Original Article

ON THE LEHMAN TYPE-2 TOP-LEONE TYPE-2 FRÉCHET DISTRIBUTION AND ITS USE IN MODELING CANCER REMISSION DURATIONS

Prof. Funke Adebisi Olumide

Department of Statistics, Federal Polytechnic Ado-Ekiti, Ekiti-State.

DOI: https://doi.org/10.5281/zenodo.17106439

Abstract: In this article we develop a new four-parameters model called the Lehman type II Top-Leone Fréchet (LT-2TLT-2F) distribution which exhibits non-monotone hazard rate. Many models such as Lehman type II Fréchet (LTIIF), Type II Top-Leone Fréchet (TIITLF), Generalized Exponentiated Fréchet (GEF), and Fréchet (F) are sub models. Some of its properties including moment, reliability, moment generating function, Incomplete moments, and hazard rate are investigated. The method of maximum likelihood is proposed to estimate the model parameters. Moreover, we give the advantage of the LT-2TLT-2F distribution by an application using two real datasets.

Keywords: moments, moment generating function, non-monotone, incomplete moments.

1.0 Introduction

Statistical distributions are very useful in describing and predicting real data analysis. Although many distributions have been developed, there are always techniques for developing distributions which are flexible for fitting real data analysis. The Fréchet distribution has found wide applications in extreme value theory. Some extensions of the Fréchet distribution are suggested to attract representing various types of data. In this article, we introduce and study mathematical properties of a new model referred to as the Lehman type II Top-Leone Type II Fréchet distribution represents a special case of the new model. We hope that it will attract wider applications in many other areas of scientific research. Some extensions of the Fréchet distribution are available in the literature, see for example [1–6]. Consider the cumulative distribution function (cdf) and probability density function (pdf) Lehman Type II Type II Top-Leone Fréchet distribution with cdf given by

$G(x) = 1 - [1 - (e^{-bx-\lambda})^2]^{av}$, (1) with corresponding pdf given by

$g(x) = 2abv\lambda bx^{-\lambda-1}(e^{-bx-\lambda})^2[1 - (e^{-bx-\lambda})^2]^{va-1}$, (2)

where a, λ and v are the two added shape parameters and b is a positive scale parameter.

The survival and the hazard function are given by

$S(x) = 1 - G(x) = [1 - (e^{-bx-\lambda})^2]^{av}$ (3)

and

$g(x) \quad 2abv\lambda bx_{-\lambda-1}(e_{-bx-\lambda})_2$

Original Article

$h(x) =$ ____ ____________________ $s(x) =$ $-bx-\lambda$ $)2)$

(3.14)

$(1 - (e$

The graphs of the cdf, pdf, s(x) and h(x) are respectively given in figures 1, 2, 3, and 4 respectively as

Graph of distribution function of LT-2TLT-2F distribution

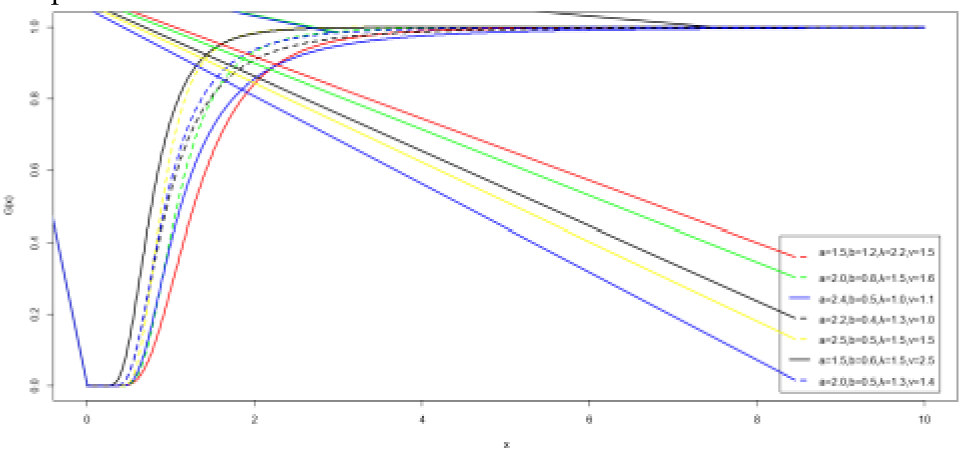

Figure 1. Graph of the distribution function of LT-2TLT-2FD

- Figure1 indicates that the Lehman Type-2 Top-Leone Type-2 Fréchet distribution has a proper probability density function which converges to one upon integration.

Graph of density function of LT-2TLT-2F distribution

Original Article

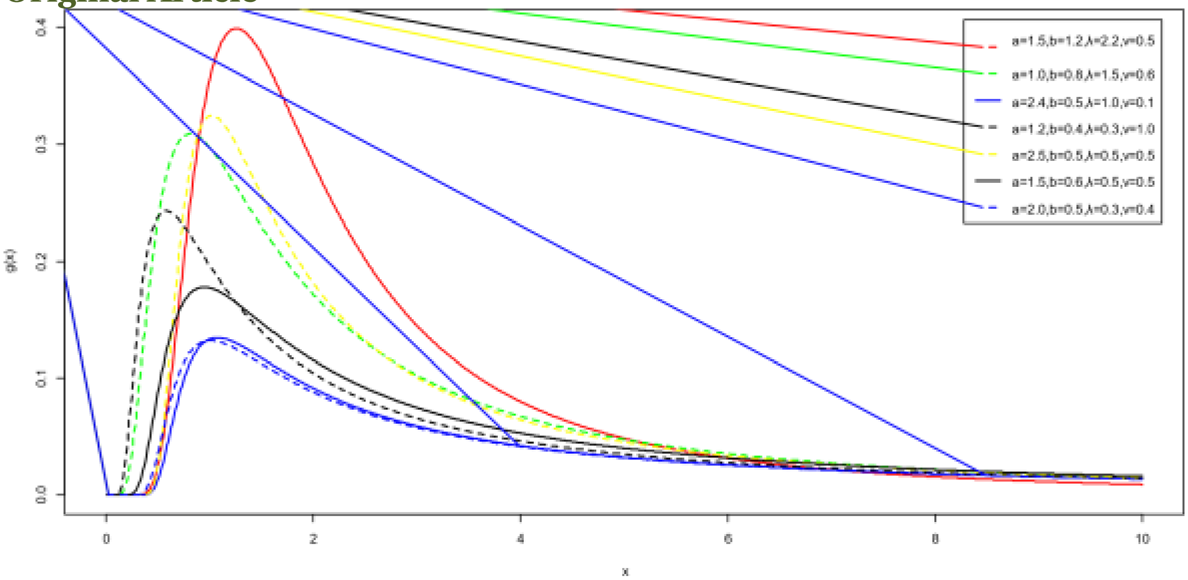

Figure 3.2. Graph of the density function of LT-2TLT-2FD

- Figure 2 indicates that the Probability density function of Lehman Type-2 Top-Leone Type-2 Fréchet distribution is non-monotone.

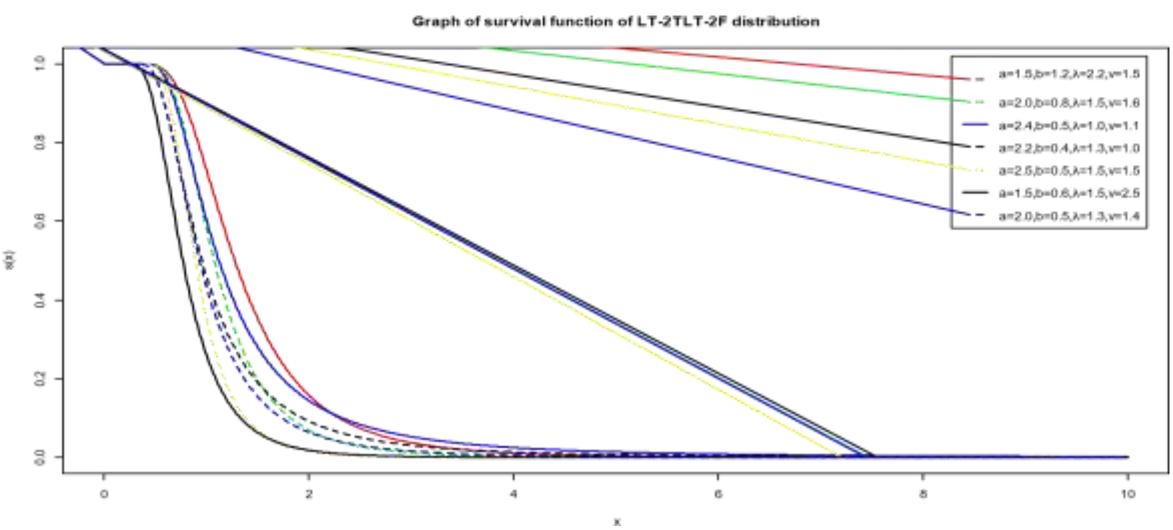

Figure 3. Graph of the survival function of LT-2TLT-2FD

- Figure 3.3 indicates that the survival function of Lehman Type-2 Top-Leone Type-2 Fréchet distribution approaches zero as time increases.

Original Article

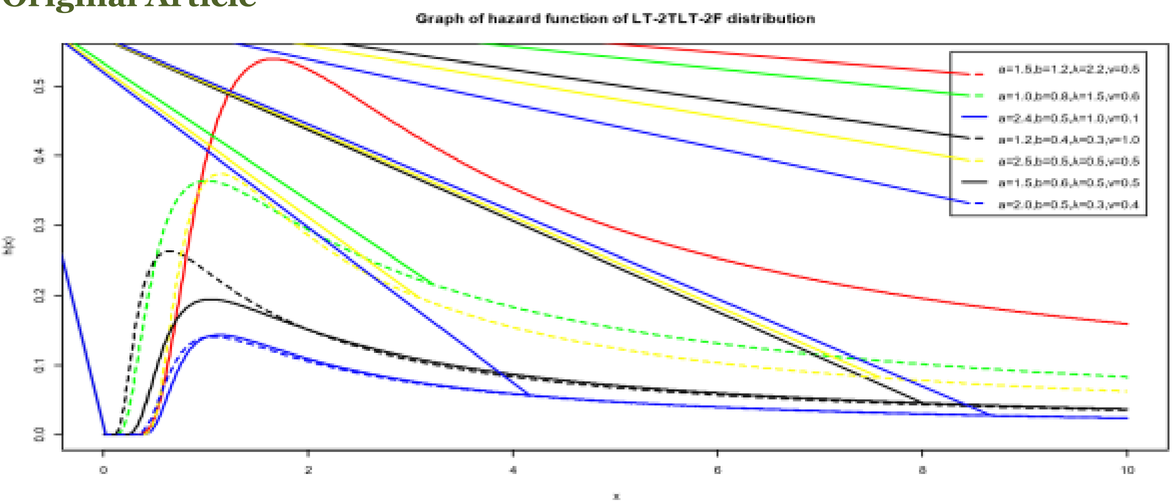

Figure 4. Graph of the hazard function of LT-2TLT-2FD

Figure 4 indicates that the shape of the hazard function of Lehman Type-2 Top-Leone Type-2 Fréchet distribution can be increasing, decreasing, non-monotonic and inverted bathtub failure rates.

A statistical expression for the reversed hazard $\gamma(x)$ and the cumulative hazard $H(x)$ functions is given by

$2abv\lambda bx_{-\lambda-1}(e_{-bx^{-\lambda}})_2[1-(e_{-bx^{-\lambda}})_2]_{va-1}$

$$\gamma(x) = \underline{\hspace{8cm}} 1-[1-(e^{-bx}{}_{-\lambda})_2]_{av}$$

(5)

and

$$H(x) = log[F(x)] = log\,(1-[1-(e^{-bx_{-\lambda}})^2]^{av}) \qquad (6)$$

3.0 Statistical properties of the $LT-2TLT-2FD$ The LT − 2TLT − 2FD can be re-written to a reduced a model using generalized binomial series.

$$(1-w)^j = \sum_{k=0}^{\infty}(-1)^k\binom{j}{k}w^k, \qquad (7)$$

where, $|w| < 1$, $k > 0$. Now using the binomial series given in (7), The pdf of $LT-2TLT-2FD$ can be can be written as a mixture model as follows:

$[1-(e_{-bx^{-\lambda}})_i$ i $i=0$

$${}^2]^{va-1} = \sum^{\infty}\binom{va-1}{}(-1)^i(e^{-bx^{-\lambda}})^2$$

Consequently, the pdf of $LT^{-2TL-T2}_{\infty}FD$ is given as

$$g(x) = 2abv\lambda\sum_{i=0}\binom{va-1}{i}(-1)^i bx^{-\lambda-1}(e^{-bx^{-\lambda}})^{2i+1}, \qquad (8)$$

And can also be written as

Original Article

$$g(x) = 2abv\lambda^{i} b x^{-\lambda-1} e^{-b(i+2)x^{-\lambda}} \sum_{i=0}^{\infty} \binom{va-1}{i} (-1)^{i}, \qquad (9)$$

The expression given in (9) shows that the $LT-2TLT-2FD$ is an infinite mixture representation of the Fréchet distribution.

3.33 Quantile function and random number generation for $LT-2TL-T2FD$

The quantile function of a distribution can be used to investigate the theoretical aspects of the probability distribution, we can employ the use of the quantile function. Mathematically, the quantile function can be expressed in form of $Q(u) = F^{-3}(u)$. Correspondingly, the quantile function of $LT-2TL-T2FD$ is obtained by inverting (1) as follows: $u = 1-[1-(e^{-bx^{-\lambda}})^{2}]^{av}$

By making u the subject of formular, we derive an expression for the quantile function of $LT-2TL-T2FD$ as

$$x_{u} = \left(-\tfrac{1}{b}\left[\log\left(1-(1-u)^{1/av}\right)^{1/2}\right]\right)^{-1/\lambda}, \qquad (11)$$

An expression given in (3.19) can be used for random number generation to validate the method of maximum likelihood used to obtain the value parameters of the distribution. The lower quartile (q_1), middle quartile (q_2) and the upper quartile (q_3)of the $LT-2TL-T2FD$ can be obtained by taking the values of $u = 0.25, 0.5$, and 0.75 respectively in (3.19) as

3.3 Moments of $LT-2TL-T2FD$

Moments are very properties for any statistical investigation, most especially in many application areas. Suppose $X \sim LT-2TL-T2FD\ (a, b, \lambda, v)$,then many important features such as dispersion, skewness, measures of central tendency, and kurtosis of the $LT-2TL-T2FD$ model can be derived by using ordinary moments. The r^{th} raw moment of the $LT-2TL-T2FD$ model is obtained as

$$q_{1} = x_{0.25} = \left(-\tfrac{1}{b}\left[\log\left(1-(0.75)^{1/av}\right)^{1/2}\right]\right)^{-1/\lambda}, \qquad (12)$$

$$q_{2} = x_{0.5} = \left(-\tfrac{1}{b}\left[\log\left(1-(0.5)^{1/av}\right)^{1/2}\right]\right)^{-1/\lambda}, \qquad (13)$$

And

$$q_{3} = x_{0.75} = \left(-\tfrac{1}{b}\left[\log\left(1-(0.25)^{1/av}\right)^{1/2}\right]\right)^{-1/\lambda}. \qquad (14)$$

Original Article

$$E(X)^r = \mu_r' = \int_{-\infty}^{\infty} x^r g(x)dx, \quad (15)$$

Inserting (9) in (15), we obtain

$$\mu_{r'} = 2abv\lambda \sum_{i=0}^{\infty} \binom{va-1}{i} (-1)^i \int_{-\infty}^{\infty} x^{r-(\lambda+1)} e^{-b(i+1)x^{-\lambda}} dx, \quad (16)$$

Letting $y = b(i+1)x^{-\lambda},\ x = [b(1+i)]^{1/\lambda} y^{-1/\lambda},\ dx = -\lambda^{-1}[b(1+i)]^{1/\lambda} y^{-1/\lambda-1} dy$, putting in (16), we have

$$\mu_r' = 2av\lambda \sum_{i=0}^{\infty} \binom{va-1}{i} (-1)^i [2(i+a)]^{\frac{r-1}{b}} \int_0^{\infty} x^{-r/b} e^{-y} dy, \quad (17)$$

Finally, we have,

$$\mu_r' = 2av\lambda \sum_{i=0}^{\infty} \binom{va-1}{i} (-1)^i [2(i+a)]^{\frac{r-1}{b}} \Gamma(1 - {}^r/b), \quad r < b. \quad (18)$$

Where $\Gamma(1 - {}^r/b)$ is an incomplete gamma function. Expression for the mean ($\mu_1' = \mu$) and the variance ($\mu_2 = \mu_2' - \mu_1'^2$) is obtained by taking $r = 1$ and 2, and is given as

$$\mu = 2av\lambda \sum_{i=0}^{\infty} \binom{va-1}{i} (-1)^i \Gamma\left(1 - {}_1/b\right). \quad (19)$$

and

$$\mu_2 = 2av\lambda \sum_{i=0}^{\infty} \binom{va-1}{i} (-1)^i [2(i+a)]^{\frac{1}{b}} \Gamma\left(1 - {}_2/b\right) - \left(2av\lambda \sum_{i=0}^{\infty} \binom{va-1}{i} (-1)^i \Gamma\left(1 - {}^1/_b\right)\right)^2. \quad (20)$$

Further, one can determine the r^{th} central moment and r^{th} cumulant of X defined respectively by,

$$\mu_r = E\{(X-\mu)^r\} = \sum_{j=0}^{r} \binom{r}{j} \mu'_{r-j}(-1)^j \mu_j, \quad \kappa_r = \mu_{r'} - \sum_{j=1}^{r-1} \binom{r-1}{j-1} \kappa_j \mu_{r'-j},$$

With $\kappa_1 = \mu$. One can express several measures of skewness and kurtosis based cumulants (central moments) Consequently, an expression for the variance, skewness and the kurtosis can respectively, be obtained as follows $\sigma^2 = \mu_2' - [\mu_1']^2, s_k = \mu_3^2(\sqrt{\mu_2})^{-3}$ and $k_u = \mu_4(\mu_2)^{-2}$ respectively. Where

Original Article

$\mu_r = E[(x - \mu_1')_r]$, $\mu_3 = -3\mu_2'\mu_1' + \mu_3' + 2(\mu_1')_3$ and $\mu_4 = 6(\mu_1')_2\mu_2' - 3(\mu_1')_4 - 4\mu_3'\mu_1' + \mu_4'$

3.4 Moment generating function of $LT - 2TLT - 2FD$

The moment generating function (MGF) of a random variable X sometimes gives an alternative method that can be used in describing the characteristics of a distribution. Mathematically, the MGF is defined as

$$\mathcal{M}_X(t) = E(e^{tX}) = E(X^r). \sum_{r=0}^{\infty} \frac{t^r}{r!} \quad (21)$$

Putting (18) in (21) for $E(X^r)$ for $LT - 2TLT - 2FD$, we obtain

$$\mathcal{M}_X(t) = 2av\lambda \sum_{i=r=0}^{\infty} \frac{t^r}{r!}\binom{va-1}{i}(-1)^i[2(i+a)]^{\frac{r-1}{b}}\, \Gamma(1 - {}^r/_b). \quad (22)$$

3.5 Entropies $LT - 2TLT - 2FD$

The Rényi entropy of a random variable X with density function $f(x)$ can be described as a measure of variation off uncertainty or randomness and its defined (for $\zeta > 0$ and $\zeta \neq 1$) as;

$$I_R(\zeta) = \frac{1}{1-\zeta} log[Z(\zeta)], \quad (23)$$

where

$$Z(\zeta) = \int_{-\infty}^{\infty} g^{\zeta}(x)\, dx \quad (24)$$

Inserting (2) in (24), we have

$$Z^{\zeta}(\zeta) = \int_{-\infty}^{\infty} [2abv\lambda b x^{-\lambda-1}(e^{-bx^{-\lambda}})^2[1 - (e^{-bx^{-\lambda}})^2]^{va-1}]\, dx \quad (25)$$

Upon simplification, we obtain

$$Z(\zeta) = n_j\Gamma\left(1 - \frac{(\lambda+1)(1+\zeta)}{\lambda}\right), \quad (26)$$

where

$$n_j = 2^{\zeta}a^{\zeta}b^{\zeta}\lambda^{\zeta-1}v^{\zeta} \sum_{j}^{\infty} \binom{\zeta(av-1)}{j}[2(\zeta+j)b]^{\frac{1-\zeta(\lambda+1)}{\lambda}}(-1)^j$$

Putting (26) in (23), we generate an expression for the Rényi entropy of $LT - 2TL - T2 - FD$ as

$$I_R(\zeta) = \frac{1}{1-\zeta} log\left[n_j\Gamma\left(1 - \frac{(\lambda+1)(1+\zeta)}{\lambda}\right)\right]. \quad (27)$$

Original Article

3.40 Order statistics

Given $x_1, x_2, x_3 \ldots, x_n$ as a random sample having CDF $F(x)$. Let $X_{1:1}, X_{2:n}, X_{3:n}, \ldots, x_{n:n}$ is the ordered sample of size n, then the density of j^{th} order statistics is given as

$$g_{j:n}(x) = W^* \sum_{i=0}^{n-r} (-1)^i \binom{n-r}{i} f(x)F(x)^{i+j-1}, \quad (28)$$

Where $W^* = \frac{n!}{(n-r)!r!}$.

Putting (1) and (2) in (28), followed by simply algebraic manipulation gives

$$g_{j:n}(x) = 2\alpha\rho\varsigma^{1/\lambda}W^* \sum_{i=k=l=m=0}^{n-r} (-1)^{i+k+m} \binom{n-r}{i} \binom{il+j-1}{k} \binom{\frac{k}{\lambda}+\frac{1}{\lambda}+}{l} \binom{}{m} (\lambda l + k)\, \varsigma^{k/\lambda}$$

$$\times \bar{\varsigma}^l x^{-(\rho+1)} e^{-(m+1)\alpha x^{-\rho}}.$$

4.0 Real Data Applications for $LT-2TLT-F$ Model

To demonstrate the flexibility proposed family of distributions, -2*log-likelihood statistic $(-2l)$, Akaike information criterion $(AIC = 2p - 2l)$, Consistent Akaike information criterion $\left(CAIC = AIC + 2\frac{p(p+1)}{n-p-1}\right)$ and Hannan–Quinn information criterion (HQIC) are calculated for $LT-ETLT-2F$ model and its sub-models, where n is the number of observations, and p is the number of estimated parameters. The goodness-of-fit statistic, Kolmogorov Smirnoff (K), Cramer–von Mises (CV), and the probability value are also presented in the Table. The best model corresponds among the class considered is the model having minimum value AIC, HQIC, CAIC, K, and CV and the largest probability value as the best model. In this study, numerical results (of maximum likelihood estimates and goodness of fit criteria) are calculated by using the goodness.fit (.) command in the Model Adequacy package available in R language. The AIC, CAIC, HQIC, CV, K and P are given for the sub-models Lehman Type-2 Top-Leone Type-2 Inverse Exponential (LT-2TLT-2IE), Lehman Type-2 Top-Leone Type-2 Inverted Weibull (LT-2TLT-2IW), Type -2 Top-Leone Frechet (T-2TLF), Lehman Type-F (LT-2F), and Frechet distribution. Two data applications are used to show how good the developed is in modeling lifetime data.

The data represent the remission times (in months) of a random sample of 128 bladder cancer patients. For previous study see Lee and Wang (2003). The Exploratory data analysis for the cancer data I of data is given in Table 4.1, the Total Time on Test (TTT) plot is given in Figure 5, Tables 1, 2 and 3 gives the exploratory data analysis for the cancer data, parameter estimates of the model and the model's measures of goodness of fit respectively.

Original Article

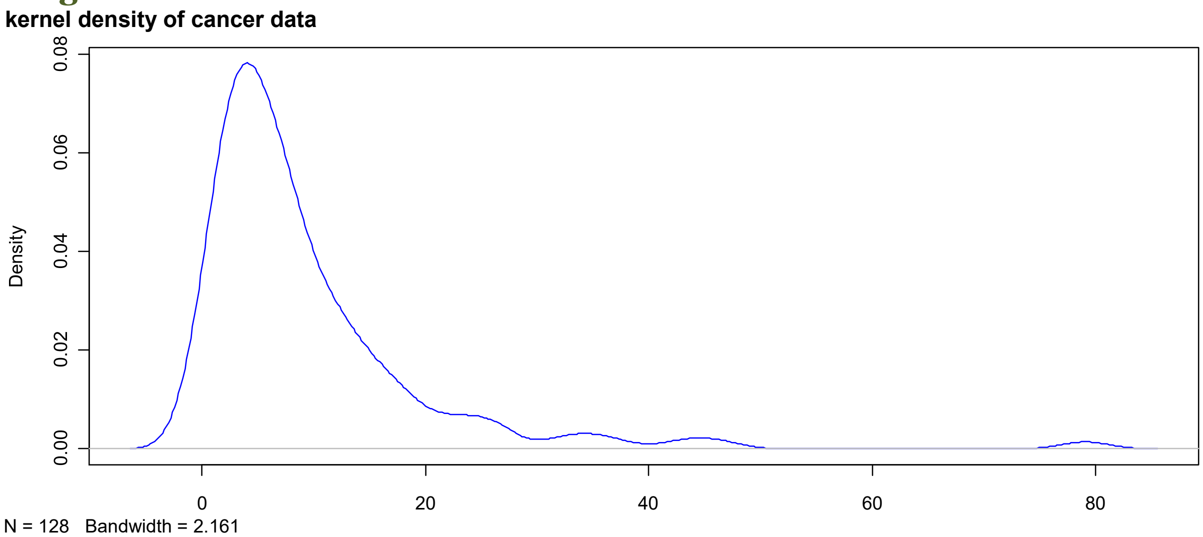

Figure 4.1: Kernel density plot for cancer remission time data

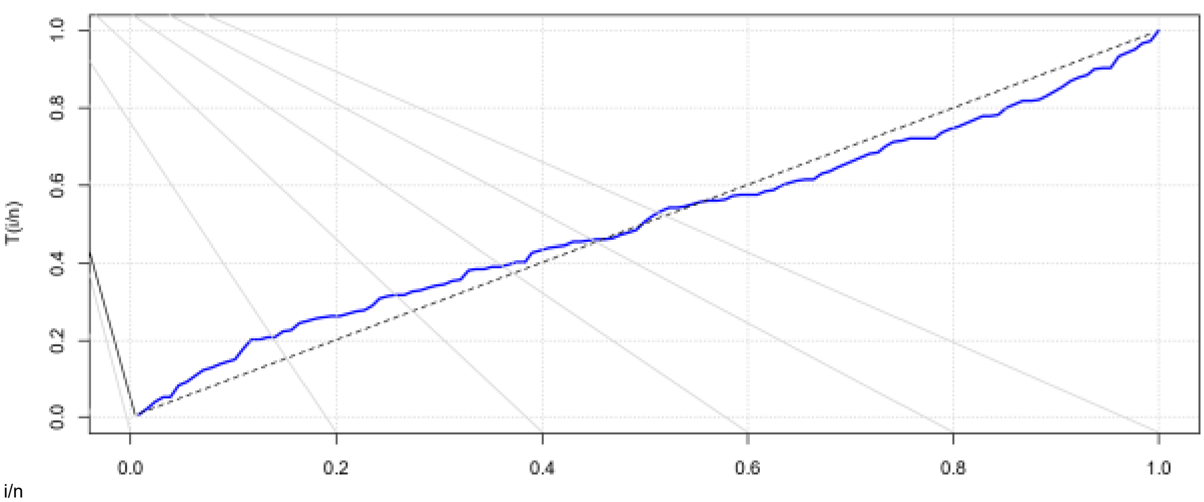

Figure 4.2 TTT plot for cancer remission time data

Table 1 Exploratory data analysis of cancer data

n	q_1	mean	q_3	Range	median	variance	Skewness	kurtosis
128	3.348	9.366	11.838	78.97	6.395	110.425	3.287	18.483

Original Article

Table 2 MLEs of the parameter, Standard error (in parenthesis) of the LT-2TLT-2FD for the cancer remission time data

$Model$	λ	b	a	v
$LT-2TLT-2F$	0.1615 (0.0594)	4.6067 (1.1834)	2.5024 (0.4559)	5.5024 (1.4512)
$LT-2TLT-2IE$	1.0005 (0.1356)	− (−)	0.3539 (1.0987)	2.112 (0.5579)
$LT-2TLT-2IW$	− (−)	0.6804 (0.0633)	0.6534 (1.1281)	1.7819 (0.0765)
$TL-F$	0.3873 (0.0587)	2.5092 (0.2876)	− (−)	8.9931 (4.4713)
$LT-2-F$	0.3553 (0.0410)	5.4365 (0.4676)	12.4994 (4.7131)	− (−)
F	0.7531 (0.0425)	2.4257 (0.2187)	− (−)	− (−)

Table 3. Measures of goodness of fit for the cancer remission time data

$Model$	$-l$	AIC	$CAIC$	$HQIC$	K	CV	PV
$LT-2TLT-2F$	411.14	830.23	830.56	843.87	0.0506	0.0527	0.8985
$LT-2TLT-2IE$	457.20	920.41	920.60	923.88	0.2062	1.1621	3.7e-5
$LT-2TLT-2IW$	445.55	897.09	897.28	900.57	0.1543	0.6211	0.0045
$TL-F$	417.23	840.47	840.66	843.94	0.0836	0.1601	0.3325
$LT-2-F$	415.66	837.33	837.52	840.80	0.0642	0.1294	0.6666
F	444.00	892.00	892.10	894.32	0.1399	0.7451	0.0134

5.0 Conclusion

We have proposed and developed the Type II Lehman Topp-Leone Type II Fréchet distribution along with its properties such as: descriptive measures based on the quantiles, moments, moment generating function, reliability model, Renyi entropy and order statistics. Maximum Likelihood estimates are computed. Goodness of fit shows

Original Article

that Type II Lehman Topp-Leone Type II Fréchet distribution is a better fit. Applications of the Type II Lehman Topp-Leone Type II Fréchet model to cancer remission time data, demonstrate its applicability.

References

Nadarajah, S., and Kotz, S. (2003). The exponentiated Fréchet distribution. Statistics on the Internet, 2003, http://interstat.Statjourna ls.Net/year/2003/articles/0312002.Pdf.

Nadarajah, S., and Gupta, A. K. (2004). The beta Fréchet distribution, Far East Journal of Theoretical Statistics,14:15-24.

Abd-Elfattah, A. M. and Omima A. M. (2009). Estimation of Unknown Parameters of Generalized Fréchet distribution, Journal of Applied Sciences Research,5(10):1398-1408.

Abd-Elfattah A. M., Fergany, H. A., Omima A. M. (2010). Goodness-Of-Fit Test for the Generalized Fréchet distribution, Australian Journal of Basic and Science,4(2):286-301.

Badr M. M. (2010). Studying the Exponentiated Fréchet Distribution, Ph.D. thesis, King AbdulAziz University, Saudi Arabia.

Abd-Elfattah A.M., Assar, S. M.and Abd Elghaffar, H. I. (2016). Exponentiated Generalized Fréchet Distribution, International Journal of Mathematical Analysis and Applications,3(5):39-48.

Lee, E. T., and Wang, J. W. (2003). Statistical methods for survival data analysis (3rd ed.), New York: Wiley. https://doi.org/10.1002/0471458546.