Biometrics & Biostatistics International Journal

Review Article

Open Access

An extended Sujatha distribution with statistical properties and applications

Abstract

In this paper, an extended Sujatha distribution has been proposed using the exponentiated technique on Sujatha distribution. Statistical properties including survival function, hazard function, harmonic mean, moment generating function, order statistics and Renyi entropy have been discussed. Moments of the proposed distribution has been obtained. The estimation of parameters using the maximum likelihood method and maximum product spacing method has been explained. The simulation study has been presented to know the performance of maximum likelihood estimates as the sample size increases. Finally, two examples of real lifetime datasets from the engineering field have been presented to demonstrate its applications and the goodness of fit of extended Sujatha distribution shows better fit over exponentiated exponential and exponentiated Aradhana distributions.

Keywords: exponentiated distribution, sujatha distribution, statistical properties, maximum likelihood estimation, applications

Volume 13 Issue 4 - 2024

Hosenur Rahman Prodhani, Rama Shanker

Department of Statistics, Assam University, Silchar, India

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

Received: August 21, 2024 | **Published:** September 12, 2024

Introduction

The exponentiation technique is a new concept of generalizing a given distribution which results into introducing an additional parameter in a distribution. Gupta et al.,[1] proposed exponentiated technique of generalizing a new distribution and proposed exponentiated exponential distribution (EED) using exponentiated technique. This family came with scale and shape parameter, similar to Weibull or Gamma families. Later, some important statistical properties of EED studied by Gupta and Kundu.[2] It has been observed that many characteristics of the new family were identical to those of the Weibull or Gamma families. Therefore, this distribution can be used as an alternative to Gamma or Weibull distributions. Two-parameter gamma and Weibull distributions are the most commonly used distributions for analyzing lifetime data. Gamma distributions have many applications beyond lifetime distributions. However, its major drawback is that its survival function cannot be obtained in a closed form unless the shape parameter is an integer. This makes the Gamma distribution a little less popular than the Weibull distribution, whose survival function and failure rate have very simple and easy-to-study forms. Presently exponentiated distributions and their mathematical properties are extensively studied for applied science experimental datasets. Pal et al.,[3] studied the exponentiated Weibull family as an extension of Weibull distribution. Following the approach of deriving EED, during recent decades several researchers attempted to derive exponentiated version of many distributions. A Generalized Lindley distribution studied by Nadarajah et al.,[4] the exponentiated generalized Lindley distribution (EGLD) studied by Rodrigues et al.,[5] exponentiated Shanker distribution (ESHD) studied by Jayakumar and Elangovan,[6] exponentiated Ishita distribution (EID) studied by Rather and Subramanian,[7] exponentiated Aradhana distribution (EAD) studied by Ganaie and Rajagopalan.[8]

Sujatha distribution has been proposed by Shanker[9] for analyzing lifetime data from engineering and biomedical science. The probability density function (pdf) and the cumulative distribution function (cdf) of Sujatha distribution can be expressed as

$$f\left(x\right)=\frac{\theta^{3}}{\theta^{2}+\theta+2}\left(1+x+x^{2}\right)e^{-\theta x};x>0,\theta>0 \qquad (1)$$

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

where θ is a scale parameter.

Statistical properties, estimation of parameter and applications of Sujatha distribution are studied by Shanker.[9] Later on, several generalizations and modifications of Sujatha distribution have been done by several researchers. For example, quasi Sujatha distribution by Shanker,[10] a generalization of Sujatha distribution by Shanker et al.,[11] a two-parameter Sujatha distribution by Mussie and Shanker,[12] a new two parameter Sujatha distribution by Mussie and Shanker,[13] another new two parameter Sujatha distribution by Shanker,[14] weighted Sujatha distribution by Shanker and Shukla,[15] power Sujatha distribution by Shanker and Shukla,[16] a new quasi Sujatha distribution by Shanker and Shukla,[17] generalized Inverse power Sujatha distribution by Okoli et al.,[18] and Marshal-Olkin Sujatha distribution by Ikechukwu and Eghwerideo[19] are some among others.

The rationale behind the introduction of the extended Sujatha distribution presented in this paper can be elucidated as follows:

Consider a scenario where we have a series system comprised of independent components, and let $X_1, X_2, ..., X_\alpha$ are independent random variables comes from a distribution with cdf $F\left(x\right)$ and represent the failure times of the respective components. It is assumed that these components are independent of each other and identical. In this context, the probability that the entire system will experience failure before a specific time x can be expressed as follows:

$$\Pr\left[\max\left(X_1,X_2,...,X_\alpha\right)\le x\right]=\Pr\left(X_1\le x\right)\Pr\left(X_2\le x\right)...\Pr\left(X_\alpha\le x\right)$$
$$=F(x)F(x)...F(x)$$

$$= \left[F(x)\right]^{\alpha}$$

Therefore, it provides the distribution that characterizes the failure of a series system consisting of independent components.

a. Its pdf has unimodal and positively skewed shape, its hazard function has monotonically increased, monotonically decreased and L- shape.

b. The proposed distribution has closed forms for cdf, survival function and hazard function.

c. It can be considered a good distribution to fit positively skewed data that other popular lifetime distributions might not be able to fit appropriately.

Various statistical properties including hazard function, order statistics and Renyi entropy have been discussed. Estimation of parameters using maximum likelihood method and maximum product spacing method have discussed. The applications and the goodness of fit of the proposed distribution have been illustrated with two examples of observed real datasets from engineering. The proposed distribution gives much closure fit then exponentiated exponential and exponentiated Aradhana distributions.

An extended sujatha distribution

Following exponentiated method is use to propose the new distribution named as extended Sujatha distribution.

If $f(x)$ and $F(x)$ are pdf and cdf respectively of a random variable X, then the new proposed exponentiated family of distribution function has the following cdf and pdf

$$G(x) = \left[F(x)\right]^{\alpha}\; ; x > 0, \alpha > 0 \quad (3)$$

$$g(x) = \alpha\left[F(x)\right]^{\alpha-1} f(x); x > 0, \alpha > 0 \quad (4)$$

Let X is a random variable which follows extended Sujatha distribution (ESD). The cdf of ESD can be obtained as

$$G(x;\theta,\alpha) = \left[1 - \left\{1 + \frac{\theta x(\theta x + \theta + 2)}{\theta^2 + \theta + 2}\right\} e^{-\theta x}\right]^{\alpha}\; ; x > 0, \theta > 0, \alpha > 0 \quad (5)$$

Thus, the pdf of ESD can be expressed as

$$g(x;\theta,\alpha) = \frac{\alpha\theta^3}{\theta^2 + \theta + 2}\left(1 + x + x^2\right) e^{-\theta x}\left[1 - \left\{1 + \frac{\theta x(\theta x + \theta + 2)}{\theta^2 + \theta + 2}\right\} e^{-\theta x}\right]^{\alpha-1}\; ; x > 0, \theta > 0, \alpha > 0 \quad (6)$$

For $\alpha = 1$ the ESD reduced to Sujatha distribution. The nature of the pdf of ESD has been shown in the following figure 1.

It is obvious that for fixed value of the parameter α and for increasing values of parameter θ, the shape of ESD approaches approximately towards leptokurtic and positively skewed. On the other hands for fixed value of the parameter θ and for increasing values of parameter α, the shape of ESD approaches approximately towards positively skewed. The behaviors of the cdf of ESD have been shown in the figure 2.

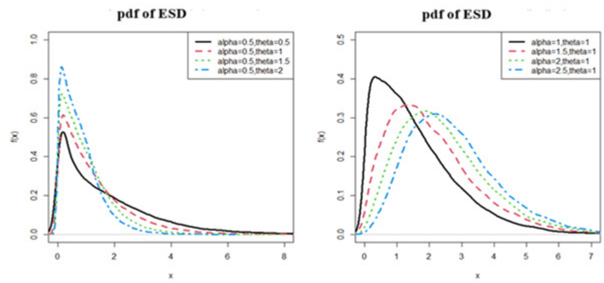

Figure 1 Graphs of the pdf of ESD for selected values of the parameters.

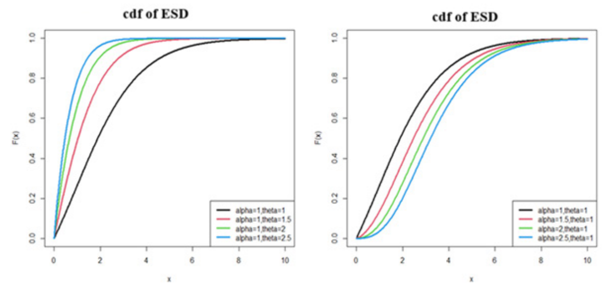

Figure 2 Graphs of the cdf of ESD for selected values of the parameters.

Statistical properties of ESD

In this section, some important statistical properties of ESD have been studied.

Survival function

The Survival function of ESD can be obtained as

$$S(x) = 1 - \left[1 - \left\{1 + \frac{\theta x(\theta x + \theta + 2)}{\theta^2 + \theta + 2}\right\} e^{-\theta x}\right]^{\alpha}\; ; x > 0, \theta > 0, \alpha > 0 \quad (7)$$

Hazard function

The hazard function of ESD can be obtained as

$$h(x) = \frac{\frac{\alpha\theta^3}{\theta^2 + \theta + 2}\left(1 + x + x^2\right) e^{-\theta x}\left[1 - \left[1 + \frac{\theta x(\theta x + \theta + 2)}{\theta^2 + \theta + 2}\right] e^{-\theta x}\right]^{\alpha-1}}{1 - \left[1 - \left[1 + \frac{\theta x(\theta x + \theta + 2)}{\theta^2 + \theta + 2}\right] e^{-\theta x}\right]^{\alpha}}\; ; x > 0, \theta > 0, \alpha > 0 \quad (8)$$

The natures of the hazard function of ESD are shown for varying values of parameters in the figure 3.

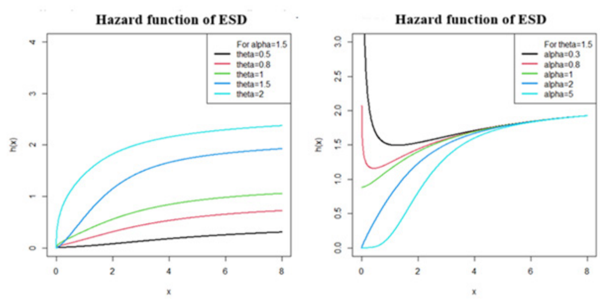

Figure 3 Graphs of hazard function of ESD for selected values of the parameters.

Citation: Prodhani HR, Shanker R. An extended Sujatha distribution with statistical properties and applications. *Biom Biostat Int J.* 2024;13(4):96‒105. DOI: 10.15406/bbij.2024.13.00420

From the figure 3 we observed that when α is fixed and for all values of θ the hazard function is monotonically increasing and when θ is fixed and $\alpha < 1$ then the hazard function is decreasing and after a certain time it is increasing and when θ is fixed and $\alpha < 1$ then hazard function is a increasing.

The linear representation

Using the binomial expansion of $(1-x)^n = \sum_{i=0}^{\infty} (-1)^i \binom{n}{i} x^i$, we can be obtained the pdf of ESD as

$$g(x) = \alpha \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} (-1)^i \binom{\alpha-1}{i} \binom{i}{j} \binom{j}{k} \frac{\theta^{k+j+3} (\theta+2)^{j-k}}{(\theta^2+\theta+2)^{j+1}} x^{j+k} \left(1+x+x^2\right) e^{-\theta(i+1)x} \quad (9)$$

Moments of ESD

Using the linear representation in 9, the r th moments about origin μ_r' of ESD can be obtained as

$$\mu_r' = E\left(X^r\right) = \int_0^{\infty} x^r g(x) dx$$

$$= \alpha \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} (-1)^i \binom{\alpha-1}{i} \binom{i}{j} \binom{j}{k} \frac{\theta^{j+k+3} (\theta+2)^{j-k}}{(\theta^2+\theta+2)^{j+1}} \int_0^{\infty} x^{j+k+r} \left(1+x+x^2\right) e^{-\theta(i+1)x} dx \quad (10)$$

Since equation (10) is a convergent series for all $r \geq 0$, all the moments or ESD exists. The first four moments about origin of ESD can thus be expressed as

$$\mu_1' = \alpha \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} (-1)^i \binom{\alpha-1}{i} \binom{i}{j} \binom{j}{k} \frac{(\theta+2)^{j-k}}{(\theta^2+\theta+2)^{j+1}} \left[\frac{\theta^2(i+1)^2\Gamma(j+k+2) + \theta(i+1)\Gamma(j+k+3) + \Gamma(j+k+4)}{\theta(i+1)^{j+k+4}} \right]$$

$$\mu_3' = \alpha \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} (-1)^i \binom{\alpha-1}{i} \binom{i}{j} \binom{j}{k} \frac{(\theta+2)^{j-k}}{(\theta^2+\theta+2)^{j+1}} \left[\frac{\theta^2(i+1)^2\Gamma(j+k+4) + \theta(i+1)\Gamma(j+k+5) + \Gamma(j+k+6)}{\theta^3(i+1)^{j+k+6}} \right]$$

$$\mu_4' = \alpha \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} (-1)^i \binom{\alpha-1}{i} \binom{i}{j} \binom{j}{k} \frac{(\theta+2)^{j-k}}{(\theta^2+\theta+2)^{j+1}} \left[\frac{\theta^2(i+1)^2\Gamma(j+k+5) + \theta(i+1)\Gamma(j+k+6) + \Gamma(j+k+7)}{\theta^4(i+1)^{j+k+7}} \right]$$

Therefore, the variance of ESD can thus be obtained using the formula

$$\mu_2 = \mu_2' - \left(\mu_1'\right)^2 .$$

Harmonic mean

Harmonic mean of ESD can be obtained as

$$E\left(\frac{1}{X}\right)=\int_0^\infty \frac{1}{x}g(x)dx$$

$$=\alpha\sum_{i=0}^\infty\sum_{j=0}^\infty\sum_{k=0}^\infty(-1)^i\binom{\alpha-1}{i}\binom{i}{j}\binom{j}{k}\frac{\theta^{j+k+3}(\theta+2)^{j-k}}{(\theta^2+\theta+2)^{j+1}}\left[\frac{\{\theta(i+1)\}^2\Gamma(j+k)+\{\theta(i+1)\}\Gamma(j+k+1)+\Gamma(j+k+2)}{\{\theta(i+1)\}^{j+k+2}}\right] \tag{11}$$

Moment generating function

Moment generating function of ESD can be obtained as

$$M_X(t)=E\left[e^{tX}\right]=\int_0^\infty e^{tx}g(x)dx$$

Using Taylor's series expansion, we get

$$M_X(t)=\int_0^\infty\left(1+tX\frac{(tX)^2}{2!}+...\right)g(x)dx=\int_0^\infty\sum_{r=0}^\infty\frac{t^r}{r!}x^r g(x)dx=\int_0^\infty\sum_{r=0}^\infty\frac{t^r}{r!}\mu_r'$$

$$=\alpha\sum_{i=0}^\infty\sum_{j=0}^\infty\sum_{k=0}^\infty\sum_{r=0}^\infty(-1)^i\binom{\alpha-1}{i}\binom{i}{j}\binom{j}{k}\frac{t^r}{r!}\frac{(\theta+2)^{j-k}}{(\theta^2+\theta+2)^{j+1}}$$

$$\times\left[\frac{\theta^2(i+1)^2\Gamma(j+k+r+1)+\theta(i+1)\Gamma(j+k+r+2)+\Gamma(j+k+r+2)}{\theta^r(i+1)^{j+k+r+3}}\right] \tag{12}$$

Order statistics of ESD

The pdf of the r th order statistics Y_r of ESD can be obtained as

$$g_{Y_r}(x)=\frac{n!}{(r-1)!(n-r)!}g(x)\left[G(x)\right]^{r-1}\left[1-G(x)\right]^{n-r}$$

$$=\frac{n!}{(r-1)!(n-r)!}\left[\frac{\alpha\theta^3}{\theta^2+\theta+2}\left(1+x+x^2\right)e^{-\theta x}\left\{1-\left(1+\frac{\theta x(\theta x+\theta+2)}{\theta^2+\theta+2}\right)e^{-\theta x}\right\}^{\alpha-1}\right] \tag{13}$$

Citation: Prodhani HR, Shanker R. An extended Sujatha distribution with statistical properties and applications. *Biom Biostat Int J.* 2024;13(4):96‒105.
DOI: 10.15406/bbij.2024.13.00420

The pdf of first order Statistics Y_1 of ESD can be obtained as

$$\times\left[1-\left\{1+\frac{\theta x\left(\theta x+\theta+2\right)}{\theta^2+\theta+2}\right\}e^{-\theta x}\right]^{\alpha(r-1)}\left[1-\left\{1-\left(1+\frac{\theta x\left(\theta x+\theta+2\right)}{\theta^2+\theta+2}\right)e^{-\theta x}\right\}^{\alpha}\right]^{n-r}$$

$$g_{Y_1}(x)=n\left[1-G(x)\right]^{n-1}g(x)$$

$$=n\left[1-\left\{1-\left(1+\frac{\theta x\left(\theta x+\theta+2\right)}{\theta^2+\theta+2}\right)e^{-\theta x}\right\}^{\alpha}\right]^{n-1}\frac{\alpha\theta^3}{\theta^2+\theta+2}\left(1+x+x^2\right)e^{-\theta x}$$

(14)

The pdf of n th order statistics Y_n of ESD can be obtained as

$$\times\left[1-\left\{1+\frac{\theta x\left(\theta x+\theta+2\right)}{\theta^2+\theta+2}\right\}e^{-\theta x}\right]^{\alpha-1} \quad (14)$$

$$g_{Y_n}(x)=n\left[G(x)\right]^{n-1}g(x)$$

$$=\frac{\alpha\theta^3}{\theta^2+\theta+2}\left(1+x+x^2\right)e^{-\theta x}\left[1-\left\{1+\frac{\theta x\left(\theta x+\theta+2\right)}{\theta^2+\theta+2}\right\}e^{-\theta x}\right]^{\alpha-1}$$

$$\times\left[\left\{1-\left(1+\frac{\theta x\left(\theta x+\theta+2\right)}{\theta^2+\theta+2}\right)e^{-\theta x}\right\}^{\alpha}\right]^{n-1} \quad (15)$$

Renyi entropy of ESD

Let X be a continuous random variable following ESD. The Renyi entropy generally deals with the measures of uncertainty or spread of X. The entropy of a random variable X defined by Renyi[20] can be obtained as

$$R_E=\frac{1}{1-\beta}\log\left(\int_0^{\infty}\left[g(x)\right]^{\beta}dx\right)$$

$$=\frac{1}{1-\beta}\left[\log\left\{\left(\frac{\alpha\theta^3}{\theta^2+\theta+2}\right)^{\beta}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\sum_{m=0}^{\infty}\sum_{r=0}^{\infty}\sum_{s=0}^{\infty}\left(-1\right)^i\binom{\beta\left(\alpha-1\right)}{i}\binom{i}{j}\binom{j}{k}\binom{k}{m}\binom{\beta}{r}\binom{r}{s}\frac{2^{k-m}}{\left(\theta^2+\theta+2\right)^j}\left\{\frac{\Gamma\left(2j+k+m+r+s+1\right)}{\theta^{r+s+1}\left(\beta+i\right)^{2j+k+m+r+s+1}}\right\}\right\}\right] \quad (16)$$

Estimation of the parameters

Maximum likelihood estimation of parameters

Let $x_1,x_2,...,x_n$ be the random sample of size n from ESD. The likelihood function can be written as

$$L\left(\alpha,\theta\right)=\left(\frac{\alpha\theta^3}{\theta^2+\theta+2}\right)^n\prod_{i=1}^{n}\left[\left(1+x+x^2\right)e^{-\theta x}\left\{1-\left(1+\frac{\theta x(\theta x+\theta+2)}{\theta^2+\theta+2}\right)e^{-\theta x}\right\}^{\alpha-1}\right]$$

The log likelihood function is given by

$$\log L(\alpha,\theta) = 3n\log\theta + n\log\alpha - n\log\left(\theta^2+\theta+2\right) + \sum_{i=1}^{n}\log\left(1+x+x^2\right) - \theta\sum_{i=1}^{n}x$$
$$+\left(\alpha-1\right)\sum_{i=1}^{n}\log\left\{1-\left(1+\frac{\theta x(\theta x+\theta+2)}{\theta^2+\theta+2}\right)e^{-\theta x}\right\} \quad (17)$$

For maximum likelihood estimates of parameters, we have to solve the following log-likelihood equations.

$$\frac{\partial \log L(\alpha,\theta)}{\partial\alpha} = \frac{n}{\alpha} + \sum_{i=1}^{n}\log\left\{1-\left(1+\frac{\theta x(\theta x+\theta+2)}{\theta^2+\theta+2}\right)e^{-\theta x}\right\} = 0$$

This gives

$$\hat{\alpha} = \frac{n}{\sum_{i=1}^{n}\log\left\{1-\left(1+\frac{\theta x(\theta x+\theta+2)}{\theta^2+\theta+2}\right)e^{-\theta x}\right\}}$$

$$\frac{\partial \log L(\alpha,\theta)}{\partial\theta} = \frac{3n}{\theta} - \frac{n\left(2\theta+1\right)}{\theta^2+\theta+2} - \sum_{i=1}^{n}x + \left(\alpha-1\right)\psi\left[1-\left(1+\frac{\theta x(\theta x+\theta+2)}{\theta^2+\theta+2}\right)e^{-\theta x}\right] = 0$$

where $\psi(.)$ is a digamma function.

Hence, it is very difficult to estimate the value of the parameter θ because the above-mentioned equation is too complicated. Therefore, we use sophisticated software like R for estimating the required parameter θ.

Maximum product spacing estimation of parameters

The maximum product spacing estimates (MPSE) $\left(\hat{\theta},\hat{\alpha}\right)$ of parameters $\left(\theta,\alpha\right)$ of ESD can be obtained by numerically by maximizing the following function with respect to θ and α.

$$MPSE = \frac{1}{n+1}\sum_{i=1}^{n+1}\log\left[F\left(x_i,\theta,\alpha\right) - F\left(x_{i-1},\theta,\alpha\right)\right]$$

Simulation study of ESD

In this section, we carried out simulation study to examine the performance of maximum likelihood estimators of the ESD. We examined the mean estimates, biases (B), mean square errors (MSEs) and variances of the maximum likelihood estimates (MLEs). The mean, bias, MSE and variance are computed using the formulae

$$Mean = \frac{1}{n}\sum_{i=1}^{n}\hat{H}_i,\quad B = \frac{1}{n}\sum_{i=1}^{n}\left(\hat{H}_i - H\right),\quad MSE = \frac{1}{n}\sum_{i=1}^{n}\left(\hat{H}_i - H\right)^2,$$

$$Variance = MSE - B^2$$

where $H = \theta,\alpha$ and $\hat{H}_i = \hat{\theta}_i,\hat{\alpha}_i$

the simulation results for different parameter values of ESD are presented in tables 1 & 2 respectively. The steps for simulation study are as follows:

a. Data is generated using the acceptance-rejection method of simulation. The acceptance-rejection method is a commonly used approach in simulation studies to generate random samples from a target distribution when inverse transform method of simulation is not feasible or efficient. Acceptance rejection method for generating random samples from the ESD consists of following steps.

i. Generate Y distributed as exponential $\left(\theta\right)$

ii. Generate U distributed as Uniform $\left(0,1\right)$

iii. If $U \le \frac{f(y)}{M\,g(y)}$, then set $X = Y$ ("accept the sample"); otherwise ("reject the sample") and if reject then repeat the process: step (i-iii) until getting the required samples.

Where M is a constant.

b. The sample sizes are taken as $n = 50, 100, 150, 200, 250, 300$

c. The parameter values are set as values $\theta = 1, \alpha = 3$ and $\theta = 1.5, \alpha = 6$

d. Each sample size is replicated 10000 times

The results obtained in Tables 1 & 2 show that as the sample size increases, biases, MSEs and variances of the MLEs of the parameters become smaller respectively. This result is in line with the first-order asymptotic theory of maximum likelihood estimators.

Variance-covariance matrix for parameter $\theta = 1, \alpha = 3$ and $\theta = 1.5, \alpha = 6$

$$\begin{array}{c} \\ \hat{\theta} \\ \hat{\alpha} \end{array}\begin{array}{cc} \hat{\theta} & \hat{\alpha} \\ \left(\begin{array}{cc} 0.00179 & 0.00027 \\ 0.00027 & 0.02140 \end{array}\right) \end{array} \quad \begin{array}{c} \\ \hat{\theta} \\ \hat{\alpha} \end{array}\begin{array}{cc} \hat{\theta} & \hat{\alpha} \\ \left(\begin{array}{cc} 0.00081 & 0.00016 \\ 0.00016 & 0.02877 \end{array}\right) \end{array}$$

Table 1 The mean values, biases, MSEs and variances of ESD for parameter $\theta = 1$, $\alpha = 3$

Sample size	Parameters	Mean	Bias	MSE	Variance
50	$\hat{\theta}$	1.01898	0.01898	0.00235	0.00199
	$\hat{\alpha}$	3.06882	0.06882	0.03697	0.03697
100	$\hat{\theta}$	1.01400	0.01400	0.00261	0.00241
	$\hat{\alpha}$	3.06173	0.06173	0.03074	0.02693
150	$\hat{\theta}$	1.01326	0.01326	0.00221	0.00204
	$\hat{\alpha}$	3.05937	0.05937	0.02867	0.02514
200	$\hat{\theta}$	1.01250	0.01250	0.00213	0.00198
	$\hat{\alpha}$	3.05253	0.05253	0.02620	0.02344
250	$\hat{\theta}$	1.01103	0.01103	0.00202	0.00190
	$\hat{\alpha}$	3.04911	0.04911	0.02441	0.02199
300	$\hat{\theta}$	1.01089	0.01089	0.00190	0.00178
	$\hat{\alpha}$	3.04626	0.04626	0.02347	0.02133

Table 2 The mean values, biases, MSEs and variances of ESD for parameter $\theta = 1.5$, $\alpha = 6$

Sample size	Parameters	Mean	Bias	MSE	Variance
50	$\hat{\theta}$	1.50881	0.00881	0.00134	0.00127
	$\hat{\alpha}$	6.12457	0.12457	0.05017	0.03465
100	$\hat{\theta}$	1.50346	0.00346	0.00114	0.00113
	$\hat{\alpha}$	6.11594	0.11594	0.05175	0.05089
150	$\hat{\theta}$	1.50133	0.00133	0.00108	0.00107
	$\hat{\alpha}$	6.11273	0.11273	0.05059	0.03788
200	$\hat{\theta}$	1.50044	0.00044	0.00094	0.00094
	$\hat{\alpha}$	6.09288	0.09288	0.04559	0.03696
250	$\hat{\theta}$	1.49980	-0.00019	0.00089	0.00089
	$\hat{\alpha}$	6.09316	0.09316	0.04067	0.03199
300	$\hat{\theta}$	1.49997	-0.00003	0.00080	0.00080
	$\hat{\alpha}$	6.09203	0.09203	0.03715	0.02868

Applications

The applications and the goodness of fit of the ESD have been discussed in this section with two examples of observed real datasets. The following datasets have been considered.

Dataset 1: The first dataset represents the breaking stress of carbon fibers of 50 mm length (GPa) reported by Nicolas and Padgett[21], the dataset is presented below

0.39,0.85,1.08,1.25,1.47,1.57,1.61,1.61,1.69,1.80,1.84,1.87,1.89,2.03,2.03,2.05,2.12,2.35,2.41,2.43,2.48,2.50,2.53,2.55,2.55,2.56,2.59,2.67,2.73,2.74,2.79,2.81,2.82,2.85,2.87,2.88,2.93,2.95,2.96,2.97,3.09,3.11,3.11,3.15,3.15,3.19,3.22,3.22,3.27,3.28,3.31,3.31,3.33,3.39,3.39,3.56,3.60,3.65,3.68,3.70,3.75,4.20,4.38,4.42,4.70,4.90.

Dataset 2: The following data represents the symmetric behavior of the tensile strength about 100 observations of carbon fibers, discussed by Nicolas and Padgett[21]the observations are:

3.7, 3.11, 4.42, 3.28, 3.75, 2.96, 3.39, 3.31, 3.15, 2.81, 1.41, 2.76, 3.19, 1.59, 2.17, 3.51, 1.84, 1.61, 1.57, 1.89, 2.74, 3.27, 2.41, 3.09, 2.43, 2.53, 2.81, 3.31, 2.35, 2.77, 2.68, 4.91, 1.57, 2.00, 1.17, 2.17, 0.39, 2.79, 1.08, 2.88, 2.73, 2.87, 3.19, 1.87, 2.95, 2.67, 4.20, 2.85, 2.55, 2.17, 2.97, 3.68, 0.81, 1.22, 5.08, 1.69, 3.68, 4.70, 2.03, 2.82, 2.50, 1.47, 3.22, 3.15, 2.97, 1.61, 2.05, 3.60, 3.11, 1.69, 4.90, 3.39, 3.22, 2.55, 3.56, 2.38, 1.92, 0.98, 1.59, 1.73, 1.71, 1.18, 4.38, 0.85, 1.80, 2.12, 3.65.

The figure 4 & 5 represents the total time on test (TTT) plot for the both observed two samples and simulated samples of ESD respectively. The figure 4 & 6 indicate that both the observed samples have decreasing failure rate.

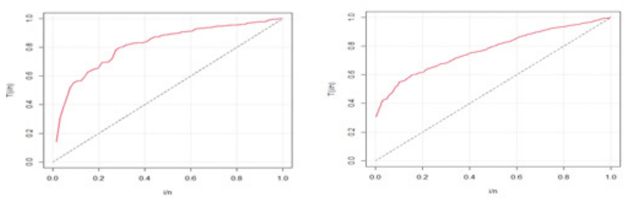

Figure 4 TTT- plot of the observed sample-1and simulated samples of ESD respectively.

The pdf of Exponentiated Exponential distribution (EED) and exponentiated Aradhana distribution (EAD) are given by

$$f\left(x;\theta,\alpha\right) = \alpha\theta e^{-\theta x}\left(1-e^{-\theta x}\right)^{\alpha-1} \quad ; x > 0, \alpha > 0, \theta > 0$$

DOI: 10.15406/bbij.2024.13.00420

$$f(x;\theta,\alpha)=\frac{\alpha\theta^3(1+x)^2e^{-\theta x}}{\theta^2+2\theta+2}\left(1-\left(1+\frac{\theta x(\theta x+2\theta+2)}{\theta^2+2\theta+2}\right)e^{-\theta x}\right)^{\alpha-1};x>0,\alpha>0,\theta>0$$

In order to compare lifetime distributions, values of $-2\log L$, Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), Consistent Akaike Information Criterion (CAIC), Hannan-Quinn Information Criterion (HQIC), Kolmogorov-Smirnov Statistics (K-S) and the corresponding probability value (p-value) for the above data set has been computed. The formulae for computing AIC, BIC, CAIC, HQIC and K-S are as follows:

$$AIC=-2\log L+2p,\ BIC=-2\log L-p\log(n),$$

$$CAIC=-2\log L+\frac{2pn}{n-p-1}$$

$$HQIC=-2\log L+2p\log\left[\log(n)\right]$$

$$,K\text{-}S=\underset{x}{Sup}\,|F_m(x)-F_o(x)|$$

where, p = number of parameters, n = sample size, $F_m(x)$ = empirical cdf of considered distribution and $F_o(x)$ = cdf of considered distribution.

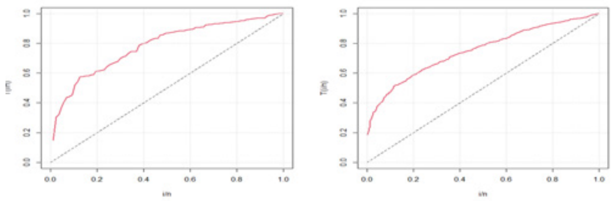

Figure 5 TTT- plot of the observed sample-2 and simulated samples of ESD respectively.

The distribution corresponding to the lower values of $-2\log L$, AIC, BIC, CAIC, HQIC and K-S Statistics and higher values of p-value is the best fit distribution. These statistical values for the two datasets have been computed and presented in tables 4 & 6 respectively. The estimation of the parameters using the method of MLE and MPSE for the two datasets is presented in the table 3 & 5 respectively.

Table 3 MLE's and MPSE's with standard errors of the considered distributions for the dataset-I

Distributions	MLE		MPSE	
	$\hat{\theta}$ SE($\hat{\theta}$)	$\hat{\alpha}$ SE($\hat{\alpha}$)	$\hat{\theta}$ SE($\hat{\theta}$)	$\hat{\theta}$ SE($\hat{\alpha}$)
Sujatha	0.8534 (0.0607)	…	0.8485 (0.0602)	…
EED	1.0075 (0.1002)	9.1992 (2.1491)	0.8068 (0.0547)	7.5849 (2.8926)
EAD	1.4808 (0.1180)	5.5104 (1.3217)	1.4839 (0.1169)	5.2793 -1.2459
ESD	1.5248 (0.1185)	5.9074 (1.4465)	1.5276 (0.1177)	5.6613 (1.3664)

Table 4 Goodness of fit measures for the dataset-I

Distributions	$-2\log L$	AIC	BIC	CAIC	HQIC	K-S	P-value
Sujatha	229.59	231.59	233.78	231.65	232.45	0.2261	0.0052
EED	190.74	194.74	199.11	194.93	196.47	0.1589	0.1105
EAD	185.15	189.15	193.53	189.34	190.88	0.1599	0.1049
ESD	184.27	188.27	192.65	188.46	190.00	0.1404	0.4108

Table 5 MLE's and MPSE's with standard errors of the considered distributions for the dataset-2

Distributions	MLE		MPSE	
	$\hat{\theta}$ SE($\hat{\theta}$)	$\hat{\alpha}$ SE($\hat{\alpha}$)	$\hat{\theta}$ SE($\hat{\theta}$)	$\hat{\alpha}$ SE($\hat{\alpha}$)
Sujatha	0.8863 (0.0550)	…	0.8819 (0.0546)	…
EED	1.0045 (0.0923)	7.7135 (1.5724)	0.8624 (0.0525)	7.2864 (2.0352)
EAD	1.4695 (0.1084)	4.5305 (0.9438)	1.5399 (0.1080)	4.9071 (0.9966)
ESD	1.5141 (0.1088)	4.8594 (1.0342)	1.5847 (0.1087)	5.2822 (1.0987)

DOI: 10.15406/bbij.2024.13.00420

From the table 4 &6 we observed that the ESD have the least -2logL, AIC, BIC, CAIC, HQIC and K-S values as compared to EED, EAD and Sujatha distribution. Hence, we may conclude that ESD provides a better fit than EED, EAD and Sujatha distribution.

Table 6 Goodness of fit measures for the dataset-2

Distributions	$-2\log L$	AIC	BIC	CAIC	HQIC	K-S	P-value
Sujatha	299.11	301.11	303.29	301.17	301.97	0.2115	0.0026
EED	254.81	258.81	263.18	259	260.54	0.0922	0.5653
EAD	250.36	254.36	258.74	254.55	256.09	0.1129	0.3084
ESD	249.70	253.70	258.07	253.89	255.43	0.0801	0.8731

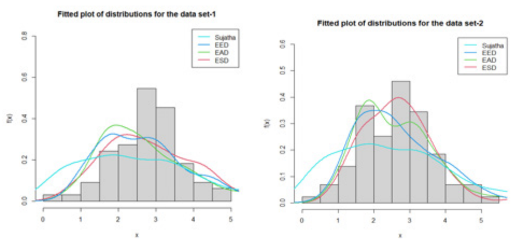

Figure 6 Fitted plot of distributions of the dataset-1 and dataset-2.

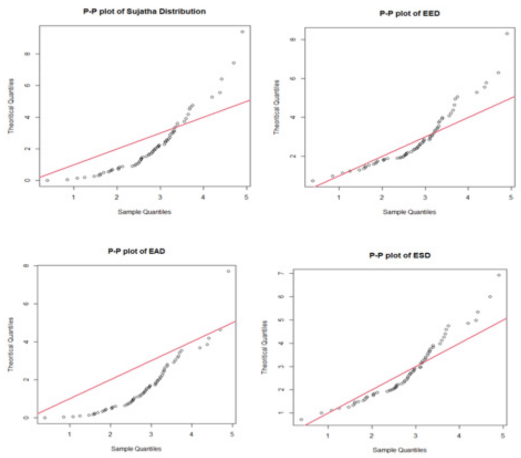

Figure 7 P-P plots of the theoretical and sample quantiles of the considered distributions of the dataset-1.

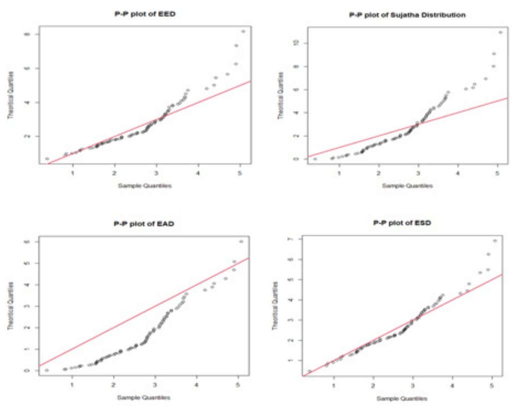

Figure 8 P-P plots of the theoretical and sample quantiles of the considered distributions of the dataset-2.

From the fitted plot and the P-P plot of the considered distribution presented in the figure 6, 7 & 8 for the both datasets also exhibit that ESD provides a better fit as compared to the considered distributions.

Conclusions

In this paper an extended Sujatha distribution (ESD) has been proposed. Its statistical properties including survival function, hazard function, harmonic mean, moment generating function, order statistics and Renyi entropy have been discussed. Moments of the proposed distribution has been obtained. The parameters of this distribution have been estimated using maximum likelihood estimation method and maximum product spacing method. The simulation study has been presented to know the performance of maximum likelihood estimates as the sample size increases. Finally, two examples of real lifetime datasets have been considered for applications and compared with the EED, EAD and Sujatha distribution. It has been found that ESD provides the best fit than the EED, EAD and Sujatha distribution.

Funding information

No funding was received from any financial organization to conduct this research.

Acknowledgements

Authors are grateful to the editor-in-chief and the anonymous reviewer for their comments which improved the quality of the paper.

Declaration of competing interest

The authors declare that they have no any known financial or non-financial competing interests in any material discussed in this paper.

References

1. Gupta RC, Gupta PL, Gupta R. Modeling failure time data by Lehman alternatives. *Communication in Statistics – Theory and Methods*. 1998:7(4):887–904.
2. Gupta RD, Kundu D. Exponentiated exponential family: an alternative to gamma and Weibull distribution. *Biometrical Journal*. 2001;43(1):117–130.
3. Pal M, Ali MM, Woo J. Exponentiated Weibull distribution. *Statistica*. 2006;66(2):139–147.
4. Nadarajah S, Bakouch HS, Tahmasbi R. A generalized Lindley distribution. *Sankhya B*. 2011;73:331–359.
5. Rodrigues JA, Percontini AC, Hamedani GG. The exponentiated generalized Lindley distribution. *Asian Research Journal of Mathematics*. 2017;5(3):1–14.
6. Jayakumar B, Elangovan R. Exponentiated Shankar distribution and their applications in breast cancer data. *Science, Technology and Development.* 2019;8(10):418–431.
7. Rather AA, Subramanian C. Exponentiated Ishita. Distribution with properties and application. *International Journal of Management Technology Engineering*. 2019;9(5):2473–2484.

8. Ganaie RA, Rajagopalan V. Exponentiated aradhana distribution with properties and applications in engineering sciences. *Journal of Scientific Research, BHU.* 2022;66(1):316–325.
9. Shanker R. Sujatha distribution and its applications. *Statistics in Transition new Series*. 2016;17(3):391–410.
10. Shanker R. A quasi sujatha distribution. *International Journal of Probability and Statistics.* 2016;5(4):89–100.
11. Shanker R, Shukla KK, Hagos F. A generalization of Sujatha distribution and its applications to real lifetime data. *Journal of Institute of Science and Technology*. 2017;22(1):77–94.
12. Mussie T, Shanker R. A two parameter Sujatha distribution. *Biom Biostat Int J*. 2018;7(3):188–197.
13. Mussie T, Shanker R. A new two parameter Sujatha distribution. *Turkiye Klinikleri Journal of Biostatistics.* 2018;10(2):96–113.
14. Shanker R. Another two parameter Sujatha distribution with properties and applications. *Journal of Mathematical Sciences and Modelling.* 2019;2(1):1–13.
15. Shanker R, Shukla KK. A two-parameter weighted Sujatha distribution and its application to model life time data. *International Journal of Mathematics and Statistics*. 2018:57(3):106–121.
16. Shanker R, Shukla KK. A two parameter power Sujatha distribution with properties and application. *International Journal of Mathematics and Statistics.* 2019;20(3):11–22.
17. Shanker R, Shukla KK. A new quasi sujatha distribution. *Statistics in Transition new Series.* 2020;21(3):53–71.
18. Okoli OM, Osuji GA, Onyekwere CK. Generalized inverse power Sujatha distribution with application. *Asian Journal of Probability and Statistics*. 2021;15(3):11–25.
19. Ikechukwu AF, Eghwerido JT. Marshall-Olkin Sujatha distribution and its application. *Thail Stat.* 2022;20(1):36–52.
20. Renyi A. On measures of entropy and information. *Symposium on Mathematical Statistics and Probability,* 1961;4(1):547–561.
21. Nicholas MD, Padgett WJ. A bootstrap control for weibull percentiles. *Quality and Reliability Engineering International.* 2006;22:141–151.