MedCrave Step into the World of Research
Biometrics & Biostatistics International Journal

Research Article

Open Access CrossMark

An extended Suja distribution with statistical properties and applications

Abstract

An extended Suja distribution, of which one parameter Suja distribution is a particular case, has been proposed. Important statistical properties of the proposed distribution based on moments, skewness, kurtosis, index of dispersion, hazard rate function, mean residual life function, stochastic ordering, mean deviations, Renyi entropy measures, and stress-strength reliability have been derived and studied. The method of moments and the method of maximum likelihood for estimating parameters have been discussed. A simulation study has been presented to know the performance of maximum likelihood estimates. Applications and goodness of fit of the proposed distribution with two real datasets have been presented.

Keywords: Suja distribution, statistical properties, parameters estimation, Goodness of fits

Volume 13 Issue 1 - 2024

Rama Shanker,[1] Ronodeep Das,[1] Kamlesh Kumar Shukla[2]

[1]Department of Statistics, Assam University, Silchar, Assam, India
[2]Department of Statistics, Jaypee Institute of Information Technology, Noida, India

Correspondence: Rama Shanker, Department of Statistics, Assam University, Silchar, India, Email shankerrama2009@gmail.com

Received: January 25, 2024 | **Published:** March 05, 2024

Introduction

The search for an appropriate statistical distribution for modeling of lifetime data is very challenging because the lifetime data are stochastic in nature. Statistical distributions are needed for modeling of lifetime data in engineering, medical science, demography, social sciences, physical sciences, finance, insurance, demography, social sciences, literature etc and during recent decades several researchers in statistics and mathematics tried to introduce lifetime distributions. In the exploration for a new lifetime distribution which can be useful to model lifetime data, Shanker[1] proposed a one parameter distribution named Suja distribution defined by its probability density function (pdf) and cumulative distribution function (cdf)

$$f(x;\theta)=\frac{\theta^5}{\theta^4+24}\left(1+x^4\right)e^{-\theta x};x>0,\theta>0 \qquad (1.1)$$

$$F(x;\theta)=1-\left[1+\frac{\theta^4x^4+4\theta^3x^3+12\theta^2x^2+24\theta x}{\theta^4+24}\right]e^{-\theta x};x>0,\theta>0 \qquad (1.2)$$

Length- biased Suja distribution, power length-biased Suja distribution and weighted Suja distribution have been proposed and studied by Al-Omari and Alsmairan,[2] Al-Omari et al.[3] and Alsmairan et al.[4] respectively. Todoka et al.[5] have studied on the cdf of various modifications of Suja distributions and discussed their applications in the field of the analysis of computer- viruses' propagation and debugging theory.

The main purpose of proposing an extended Suja distribution is to see the impact of additional parameter in the distribution over one parameter and other two-parameter distributions. Various descriptive measures, reliability properties and estimation parameters using both the method of moments and the method of maximum likelihood have been discussed. The applications and the goodness of fit of the distribution with two real lifetime datasets have been presented.

An extended Suja distribution

Taking the convex combination of exponential(θ)distribution and gamma$(5,\theta)$distribution with mixing proportion $p=\frac{\alpha\theta^4}{\alpha\theta^4+24}$, the pdf of extended Suja distribution can be expressed as:

$$f(x;\theta,\alpha)=\frac{\theta^5}{\alpha\theta^4+24}\left(\alpha+x^4\right)e^{-\theta x};x>0,\theta>0,\alpha>0 \qquad (2.1)$$

We would call this a two-parameter Suja distribution (TPSD). The corresponding cdf and survival function of TPSD are thus obtained as:

$$F(x;\theta,\alpha)=1-\left[1+\frac{\theta^4x^4+4\theta^3x^3+12\theta^2x^2+24\theta x}{\alpha\theta^4+24}\right]e^{-\theta x};x>0,\theta>0,\alpha>0 \qquad (2.2)$$

$$S(x;\theta,\alpha)=\left[\frac{\theta^4x^4+4\theta^3x^3+12\theta^2x^2+24\theta x+\left(\alpha\theta^4+24\right)}{\alpha\theta^4+24}\right]e^{-\theta x};x>0,\theta>0,\alpha>0.$$

At $\alpha=1$, TPSD reduces to Suja distribution. Also, for $\alpha\to\infty$, TPSD reduces to exponential distribution. The pdf and the cdf of TPSD for varying values of parameters are shown in the Figures 1 & 2 respectively.

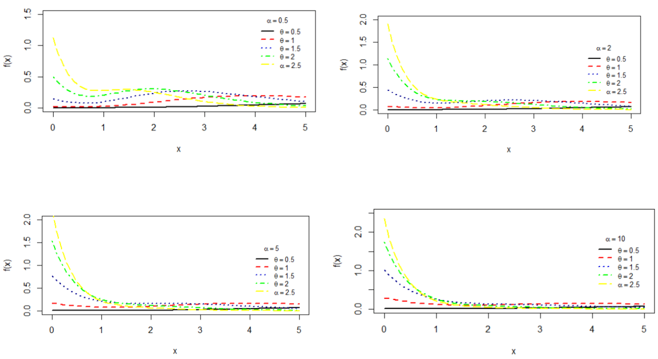

Figure 1 pdf of TPSD.

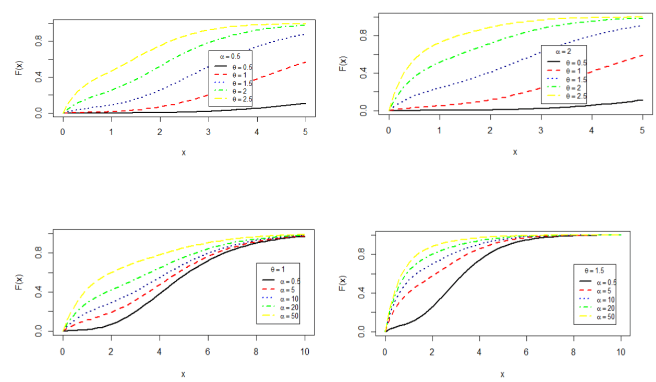

Figure 2 cdf of TPSD.

Measures based on moments

The r th moment about origin (raw moment) μ_r' of TPSD can be obtained as

$$\mu_r' = \frac{r!\left\{\alpha\theta^4 + (r+1)(r+2)(r+3)(r+4)\right\}}{\theta^r\left(\alpha\theta^4+24\right)}; r = 1,2,3,...$$

Thus first four raw moments of TPSD can be expressed as

$$\mu_1' = \frac{\alpha\theta^4+120}{\theta\left(\alpha\theta^4+24\right)},\ \mu_2' = \frac{2\left(\alpha\theta^4+360\right)}{\theta^2\left(\alpha\theta^4+24\right)},$$

$$\mu_3' = \frac{6\left(\alpha\theta^4+840\right)}{\theta^3\left(\alpha\theta^4+24\right)} \text{ and } \mu_4' = \frac{24\left(\alpha\theta^4+1680\right)}{\theta^4\left(\alpha\theta^4+24\right)}.$$

The central moments of TPSD are thus obtained as

$$\mu_2 = \frac{\alpha^2\theta^8+528\alpha\theta^4+2880}{\theta^2\left(\alpha\theta^4+24\right)^2}$$

$$\mu_3 = \frac{2\left(\alpha^3\theta^{12}+1512\alpha^2\theta^8+1728\alpha\theta^4+69120\right)}{\theta^3\left(\alpha\theta^4+24\right)^3}$$

$$\mu_4 = \frac{9\left(\alpha^4\theta^{16}+2656\alpha^3\theta^{12}+58752\alpha^2\theta^8+1234944\alpha\theta^4+3870720\right)}{\theta^4\left(\alpha\theta^4+24\right)^4}$$

The descriptive measures based on moments of TPSD such as coefficient of variation (C.V), coefficient of skewness($\sqrt{\beta_1}$), coefficient of kurtosis (β_2) and index of dispersion (γ) are obtained as

$$C.V. = \frac{\sqrt{\mu_2}}{\mu_1'} = \frac{\sqrt{\alpha^2\theta^8+528\alpha\theta^4+2880}}{\alpha\theta^4+120}$$

$$\sqrt{\beta_1} = \frac{\mu_3}{(\mu_2)^{3/2}} = \frac{2\left(\alpha^3\theta^{12}+1512\alpha^2\theta^8+1728\alpha\theta^4+69120\right)}{\left(\alpha^2\theta^8+528\alpha\theta^4+2880\right)^{3/2}}$$

$$\beta_2 = \frac{\mu_4}{\mu_2^{\,2}} = \frac{9\left(\alpha^4\theta^{16}+2656\alpha^3\theta^{12}+58752\alpha^2\theta^8+1234944\alpha\theta^4+3870720\right)}{\left(\alpha^2\theta^8+528\alpha\theta^4+2880\right)^2}$$

$$\gamma = \frac{\mu_2}{\mu_1'} = \frac{\alpha^2\theta^8+528\alpha\theta^4+2880}{\theta\left(\alpha\theta^4+24\right)\left(\alpha\theta^4+120\right)}$$

The behaviors of these descriptive measures are shown in the Figures 3-6 respectively.

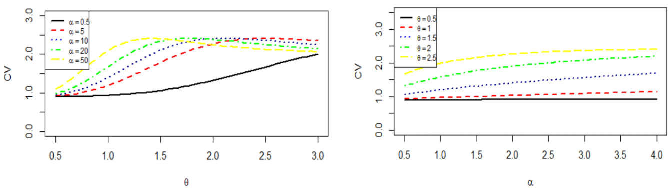

Figure 3 Coefficient of variation of TPSD.

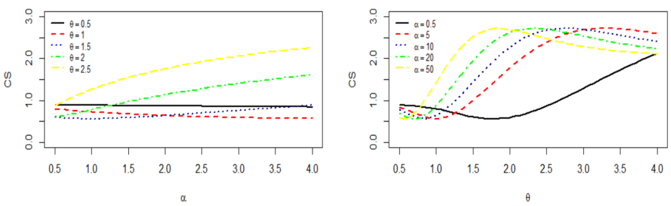

Figure 4 Coefficient of skewness of TPSD.

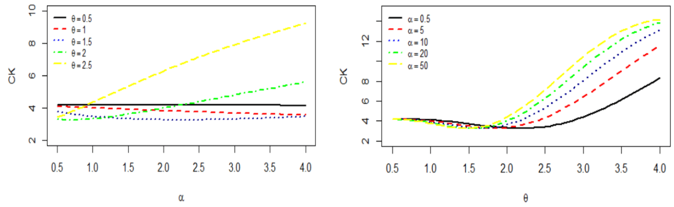

Figure 5 Coefficient of kurtosis of TPSD.

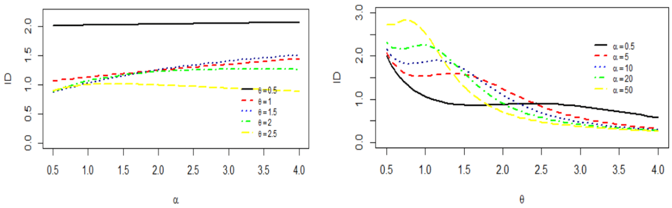

Figure 6 Index of dispersion of TPSD.

Reliability measures

The hazard rate function $h(x)$ and the mean residual life function $m(x)$ of a random variable X having pdf $f(x)$ and cdf $F(x)$ are defined as:

$$h(x) = \lim_{\Delta x\to 0}\frac{P\left(x<X<x+\Delta x\,|\,X>x\right)}{\Delta x} = \frac{f(x)}{1-F(x)} \text{ and}$$

$$m(x) = E\left[X-x\,|\,X>x\right] = \frac{1}{1-F(x)}\int_x^{\infty}\left[1-F(t)\right]dt = \frac{1}{S(x)}\int_x^{\infty} tf(t)dt - x.$$

Thus $h(x)$ and $m(x)$ of the TPSD are obtained as:

$$h(x) = \frac{\theta^4\left(\alpha+x^4\right)}{\theta^4x^4+4\theta^3x^3+12\theta^2x^2+24\theta x+\left(\alpha\theta^4+24\right)}$$

$$\text{and } m(x) = \frac{\theta^4x^4+8\theta^3x^3+36\theta^2x^2+96\theta x+\left(\alpha\theta^4+120\right)}{\theta\left[\theta^4x^4+4\theta^3x^3+12\theta^2x^2+24\theta x+\left(\alpha\theta^4+24\right)\right]}.$$

The $h(x)$ and $m(x)$ of TPSD are shown in Figures 7 & 8 respectively.

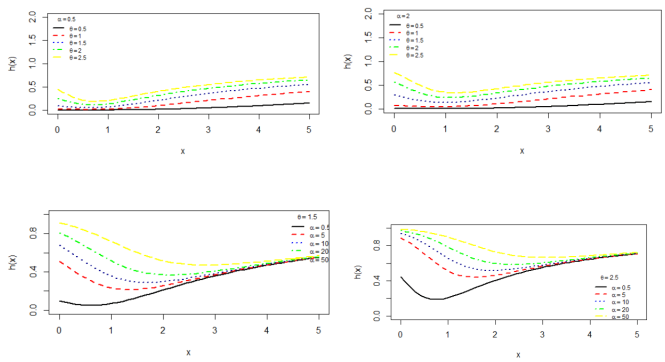

Figure 7 Hazard rate function of TPSD.

Mean deviations

The mean deviation about the mean and the mean deviation about the median are defined as

$$\delta_1(X) = \int_0^{\infty}|x-\mu|f(x)dx = 2\mu F(\mu) - 2\int_0^{\mu} xf(x)dx$$

$$\text{and } \delta_2(X) = \int_0^{\infty}|x-M|f(x)dx = \mu - 2\int_0^{M} xf(x)dx, \text{ respectively,}$$

Citation: Shanker R, Das R, Shukla KK. An extended Suja distribution with statistical properties and applications. *Biom Biostat Int J.* 2024;13(1):16–21. DOI: 10.15406/bbij.2024.13.00409

where $\mu = E(X)$ and $M =$ Median (X).

We have

$$\int_0^{\mu} x f(x;\theta,\alpha)dx = \mu - \frac{\left\{\theta^5\mu^5 + 5\theta^4\mu^4 + 20\theta^3\mu^3 + 60\theta^2\mu^2 + \left(\alpha\theta^4+120\right)\theta\mu + \left(\alpha\theta^4+120\right)\right\}e^{-\theta\mu}}{\theta\left(\alpha\theta^4+24\right)}$$

$$\int_0^{M} x f(x;\theta,\alpha)dx = \mu - \frac{\left\{\theta^5M^5 + 5\theta^4M^4 + 20\theta^3M^3 + 60\theta^2M^2 + \left(\alpha\theta^4+120\right)\theta M + \left(\alpha\theta^4+120\right)\right\}e^{-\theta M}}{\theta\left(\alpha\theta^4+24\right)}$$

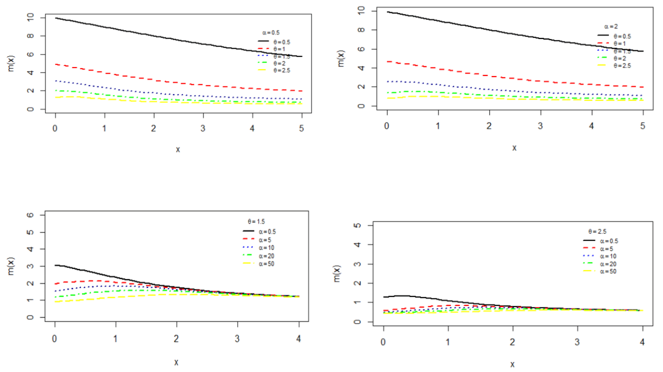

Figure 8 Mean residual life function of TPSD.

Using above expressions and after little simplifications, the mean deviation about mean, $\delta_1(X)$ and the mean deviation about median, $\delta_2(X)$ of TPSD are obtained as:

$$\delta_1(X) = \frac{2\left\{\theta^4\mu^4 + 8\theta^3\mu^3 + 36\theta^2\mu^2 + 96\theta\mu + \left(\alpha\theta^4+120\right)\right\}e^{-\theta\mu}}{\theta\left(\alpha\theta^4+24\right)}$$

$$\delta_2(X) = \frac{2\left\{\theta^5M^5 + 5\theta^4M^4 + 20\theta^3M^3 + 60\theta^2M^2 + \left(\alpha\theta^4+120\right)\theta M + \left(\alpha\theta^4+120\right)\right\}e^{-\theta M}}{\theta\left(\alpha\theta^4+24\right)} - \mu .$$

Order statistics

Let $X_{(1)} < X_{(2)} < ... < X_{(n)}$ denote the order statistics corresponding to random sample $(X_1, X_2, ..., X_n)$. The pdf and the cdf of the k th order statistic, say $Y = X_{(k)}$ are given by

$$f_Y(y) = \frac{n!}{(k-1)!(n-k)!}F^{k-1}(y)\left\{1-F(y)\right\}^{n-k} f(y)$$

$$= \frac{n!}{(k-1)!(n-k)!}\sum_{l=0}^{n-k}\binom{n-k}{l}(-1)^l F^{k+l-1}(y) f(y)$$

and

$$F_Y(y) = \sum_{j=k}^{n}\binom{n}{j}F^j(y)\left\{1-F(y)\right\}^{n-j} = \sum_{j=k}^{n}\sum_{l=0}^{n-j}\binom{n}{j}\binom{n-j}{l}(-1)^l F^{j+l}(y).$$

The pdf and the cdf of the k th order statistics of TPSD are thus obtained as:

and

$$F_Y(y) = \sum_{j=k}^{n}\sum_{l=0}^{n-j}\binom{n}{j}\binom{n-j}{l}(-1)^l\left[1-\left\{\frac{\theta^4x^4 + 4\theta^3x^3 + 12\theta^2x^2 + 24\theta x + \left(\alpha\theta^4+24\right)}{\alpha\theta^4+24}\right\}e^{-\theta x}\right]^{j+l}$$

Stochastic orderings

Stochastic ordering of positive continuous random variables is an important tool for judging their comparative behavior. A random variable Y is said to be greater than a random variable X in the

i. Stochastic order $(X \le_{st} Y)$ if $F_X(x) \ge F_Y(x)$ for all x

ii. Hazard rate order $(X \le_{hr} Y)$ if $h_X(x) \ge h_Y(x)$ for all x

iii. Mean residual life order $(X \le_{mrl} Y)$ if $m_X(x) \le m_Y(x)$ for all x

iv. Likelihood ratio order $(X \le_{lr} Y)$ if $\frac{f_X(x)}{f_Y(x)}$ decreases in x.

The well-known results due to Shaked and Shanthikumar[6] for establishing stochastic ordering of distributions is

$$X \le_{lr} Y \Rightarrow X \le_{hr} Y \Rightarrow X \le_{mrl} Y$$
$$\Downarrow$$
$$X \le_{st} Y$$

Using above results, we have shown in the following theorem that TPSD is ordered with respect to the strongest 'likelihood ratio' ordering.

Theorem: Let $X \sim$ TPSD(θ_1,α_1) and $Y \sim$ TPSD(θ_2,α_1). If $\alpha_1 > \alpha_2$ and $\theta_1 = \theta_2$, or

$\alpha_1 = \alpha_2$ and $\theta_1 = \theta_2$ then $X \le_{lr} Y$ and hence $X \le_{hr} Y$, $X \le_{mrl} Y$ and $X \le_{st} Y$.

Proof: We have $\frac{f_X(x)}{f_Y(x)} = \frac{\theta_1^5\left(\alpha_2\theta_2^4+24\right)}{\theta_2^5\left(\alpha_1\theta_1^4+24\right)}\left(\frac{\alpha_1+x^4}{\alpha_2+x^4}\right)e^{-(\theta_1-\theta_2)x}; x>0$

Now $\ln\frac{f_X(x)}{f_Y(x)} = \ln\frac{\theta_1^5\left(\alpha_2\theta_2^4+24\right)}{\theta_2^5\left(\alpha_1\theta_1^4+24\right)} + \ln\left(\frac{\alpha_1+x^4}{\alpha_2+x^4}\right) - (\theta_1-\theta_2)x$.

This gives $\frac{d}{dx}\ln\frac{f_X(x)}{f_Y(x)} = \frac{4(\alpha_2-\alpha_1)x^3}{\left(\alpha_1+x^4\right)\left(\alpha_2+x^4\right)} - (\theta_1-\theta_2)$

Thus, for $\alpha_1 > \alpha_2$ and $\theta_1 = \theta_2$, or $\alpha_1 = \alpha_2$ and $\theta_1 > \theta_2$, $\frac{d}{dx}\ln\frac{f_X(x)}{f_Y(x)} < 0$. This means that $X \le_{lr} Y$ and hence $X \le_{hr} Y$, $X \le_{mrl} Y$ and $X \le_{st} Y$.

Renyi entropy measure

A measure of variation of uncertainty of a random variable X is known as Renyi entropy measure and given by Renyi.[7] If X is a continuous random variable having pdf $f(.)$, then Renyi entropy is defined as:

$$T_R(\gamma) = \frac{1}{1-\gamma}\log\left\{\int f^{\gamma}(x)dx\right\}, \text{ where } \gamma > 0 \text{ and } \gamma \ne 1.$$

Thus, the Renyi entropy of TPSD can be obtained as

$$T_R(\gamma) = \frac{1}{1-\gamma}\log\left[\int_0^{\infty}\frac{\theta^{5\gamma}}{\left(\alpha\theta^4+24\right)^{\gamma}}\left(\alpha+x^4\right)^{\gamma}e^{-\theta\gamma x}dx\right]$$

$$= \frac{1}{1-\gamma}\log\left[\sum_{j=0}^{\infty}\binom{\gamma}{j}\frac{\theta^{5\gamma-4j-1}\alpha^{\gamma-j}}{\left(\alpha\theta^4+24\right)^{\gamma}}\frac{\Gamma(4j+1)}{(\gamma)^{4j+1}}\right].$$

Stress-strength reliability

Let X and Y denote the strength and the stress of a component. The stress- strength reliability describes the life of a component whose random strength is subjected to a random stress. When $X < Y$, the component fails instantly and the component will function satisfactorily till $X > Y$. Therefore, $R = P(Y < X)$ is the measure of component reliability and is known as stress-strength parameter. It has wide applications in engineering, biomedical science, social science etc.

Let X and Y are independent strength and stress random variables having TPSD with parameter (θ_1,α_1) and (θ_2,α_2), respectively.

Citation: Shanker R, Das R, Shukla KK. An extended Suja distribution with statistical properties and applications. *Biom Biostat Int J.* 2024;13(1):16‒21. DOI: 10.15406/bbij.2024.13.00409

Then, the stress-strength reliability R of TPSD can be obtained as:

$$R = P(Y < X) = \int_0^\infty P(Y < X \mid X = x) f_X(x)\,dx = \int_0^\infty f(x;\theta_1,\alpha_1)\ F(x;\theta_2,\alpha_2)\,dx$$

$$= 1 - \frac{\theta_1^5\left[\begin{array}{l} 40320\theta_2^4 + 20160\theta_2^3(\theta_1+\theta_2) + 8640\theta_2^2(\theta_1+\theta_2)^2 + 2880\theta_2(\theta_1+\theta_2)^3 \\ +24\left(2\alpha_1\theta_2^4+24\right)(\theta_1+\theta_2)^4 + 24\alpha_1\theta_2^3(\theta_1+\theta_2)^5 + 24\alpha_1\theta_2^2(\theta_1+\theta_2)^6 \\ +24\alpha_1\theta_2(\theta_1+\theta_2)^7 + \alpha_1\left(\alpha_2\theta_2^4+24\right)(\theta_1+\theta_2)^8 \end{array}\right]}{\left(\alpha_1\theta_1^4+24\right)\left(\alpha_2\theta_2^4+24\right)(\theta_1+\theta_2)^9}.$$

Estimation of parameters

Method of moments

Since TPSD has two parameters to be estimated, the first two moments about origin are required to estimate its parameters. We have

$$\frac{\mu_2'}{(\mu_1')^2} = \frac{2\left(\alpha\theta^4+360\right)\left(\alpha\theta^4+24\right)}{\left(\alpha\theta^4+120\right)^2} = k \text{ (Say)}$$

Taking $b = \alpha\theta^4$, above equation becomes

$$\frac{2(b+360)(b+24)}{(b+120)^2} = k$$

$$\frac{2\left(b^2+384b+8670\right)}{b^2+240b+14400} = k$$

$$(k-2)b^2 + (240k-768)b + (14400k-17340) = 0 \qquad (11.1.1)$$

Now, for real root of b, the discriminant of the above equation should be greater than and equal to zero. That is

$$(240k-768)^2 - 4(k-2)(14400k-17340) \ge 0 \Rightarrow k \le 2.45.$$

This means that the method of moments estimate is applicable if $k = \frac{m_2'}{(\bar{x})^2} \le 2.45$, where m_2' is the second moment about origin and $\bar{x}$ is the sample mean of the dataset. Now taking $b = \alpha\theta^4$ in the expression for mean, we get the moment estimate $\tilde{\theta}$ of θ as

$$\frac{\alpha\theta^4+120}{\theta\left(\alpha\theta^4+24\right)} = \frac{b+120}{\theta(b+24)} = \bar{x} \Rightarrow \tilde{\theta} = \frac{b+120}{(b+24)\bar{x}}.$$

Using the moment estimate of θ in $b = \alpha\theta^4$, we get the moment estimate $\tilde{\alpha}$ of α as

$$\tilde{\alpha} = \frac{b}{(\tilde{\theta})^4} = \frac{b(b+124)^4(\bar{x})^4}{(b+120)^4}$$

Thus the method of moment estimates $(\tilde{\theta},\tilde{\alpha})$ of parameters (θ,α) of TPSD are given by

$$(\tilde{\theta},\tilde{\alpha}) = \left(\frac{b+120}{(b+24)\bar{x}}, \frac{b(b+124)^4(\bar{x})^4}{(b+120)^4}\right),$$

where b is the value of the quadratic equation in (11.1.1).

Method of maximum likelihood

Let $(x_1,x_2,x_3,\ldots,x_n)$ be a random sample of size n from TPSD (θ,α). Then the log- likelihood function of TPSD is given by

$$\log L = n\left[5\log\theta - \log\left(\alpha\theta^4+24\right)\right] + \sum_{i=1}^{n}\log\left(\alpha + x_i^4\right) - n\theta\bar{x}.$$

The maximum likelihood estimates $(\hat{\theta},\hat{\alpha})$ of parameters (θ,α) are the solution of the following log-likelihood equations

$$\frac{\partial \log L}{\partial\theta} = \frac{5n}{\theta} + \frac{4n\theta\alpha}{\alpha\theta^4+24} - n\bar{x} = 0$$

$$\frac{\partial \log L}{\partial\alpha} = \frac{-n\theta^4}{\alpha\theta^4+24} + \sum_{i=1}^{n}\frac{1}{\alpha+x_i^4} = 0$$

We have to use Fisher's scoring method for solving these two log-likelihood equations because these two log-likelihood equations cannot be solved directly. We have

$$\frac{\partial^2 \log L}{\partial\theta^2} = \frac{-5n}{\theta^2} + \frac{4n\alpha^2\theta^6 - 288n\alpha\theta^2}{\left(\alpha\theta^4+24\right)^2}$$

$$\frac{\partial^2 \log L}{\partial\alpha^2} = \frac{n\theta^8}{\left(\alpha\theta^4+24\right)^2} - \sum_{i=1}^{n}\frac{1}{\left(\alpha+x_i^4\right)^2}$$

$$\frac{\partial^2 \log L}{\partial\theta\partial\alpha} = \frac{-96n\theta^3}{\left(\alpha\theta^4+24\right)^2} = \frac{\partial^2 \log L}{\partial\alpha\partial\theta}.$$

The following equations can be solved for MLEs $(\hat{\theta},\hat{\alpha})$ of (θ,α) of TPSD

$$\begin{bmatrix} \frac{\partial^2 \ln L}{\partial\theta^2} & \frac{\partial^2 \ln L}{\partial\theta\partial\alpha} \\ \frac{\partial^2 \ln L}{\partial\theta\partial\alpha} & \frac{\partial^2 \ln L}{\partial\alpha^2} \end{bmatrix}_{\substack{\hat{\theta}=\theta_0 \\ \hat{\alpha}=\alpha_0}} \begin{bmatrix} \hat{\theta}-\theta_0 \\ \hat{\alpha}-\alpha_0 \end{bmatrix} = \begin{bmatrix} \frac{\partial \ln L}{\partial\theta} \\ \frac{\partial \ln L}{\partial\alpha} \end{bmatrix}_{\substack{\hat{\theta}=\theta_0 \\ \hat{\alpha}=\alpha_0}}$$

where θ_0 and α_0 are the initial values of θ and α, as given by the method of moments. These equations are solved iteratively till close estimates of parameters are obtained.

A simulation study

A simulation study has been carried out to check the performance of maximum likelihood estimates by taking sample sizes (n = 20,40,60,80) for values of $\theta = 0.5, 1.0, 1.5, 2.0$ and $\alpha = 0.5$ and 4. Acceptance and rejection method is used to generate random number for data simulation using R-software. The process was repeated 1,000 times for the calculation of Average Bias error (ABE) and MSE (Mean square error) of parameters θ and α are presented in Tables 1 &2 respectively. For the TPSD decreasing trend has been observed in ABE and MSE as the sample size increases and this shows that the performance of maximum likelihood estimators is quite good and consistent.

Table 1 ABE and MSE of parameters at fixed value $\alpha = 0.5$

Sample	θ	ABE(θ)	MSE (θ)	ABE (α)	MSE (α)
20	0.5	0.0323	0.02083	0.0645	0.7180
	1.0	0.0073	0.0010	0.1145	0.2621
	1.5	-0.0177	0.0063	0.0645	0.0831
	2.0	-0.0427	0.0365	0.0144	0.0041
40	0.5	0.0168	0.0113	-0.0074	0.1210
	1.0	0.0043	0.0007	0.0175	0.0122
	1.5	-0.0081	0.0026	-0.0074	0.0022
	2.0	-0.0206	0.0170	-0.0324	0.0422
60	0.5	0.0098	0.0058	-0.0011	0.0982
	1.0	0.0015	0.0001	0.0143	0.0154
	1.5	-0.0067	0.0027	-0.0011	0.0008
	2.0	-0.0151	0.0136	-0.0178	0.0191
80	0.5	0.0057	0.0026	0.0292	0.2932
	1.0	-0.0004	0.0001	0.0417	0.1397
	1.5	-0.0067	0.0035	0.0292	0.0686
	2.0	-0.0129	0.01342	0.0167	0.0225

Citation: Shanker R, Das R, Shukla KK. An extended Suja distribution with statistical properties and applications. *Biom Biostat Int J.* 2024;13(1):16‒21. DOI: 10.15406/bbij.2024.13.00409

Table 2 ABE and MSE of parameters at fixed value of $\alpha = 4$

Sample	θ	ABE (θ)	MSE (θ)	ABE (α)	MSE (α)
20	0.5	0.0156	0.0048	0.0387	0.5365
	1.0	-0.0093	0.0017	0.0887	0.1576
	1.5	-0.0343	0.0236	0.0387	0.0300
	2.0	-0.0593	0.0704	-0.0112	0.0025
40	0.5	0.0168	0.0113	-0.0074	0.1210
	1.0	0.0043	0.0007	0.01750	0.0122
	1.5	-0.0081	0.0026	-0.0074	0.0022
	2.0	-0.0206	0.0170	-0.0324	0.0422
60	0.5	0.0100	0.0060	-0.0064	0.0745
	1.0	0.0016	0.0001	0.0102	0.0063
	1.5	-0.0066	0.0026	-0.0064	0.0024
	2.0	-0.0149	0.0134	-0.0230	0.0319
80	0.5	0.0057	0.0026	0.0306	0.3063
	1.0	-0.0004	0.0017	0.0431	0.1488
	1.5	-0.0068	0.0036	0.0306	0.0750
	2.0	-0.0129	0.0134	0.0181	0.0262

Applications

The goodness of fit of TPSD along with its comparison with one parameter Suja distribution and two-parameter lifetime distributions including quasi Lindley distribution (QLD) of Shanker and Mishra,[8] a two-parameter Lindley distribution (TPLD-I b) of Shanker and Mishra,[9] a two-parameter Lindley distribution (TPLD-II) of Shanker et al.[10] for two real lifetime datasets relating to failure times have been discussed. The applications of the TPSD can also be extended to model the survival times of patients suffering from serious disease in medical sciences. The pdf and the cdf of these distributions are presented in the following Table 3.

The two datasets considered for testing the goodness of fit of TPSD over other one parameter and two-parameter lifetime distributions are as follows:

Dataset 1: The positively skewed data relating to the accelerated life testing of item ($n = 55$) with changes in stress from 100 to 150 at time $t = 15$, available in Murthy et al (2004).

0.032, 0.035, 0.104, 0.169, 0.196, 0.260, 0.326, 0.445, 0.449, 0.496, 0.543, 0.544, 0.577, 0.648, 0.666, 0.742, 0.757, 0.808, 0.857, 0.858, 0.882, 1.005, 1.025, 1.472, 1.916, 2.313, 2.457, 2.530, 2.543, 2.617, 2.835, 2.940, 3.002, 3.158, 3.430, 3.459, 3.502, 3.691, 3.861, 3.952, 4.396, 4.744, 5.346, 5.479, 5.716, 5.825, 5.847, 6.084, 6.127, 7.241, 7.560, 8.901, 9.000, 10.482, 11.133.

Dataset 2: The positively skewed failure time data ($n = 40$), available in Murthy *et al* (2004).

0.13, 0.62, 0.75, 0.87, 1.56, 2.28, 3.15, 3.25, 3.55, 4.49, 4.50, 4.61, 4.79, 7.17, 7.31, 7.43, 7.84, 8.49, 8.94, 9.40, 9.61, 9.84, 10.58, 11.18, 11.84, 13.28, 14.47, 14.79, 15.54, 16.90, 17.25, 17.37, 18.69, 18.78, 19.88, 20.06, 20.10, 20.95, 21.72, 23.87.

The corresponding maximum likelihood estimates of parameters along with $-2\log L$, AIC, kolmogorov-Smirnov (K-S) and p-values of the considered datasets for the given distributions are presented in Table 4 & 5, respectively. The fitted plots of the distributions for the considered two datasets have been shown in Figures 9 & 10 respectively. The goodness of fit in Tables 4 & 5 and the fitted plots in Figures 9 & 10 shows that TPSD gives much closer fit for the considered datasets in Table 4 while in Table 5 TPLD-1 gives better fit over other distributions. Therefore, it can be concluded that TPSD and TPLD-1 can be considered the best distributions for lifetime data.

Table 3 pdf and the cdf of two-parameter distributions

Distributions	pdf	Cdf
TPLD-I	$f(x;\theta,\alpha)=\frac{\theta^2}{\theta\alpha+1}(\alpha+x)e^{-\theta x};x>0,\theta>0,\theta\alpha>-1$	$F(x;\theta,\alpha)=1-\left(1+\frac{\theta x}{\alpha\theta+1}\right)e^{-\theta x}$
TPLD-II	$f(x;\theta,\alpha)=\frac{\theta^2}{\theta+\alpha}(1+\alpha x)e^{-\theta x};x>0,\theta>0,\alpha>0$	$F(x;\theta,\alpha)=1-\left(1+\frac{\alpha\theta x}{\theta+\alpha}\right)e^{-\theta x}$
QLD	$f(x;\theta,\alpha)=\frac{\theta}{\alpha+1}(\alpha+\theta x)e^{-\theta x}$	$F(x;\theta,\alpha)=1-\left(1+\frac{\theta x}{\alpha+1}\right)e^{-\theta x}$

Table 4 ML estimates of the parameters of distributions and values of $-2\log L, AIC, K-S, p-value$ for data set 1

Distributions	ML estimates		$-2\log L$	AIC	K-S	p-value
	θ	α				
TPSD	0.9563	32.1684	226.65	230.65	0.086	0.774
QLD	0.3848	5.19455	231.44	235.44	0.135	0.244
TPLD-I	0.3907	11.6595	231.45	235.45	0.136	0.235
TPLD-II	0.383	0.07082	231.44	235.44	0.134	0.246
SD	1.4504	……..	265.86	267.86	0.282	0.0002

DOI: 10.15406/bbij.2024.13.00409

Table 5 ML estimates of the parameters of distributions values of $-2\log L, AIC, K-S, p-value$ for data

Distributions	ML estimates		$-2\log L$	AIC	K-S	p-value
	θ	α				
TPSD	0.4175	158.423	262.10	266.10	0.136	0.406
QLD	0.16453	0.3914	263.24	267.24	0.107	0.708
TPLD-I	0.16456	2.3745	263.25	263.25	0.106	0.711
TPLD-II	0.16453	0.42038	263.24	267.24	0.107	0.709
SD	0.4778	……..	301.17	303.17	0.24	0.015

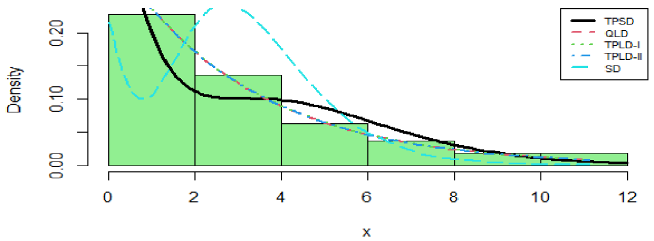

Figure 9 Fitted plots of distributions for data set1.

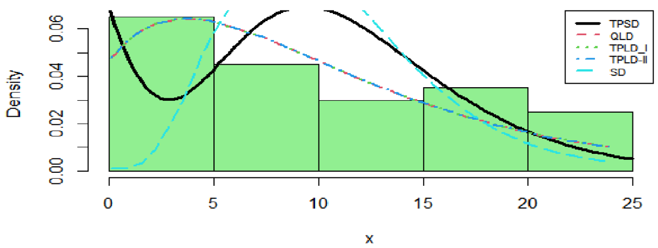

Figure 10 Fitted plots of distributions for datasets 2.

Conclusion

In this paper, a two-parameter Suja distribution has been proposed by introducing an additional parameter in one parameter Suja distribution to see its effect regarding goodness of fit over Suja distribution and other two-parameter lifetime distributions. Its various descriptive measures based on moments and reliability properties have been discussed. The estimation of parameters using method of moments and maximum likelihood method has been discussed. A simulation study has been presented to know the performance of maximum likelihood estimates. The goodness of fit of the proposed distribution has been presented with two real lifetime datasets.

Acknowledgments

Authors are grateful to the editor in chief and the anonymous reviewer for some minor comments which improved both the quality and the presentation.

Conflicts of interest

The authors declare that there are no conflicts of interest.

Funding

None.

References

1. Shanker R. Suja distribution and its application. *International Journal of Probability and Statistics*. 2017;6(2):11–19.
2. Al-Omari AI, Alsmairam I.K. Length-biased Suja distribution and Its application, *Journal of Applied Probability and Statistics*. 2019;14(3):95–116.
3. Al-Omari AI, Alhyasat IK, Abu B.M.A. Power length-biased Suja distribution –properties and application. *Electronic Journal of Applied Statistical Analysisl*. 2019;12(2):429–452.
4. Alsmairan IK, Al-Omari AI. Weighted Suja distribution with application to ball bearings data, *Life Cycle Reliability and Safety Engineering*. 2020;9:195–211.
5. Todorka T, Anton I, Asen R, et al. Comments on some modification of Suja Cumulative Functions with applications to the theory of computer viruses propagation. *International Journal of Differential Equations and Applications*. 2020;19(1):83–95.
6. Shaked M, Shanthikumar JG. Stochastic orders and their applications. Academic Press; New York. 1994.
7. Renyi A. On measures of entropy and information. *Berkeley Symp on Math Statist Prob*. 1961;1:547–561.
8. Shanker R, Mishra A. A quasi Lindley distribution. *African Journal of Mathematics and Computer Science Research.* (2013 a);6(4):64– 71.
9. Shanker R, Mishra A. A two-parameter Lindley distribution**.** *Statistics in Transition New Series*. 2013(b);14(1):45–56.
10. Shanker R, Sharma S, Shanker R. A two -parameter Lindley distribution for modeling waiting and survival times data. *Applied Mathematics*. 2013;4(2):363–368.
11. Murthy DNP, Xie M, Jiang R. *Weibull Models.* John Wiley& sons Inc; Hoboken. 2004.
12. Bonferroni CE. *Elementi di Statistca generale*. Seeber; Firenze. 1930.