. Reliability: Theory & Applications. 2023;18(1):412-24. [Link]
22. Sule OB. A new extended generalized inverse exponential distribution: properties and applications. Asian Journal of Probability and Statistics. 2021;11(2):30-46. [Crossref]
23. Isa AM, Sule OB, Alhaji AB, Adepoju AA, Itopa II. Sine-exponential distribution: its mathematical properties and application to real datasets. UMYU Scientifica, 2022;1(1):127-31. [Crossref]
24. Halid OY, Sule OB. A classical and bayesian estimation techniques for gompertz inverse rayleigh distribution: properties and applications. Pakistan Journal of Statistics. 2022;38(1):49-76. [Link]
25. Trayer VN. Inverse Rayleigh (IR) model. Proceedings of the Academy of Science. Belarus: USSR; 1964. [Link]
26. Al-Marzouki S, Jamal F, Chesneau C, Elgarhy M. Topp-leone odd frechet generated family of distributions with applications to COVID-19. Data Sets, Computer Modeling in Engineering and Sciences. 2020;125(1):437-58. [Crossref]
27. Andrews DF, Herzberg AM. Data: A Collection of Problems from Many Fields for the Student and Research Worker. New York: Springer Series in Statistics; 1985.
28. Barlow RE, Toland RH, Freeman T. A Bayesian analysis of stress rupture life of Kevlar 49/epoxy spherical pressure vessels. Procedings Conference on Applications of Statistics. New York: Marcel Dekker; 1984.