JJMS 17, No. 3, 437-462 (2024)

Jordan Journal of Mathematics and Statistics.
Yarmouk University

DOI:https://doi.org/10.47013/17.3.10

Sine exponential pareto distribution: properties, estimation, and applications

Doaa abd Khalil ELhertaniy

Department of Statistics, The Holy Quran and Taseel of Science university,Sudan.

Received: July 4, 2023 Accepted : April 30, 2024

Abstract: The work of the Sine-G distribution family is extended in this paper. The applicability of the sine exponential Pareto distribution is highlighted via the goodness-of-fit approach to data. Various properties of the suggested distribution, including moments, quantiles, entropy, and order statistics are acquired. For model parameters estimation, the maximum likelihood technique is used. Four extensive data sets are empirically used to demonstrate the potential significance and applicability of the proposed distribution. The results of the investigation indicated that the Sine Exponential Pareto distribution was superior than numerous other competing distributions.

Keywords: Sine-G distribution; Sine Exponential Pareto distribution; entropy.

2010 Mathematics Subject Classification. 26A25; 26A35.

1 Introduction

In general operator theory, inverse and half-inverse problems have received a lot of attention lately[5],[6], and[11]. Furthermore, a lot of authors have emphasized the importance of generated families of distributions, such as the Kumaraswamy-G[10], sine generated (S-G) family by[18], Weibull-G[7], odd Frechet-G[12], Burr type X-G[31], truncated Cauchy power-G[3], Type-I half-logistic Burr X-G[4], generalized odd half-Cauchy-G[9], exponentiated sine-G family[23], Sine Half-Logistic Inverse Rayleigh [28], Sine Modified Lindley [29], Sin Topp-Leone-generated family[1], beyond the Sin-G family[15] and others. The Sine-G (S-G) family of distributions, a recently developed family of distributions, was first introduced in [18]. The cumulative distribution function (CDF) of the S-G family.

$$F(x;\Psi) = sin[\frac{\pi}{2}G(x;\Psi)], \quad x \in R. \tag{1}$$

where $G(x;\Psi)$ is the CDF of baseline model with parameter vector Ψ.

The probability density function (PDF) of the S-G family is

$$f(x;\Psi) = \frac{\pi}{2}g(x;\Psi)cos[\frac{\pi}{2}G(x;\Psi)], \quad x \in R. \tag{2}$$

A well-known probability model for modeling and forecasting many socioeconomic elements is the Pareto distribution, which bears the name of swiss economist Vilfredo Pareto (1848-1923).

Pareto distribution has been studied in literature in numerous applications in lifetests, climate science, economics, finance, biology, physics, and actuarial science. Even though the distribution has many applications, studying the distribution of income is one of its most significant and significant applications. [25] pioneered this approach in his personal economic texts. Studying some of the distribution's applications in simulating earthquakes, forest fire zones, oil

* Corresponding author e-mail: dalhirtani@gmail.com

and gas field sizes [8].[2]proposed a novel distribution known as the exponential Pareto distribution with (CDF) and (PDF) are given by

$$G(x) = 1 - e^{(-\sigma\left(\frac{x}{\mu}\right)^k)}, x > 0. \tag{3}$$

$$g(x) = \frac{\sigma k x^{k-1}}{\mu^k} e^{(-\sigma\left(\frac{x}{\mu}\right)^k)}, x > 0. \tag{4}$$

where σ, $k \in R^+$ are the shape parameters and $\mu > 0$ is a scale parameters.

To improve the flexibility of the S-G family of distributions without adding any additional parameters, to produce heavy-tailed distributions with fewer parameters that provide a better parametric fit to a given data set than some existing distributions, to produce distributions that are roughly symmetric, right-skewed, and reversed-J shaped, and to generate distributions capable of mode estimation are the main reasons for our proposal. To include all of these variables, it is necessary to expand the S-G family of distributions. As a result, the present paper is strivingto develop a new distribution that is more flexible than the Exponential Pareto model using the S-G family. It may be viewed as a useful model for fitting asymmetric data that may not be well fitted by some common models, and it can be used to solve a variety of problems in many fields, as well as to provide more accurate fits than some popular models with good results for some existing distributions. The new model is called the Sine Exponential Pareto (S-EP) model. The manuscript is organized as follows. Section 2 is the S-EP Distribution. Some distributional properties of the proposed (S-EP) distribution are derived in Section 3. Entropy is discussed in Section 4. Section 5 presents the distributions of various order statistics. Section 6 discusses maximum likelihood estimation of model parameters. Outlines the suggested model's Monte Carlo simulation are presented in Section 7. Section 8 includes the application of the distribution to four data, and Section 9 concludes the study.

2 The S-EP Distribution

Letting a random variable X have S-EP distribution, then the CDF, PDF, survival function (SF), and hazard rate function (HRF) of X are

$$F(x;\sigma,\mu,k) = \sin\left[\frac{\pi}{2}[1 - e^{\left(-\sigma\left(\frac{x}{\mu}\right)^k\right)}]\right], \; x,\sigma,\mu,k > 0, \tag{5}$$

$$f(x;\sigma,\mu,k) = \frac{\pi\sigma k x^{k-1}}{2\mu^k} e^{\left(-\sigma\left(\frac{x}{\mu}\right)^k\right)} \cos\left[\frac{\pi}{2}[1 - e^{\left(-\sigma\left(\frac{x}{\mu}\right)^k\right)}]\right], \tag{6}$$

$$R(x) = 1 - \sin\left[\frac{\pi}{2}[1 - e^{\left(-\sigma\left(\frac{x}{\mu}\right)^k\right)}]\right], \tag{7}$$

$$h(x) = \frac{\frac{\pi\sigma k x^{k-1}}{2\mu^k} e^{\left(-\sigma\left(\frac{x}{\mu}\right)^k\right)} \cos\left[\frac{\pi}{2}[1 - e^{\left(-\sigma\left(\frac{x}{\mu}\right)^k\right)}]\right]}{1 - \sin\left[\frac{\pi}{2}[1 - e^{\left(-\sigma\left(\frac{x}{\mu}\right)^k\right)}]\right]}. \tag{8}$$

where $\sigma > 0$ and $k > 0$ are shape parameters and $\mu > 0$ is a scale parameter.

Figure 1 and 2 show various shapes of S-EP distribution for the PDF , HRF and CDF, SF at different parameter values of σ, μ and k. PDF has various shapes including the reverse-J, unimodal, right-skewed, decreasing, and approximately symmetric shapes, and various degrees of kurtosis. Moreover, the HRF can take several shapes, such as the increasing, decreasing, and bathtub shapes. These different behaviours indicate the flexibility and adaptability for the S-EP distribution to fit a variety of data shapes.

3 Statistical Properties

Several statistical features of the S-EP distribution are shown in this section. These qualities include moments, the moment generating function, the quantile function, and random sample simulation.

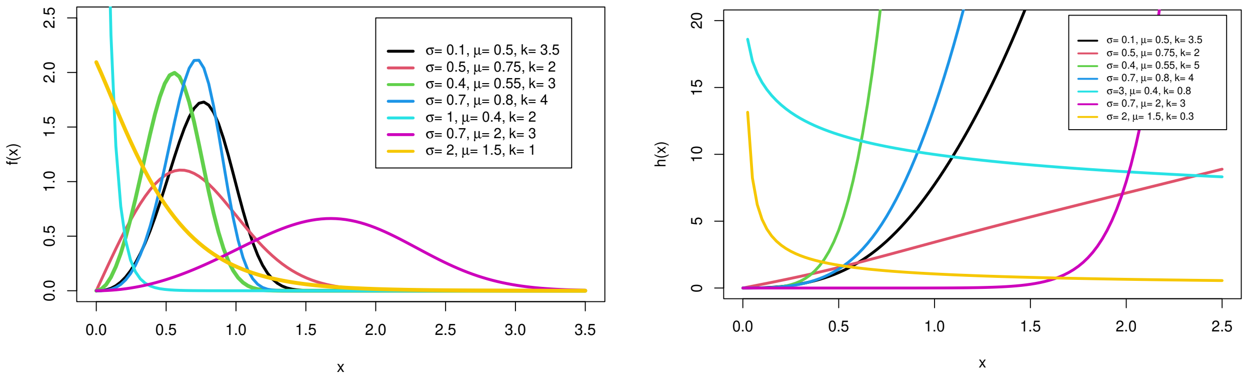

Fig. 1: The PDF and HRF for the S-EP distribution

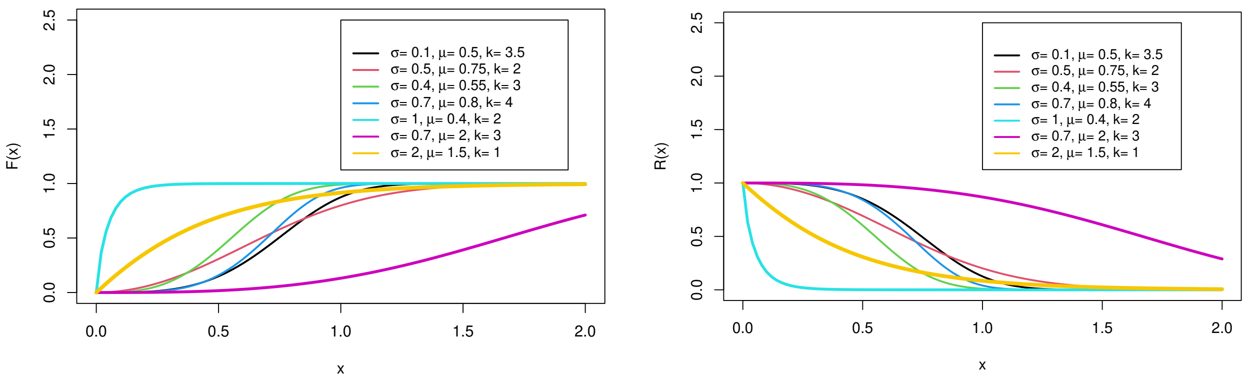

Fig. 2: The CDF and SF for the S-EP distribution

3.1 The Moments

Theorem 1.*Let X be a random variable having the S-EP distribution, then the r^{th} moment of X about the origin is*

$$E\left(X^r\right) = \frac{\mu^{k+r}}{k}\sum_{i=0}^{\infty}\sum_{j=0}^{2i}\varphi_{i,j}\frac{\Gamma\left(1+\left(\frac{r}{k}\right)\right)}{\left(\sigma(1+j)\right)^{1+\frac{r}{k}}},$$

$r = 1,2,\ldots$ *,where $\Gamma(1+r/k)$ is the gamma function.*

*Proof.*The r^{th} moment is given by

$$E\left(X^r\right) = \int_0^{\infty} x^r f(x)\,dx$$

$$= \int_0^{\infty} x^r \frac{\pi\sigma k x^{k-1}}{2\mu^k} e^{\left(-\sigma\left(\frac{x}{\mu}\right)^k\right)} \cos\left[\frac{\pi}{2}[1-e^{\left(-\sigma\left(\frac{x}{\mu}\right)^k\right)}]\right] dx$$

$$= \int_0^{\infty} \frac{\pi\sigma k x^{k+r-1}}{2\mu^k} e^{\left(-\sigma\left(\frac{x}{\mu}\right)^k\right)} \cos\left[\frac{\pi}{2}[1-e^{\left(-\sigma\left(\frac{x}{\mu}\right)^k\right)}]\right] dx$$

By inserting the expansion $cos[z(x)] = \sum_{i=0}^{\infty} \frac{(-1)^i}{(2i)!} z(x)^{2i}$, then

$$= \sum_{i=0}^{\infty} \frac{\sigma k(-1)^i}{\mu^k (2i)!} \left(\frac{\pi}{2}\right)^{2i+1} \int_0^{\infty} x^{k+r-1} e^{\left(-\sigma\left(\frac{x}{\mu}\right)^k\right)} [1 - e^{\left(-\sigma\left(\frac{x}{\mu}\right)^k\right)}]^{2i} dx \tag{9}$$

By the general binomial expansion, we have

$$[1 - e^{(-\sigma(\frac{x}{\mu})^k)}]^{2i} = \sum_{j=0}^{2i} \binom{2i}{j} (-1)^j e^{(-\sigma j(\frac{x}{\mu})^k)} \tag{10}$$

Substitute from (10) in (9), to get

$$E\left(X^r\right) = \sum_{i=0}^{\infty} \sum_{j=0}^{2i} \frac{\sigma k(-1)^{i+j}}{\mu^k (2i)!} \left(\frac{\pi}{2}\right)^{2i+1} \binom{2i}{j} \int_0^{\infty} x^{k+r-1} e^{\left(-\sigma(1+j)\left(\frac{x}{\mu}\right)^k\right)} dx$$

The last equation can be rewritten as follows:

$$E\left(X^r\right) = \sum_{i=0}^{\infty} \sum_{j=0}^{2i} \varphi_{i,j} \int_0^{\infty} x^{k+r-1} e^{\left(-\sigma(1+j)\left(\frac{x}{\mu}\right)^k\right)} dx,$$

Where $\varphi_{i,j} = \frac{\sigma k(-1)^{i+j}}{\mu^k (2i)!} \left(\frac{\pi}{2}\right)^{2i+1} \binom{2i}{j}$

Using the transformation $y = (\frac{x}{\mu})^k$ then $x = \mu \sqrt[k]{y}$ and $dx = \frac{\mu}{k} \sqrt[k]{y^{1-k}} dy$ using y in the above equation, we get

$$E\left(X^r\right) = \sum_{i=0}^{\infty} \sum_{j=0}^{2i} \varphi_{i,j} \int_0^{\infty} \mu^{k+r-1} y^{\frac{k+r-1}{k}} e^{(-\sigma(1+j)y)} \frac{\mu}{k} \sqrt[k]{y^{1-k}} dy$$

$$= \frac{\mu^{k+r}}{k} \sum_{i=0}^{\infty} \sum_{j=0}^{2i} \varphi_{i,j} \int_0^{\infty} y^{\frac{r}{k}} e^{-(\sigma(1+j))y} dy$$

Using the relation $\int_0^{\infty} t^b e^{-at} dt = \frac{\Gamma(1+b)}{a^{(1+b)}}$, then $E\left(X^r\right)$ becomes

$$E\left(X^r\right) = \frac{\mu^{k+r}}{k} \sum_{i=0}^{\infty} \sum_{j=0}^{2i} \varphi_{i,j} \frac{\Gamma(1 + \left(\frac{r}{k}\right))}{(\sigma(1+j))^{1+\frac{r}{k}}}$$

3.2 *The Moment Generating Function*

Theorem 2.*Let X be a random variable which has the S-EP distribution, then the moment generating function (MGF) of X is*

$$M_X(t) = \frac{\mu^{k+r}}{k} \sum_{r,i=0}^{\infty} \sum_{j=0}^{2i} \varphi_{i,j} \frac{t^r}{r!} \frac{\Gamma(1 + \left(\frac{r}{k}\right))}{(\sigma(1+j))^{1+\frac{r}{k}}}$$

*Proof.*We know that

$$M_X(t) = E\left(e^{tx}\right) = \int_{-\infty}^{\infty} e^{tx} f(x)\, dx$$

Using series expansion of e^{tx},

$$M_X(t) = \sum_{r=0}^{\infty} \frac{t^r}{r!} \int_0^{\infty} x^r f(x)\, dx = \sum_{r=0}^{\infty} \frac{t^r}{r!} E\left(x^r\right)$$

Then

$$M_X(t) = \frac{\mu^{k+r}}{k} \sum_{r,i=0}^{\infty} \sum_{j=0}^{2i} \varphi_{i,j} \frac{t^r}{r!} \frac{\Gamma(1 + \left(\frac{r}{k}\right))}{(\sigma(1+j))^{1+\frac{r}{k}}}$$

Some moments for selected parameters values in order (σ,μ,k): (0.1, 0.5, 3.5), (0.1, 0.5, 4), (0.1, 0.5, 5), (0.4, 0.5, 3.5), (0.7, 0.5, 3.5), (0.4, 0.7, 3.5) and (0.4, 1.5, 3.5) are given in Table 1, where SD, CV, CS, and CK represent the standard deviation, coefficient of variation, skewness, and kurtosis, respectively, and 3D plots of skewness and kurtosis for the S-EP distribution are given in Figures 2. We observe that When μ is held constant, we notice that an increase in k causes an increase in skewness and kurtosis, as illustrated in Figure 2.

	(0.1, 0.5, 3.5)	**(0.1, 0.5, 4)**	**(0.1, 0.5, 5)**	**(0.4, 0.5, 3.5)**	**(0.7, 0.5, 3.5)**	**(0.4, 0.7, 3.5)**	**(0.4, 1.5, 3.5)**
E(X)	0.6026522	0.6486755	0.64963498	0.49863823	0.425084887	0.6168918	0.11154194
$E(X^2)$	0.4460229	0.4712642	0.44305775	0.27142858	0.197252155	0.4435175	0.09078230
$E(X^3)$	0.3468944	0.3584487	0.31363699	0.15785282	0.097806623	0.3362790	0.07649048
$E(X^4)$	0.2798668	0.2824809	0.22898408	0.09680572	0.051153266	0.2653076	0.06606055
$E(X^5)$	0.2322340	0.2289977	0.17162989	0.06205734	0.027975064	0.2158695	0.05811839
$E(X^6)$	0.1970410	0.1899718	0.13160518	0.04132305	0.015899286	0.1800124	0.05187149
SD	0.1702112	0.1606417	0.10295618	0.02844640	0.009347271	0.1531403	0.04683102
CV	0.1492225	0.1380442	0.08199076	0.02016803	0.005664119	0.1324482	0.04267927
CS	0.1324445	0.1202609	0.06634507	0.01468140	0.003527533	0.1161469	0.03920086
CK	0.1187838	0.1060074	0.05446345	0.01094513	0.002252559	0.1030520	0.03624470

Table 1: Moments for selected parameters for S-EP distribution

From Table 1 observe that:

1) when we fix the parameters σ,μ, and parameter k increase leads to increase in mean and decrease in standard deviation, coefficient of variation, skewness, and kurtosis.

2) when we fix the parameters μ and k, and parameter σ increase leads to decrease in mean, standard deviation, coefficient of variation, skewness, and kurtosis.

3) when we fix the parameters σ and k, and parameter μ increase leads to increase and decrease in mean, standard deviation, coefficient of variation, skewness, and kurtosis.

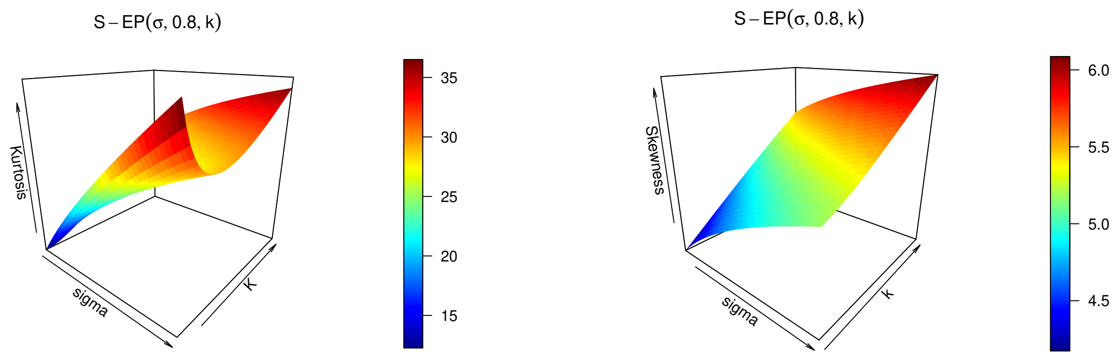

Fig. 3: Kurtosis and Skewness for the S-EP distribution for some selected parameter values

3.3 Quantile function

The quantile function theorem of the S-EP distribution is stated as follow;

Theorem 3. *Let X be a random variable from the S-EP distribution with parameters, $\sigma, \mu, k > 0$. Then the quantile function of X is given by*

$$Q(u) = \mu \left[\frac{-ln(-\frac{2}{\pi}\sin^{-1}(u)+1)}{\sigma} \right]^{\frac{1}{k}}$$

where u is a uniform random variable on (0, 1).

Proof. To compute the quantile function of the S-EP distribution, we substitute $F(x)$ by u where $0 < u < 1$ in (5) to get the equation

$$u = \sin\left[\frac{\pi}{2}[1 - e^{\left(-\sigma\left(\frac{x}{\mu}\right)^k\right)}]\right] \tag{11}$$

Then, we solve the equation (11)

$$\sin^{-1}(\text{u}) = \left[\frac{\pi}{2}[1 - e^{\left(-\sigma\left(\frac{x}{\mu}\right)^k\right)}]\right] \tag{12}$$

Then,

$$-\frac{2}{\pi}\ \sin^{-1}(\text{u}) + 1 = e^{\left(-\sigma\left(\frac{x}{\mu}\right)^k\right)} \tag{13}$$

Take Logarithm to both sides to get

$$\ln(1 - \frac{2}{\pi}\ \sin^{-1}(\text{u})) = -\sigma\left(\frac{x}{\mu}\right)^k$$

Then, we get the equation

$$\frac{-\ln(1-\frac{2}{\pi}\ \sin^{-1}(\text{u}))}{\sigma} = \left(\frac{x}{\mu}\right)^k$$

then,

$$\mu(\frac{-\ln(1-\frac{2}{\pi}\ \sin^{-1}(\text{u}))}{\sigma})^{\frac{1}{k}} = x$$

which finally gives the quantile function of S-EP the distributions as

$$x = F^{-1}(u) = Q(u) = \mu(\frac{-\ln(1-\frac{2}{\pi}\ \sin^{-1}(\text{u}))}{\sigma})^{\frac{1}{k}}$$

and by setting u = 0.5, we get the median as

$$Q(0.5) = \mu \left[\frac{-ln(-\frac{2}{\pi}sin^{-1}(0.5)+1)}{\sigma} \right]^{\frac{1}{k}}$$

The equivalent Bowley skewness[16] and Moors kurtosis can also be expressed in a similar manner ([22]).
The Bowley skewness based on quartiles is given by

$$B = \frac{Q\left(\frac{3}{4}\right) + Q\left(\frac{1}{4}\right) - 2Q\left(\frac{2}{4}\right)}{Q\left(\frac{3}{4}\right) - Q\left(\frac{1}{4}\right)}$$

The Moors kurtosis based on octiles is given by

$$M = \frac{Q\left(\frac{3}{8}\right) - Q\left(\frac{1}{8}\right) + Q\left(\frac{7}{8}\right) - Q\left(\frac{5}{8}\right)}{Q\left(\frac{6}{8}\right) - Q\left(\frac{2}{8}\right)}$$

as Table2 indicates the values of the first quartile, median, third quartile, Bowley skewness and Moors kurtosis of the S-EP distribution for the following selected parameters values in order (σ, μ , k): (0.5, 0.75, 2), (0.8, 0.75, 2) (0.4, 0.55, 3), (0.1, 0.5, 3.5), and (0.7, 0.8, 4). We observe that the Bowley skewness is minus if $k > 2$ and Bowley skewness and Moors kurtosis depend on the all parameter (σ, μ ,k) .

	Q(1/4)	median	Q(3/4)	B	M
(0.5, 0.75, 2)	0.4442	0.6754	0.9345	0.05690	1.204568
(0.8, 1.5, 2)	0.7023	1.0679	1.4776	0.05688	1.204566
(0.4, 0.55, 3)	0.4178	0.5525	0.6860	-0.0045	1.2133
(0.1, 0.5, 3.5)	0.5871	0.7459	0.8980	-0.0216	1.2193
(0.7, 0.8, 4)	0.5660	0.6979	0.8210	-0.0345	1.2249

Table 2: First quartile, median, third quartile, Bowley skewness and Moors kurtosis of the S-EP distribution

4 Entropy

In investigations of dependability and risk, measurement of entropy is important.It is widely used in a variety of biological, medical and physical applications.

4.1 Rényi Entropy

If X is a non-negative continuous random variable with PDF $f(x)$, then the Rényi entropy of order δ [26] of X is defined as

$$H_{\delta}(x) = \frac{1}{1-\delta} log \int_{0}^{\infty} [f(x)]^{\delta} dx, \forall \delta > 0, (\delta \neq 1) \tag{14}$$

We had the representation supplied by[32] for a power series raised to the power of a positive integer as the last significant result that was known, and we referred to it in the following lemma.

Lemma 1.*Established by[32]. For a given power series of the form* $(\sum_{k=0}^{\infty} a_k x^k)$ *and a positive integer n, we had:*

$$\left(\sum_{k=0}^{\infty} a_k x^k\right)^n = \sum_{k=0}^{\infty} c_k x^k, \tag{15}$$

where $c_0 = a_0^n$, $c_m = [1/(ma_0)] \sum_{k=1}^{m} (kn - m + k) a_k c_{m-k}$ *for* $m \geq 1$.

Theorem 4.*The Rényi entropy of a random variable X* $\sim$ *S-EP distribution* (μ,σ,k), *with,* μ,σ, $k > 0$*is given by*

$$H_{\delta}(x) = \frac{1}{1-\delta} log \left[b(\sigma, \mu, k, j, i, \delta, c_i) \times \frac{(\sigma(j+\delta)(\frac{1}{\mu})^k)^{\frac{-1+\delta-k\delta}{k}} \Gamma[\frac{1-\delta+k\delta}{k}]}{k} \right]$$

Where $b(\sigma,\ \mu,\ k,\ j,\ i\ ,\delta,\ c_i) = \frac{\pi^{\delta} \sigma^{\delta} k^{\delta}}{2^{\delta} \mu^{\delta k}} \sum_{i=0}^{\infty} \sum_{j=0}^{2i} c_i \binom{2i}{j} (-1)^j$ *and*
$c_0 = a_0^{\delta}$, $c_m = [1/(ma_0)] \sum_{i=1}^{m} (i\delta - m + i) a_i c_{m-i}$ *for* $m \geq 1$.

*Proof.*Suppose X has the PDF in Eq.(6) Then, one can calculate

$$[f(x)]^{\delta} = \frac{\pi^{\delta} \sigma^{\delta} k^{\delta} x^{\delta(k-1)}}{2^{\delta} \mu^{\delta k}} e^{\left(-\delta\sigma\left(\frac{x}{\mu}\right)^k\right)} \left(\cos\left[\frac{\pi}{2}[1 - e^{\left(-\sigma\left(\frac{x}{\mu}\right)^k\right)}]\right]\right)^{\delta}$$

By inserting the expansion cos[z(x)], we have

$$[f(x)]^{\delta} = \frac{\pi^{\delta} \sigma^{\delta} k^{\delta} x^{\delta(k-1)}}{2^{\delta} \mu^{\delta k}} e^{\left(-\delta\sigma\left(\frac{x}{\mu}\right)^k\right)} \left(\sum_{i=0}^{\infty} \frac{(-1)^i}{(2i)!} \left(\frac{\pi}{2}\right)^{2i} \left[1 - e^{\left(-\sigma\left(\frac{x}{\mu}\right)^k\right)}\right]^{2i}\right)^{\delta}$$

let $a_n = \frac{(-1)^i}{(2i)!}\left(\frac{\pi}{2}\right)^{2i}$, $G(x;\xi) = \left[1 - e^{\left(-\sigma\left(\frac{x}{\mu}\right)^k\right)}\right]$ and Eq.(15) we obtained:

$$\left(\sum_{i=0}^{\infty} a_i G(x;\xi)^{2i}\right)^{\delta} = \sum_{i=0}^{\infty} c_i G(x;\xi)^{2i} = \sum_{i=0}^{\infty} c_i \left[1 - e^{\left(-\sigma\left(\frac{x}{\mu}\right)^k\right)}\right]^{2i} \tag{16}$$

where $c_0 = a_0^{\delta}$, $c_m = [1/(ma_0)]\sum_{i=1}^{m}(i\delta - m + i)a_i c_{m-i}$ for $m \geq 1$.

Then by Eq.(16) and Eq.(10), we have

$$[f(x)]^{\delta} = \frac{\pi^{\delta}\sigma^{\delta}k^{\delta}x^{\delta(k-1)}}{2^{\delta}\mu^{\delta k}} e^{\left(-\delta\sigma\left(\frac{x}{\mu}\right)^k\right)} \left(\sum_{i=0}^{\infty} c_i \sum_{j=0}^{2i}\binom{2i}{j}(-1)^j e^{\left(-\sigma j\left(\frac{x}{\mu}\right)^k\right)}\right)$$

$$= \frac{\pi^{\delta}\sigma^{\delta}k^{\delta}x^{\delta(k-1)}}{2^{\delta}\mu^{\delta k}} e^{\left(-\delta\sigma\left(\frac{x}{\mu}\right)^k\right)} \left(\sum_{i=0}^{\infty}\sum_{j=0}^{2i} c_i \binom{2i}{j}(-1)^j e^{\left(-\sigma j\left(\frac{x}{\mu}\right)^k\right)}\right)$$

Let b(σ, μ, k, j, i ,δ, c_i) = $\frac{\pi^{\delta}\sigma^{\delta}k^{\delta}}{2^{\delta}\mu^{\delta k}}\sum_{i=0}^{\infty}\sum_{j=0}^{2i} c_i\binom{2i}{j}(-1)^j$, then

$$[f(x)]^{\delta} = b(\sigma,\mu,k,j,i,\delta,c_i)x^{\delta(k-1)}e^{\left(-\sigma(\delta+j)\left(\frac{x}{\mu}\right)^k\right)} \tag{17}$$

To find $H_{\delta}(x)$, we substitute from (17) in (14)

$$H_{\delta}(x) = \frac{1}{1-\delta} log\left[b(\sigma,\mu,k,j,i,\delta,c_i) \times \int_0^{\infty} x^{\delta(k-1)} e^{\left(-\sigma(\delta+j)\left(\frac{x}{\mu}\right)^k\right)} dx\right]$$

We obtain the R'enyi entropy of the S-EP distribution after solving the integral.

$$H_{\delta}(x) = \frac{1}{1-\delta} log\left[b(\sigma,\mu,k,j,i,\delta,c_i) \times \frac{\left(\sigma(j+\delta)\left(\frac{1}{\mu}\right)^k\right)^{\frac{-1+\delta-k\delta}{k}} \Gamma\left[\frac{1-\delta+k\delta}{k}\right]}{k}\right]$$

4.2 q-Entropy

The q-entropy was introduced by[13]. It is the one-parameter generalization of the Shannon entropy[30] defined the q-entropy as

$$I_H(q) = \frac{1}{1-q}\left[1 - \int_0^{\infty} f(x)^q dx\right], where\ q > 0, and\ q \neq 1 \tag{18}$$

Theorem 5. *The q-entropy of a random variable X ∼S-EP distribution (μ,σ,k), with $\mu,\sigma,k > 0$ is given by*

$$I_H(q) = \frac{1}{1-q}\left[1 - b(\sigma,\mu,k,j,i,q,c_i) \times \frac{\left(\sigma(j+q)\left(\frac{1}{\mu}\right)^k\right)^{\frac{-1+q-kq}{k}} \Gamma\left[\frac{1-q+kq}{k}\right]}{k}\right]$$

Proof. To find $I_H(q)$, we substitute(17) in (18).

4.3 Shannon's Entropy

The Shannon's entropy[27] of a non-negative continuous random variable X with PDF $f(x)$ is defined as

$$H(f) = E[-\log f(x)] = -\int_0^\infty f(x)\log(f(x))dx \tag{19}$$

Below, we are going to use the expansion of the logarithm function (Taylor series at 1),

$$\log(x) = \sum_{m=0}^{\infty}(-1)^{m-1}\frac{(x-1)^m}{m}, |x| < 1 \tag{20}$$

The Shannon entropy of a random variable $X \sim$ S-EP distribution (μ,σ,k), with $\mu,\sigma,k > 0$ is given by

$$H(f) = \sum_{m=1}^{\infty}\sum_{n=0}^{m}(-1)^n\binom{m}{n}\frac{1}{m}\,\mathrm{c}(\sigma,\ \mu,\mathrm{k},\ \mathrm{j},\ \mathrm{i}\ ,\mathrm{n}+1) \tag{21}$$

$$\times\frac{\left((1+j)(1+n)(\frac{1}{\mu})^k\sigma\right)^{-1+(-1+\frac{1}{k})n}\Gamma[1+n-\frac{n}{k}]}{k}$$

*Proof.*By the Expansion of the Logarithm function (20)

$$\log(f(x)) = \sum_{m=1}^{\infty}(-1)^{m-1}\frac{(f(x)-1)^m}{m},$$

and by Binomial theorem,

$$= \sum_{m=1}^{\infty}(-1)^{m-1}\frac{1}{m}\left\{\sum_{n=0}^{m}(-1)^{m-n}\binom{m}{n}f^n(x)\right\}$$

$$\log(f(x)) = \sum_{m=1}^{\infty}\sum_{n=0}^{m}(-1)^{n+1}\binom{m}{n}\frac{1}{m}f^n(x) \tag{22}$$

For calculating the Shannon's entropy of X, substitute from Eq.(22) in Eq.(19)

$$H(f) = -\int_0^\infty f(x)\log(f(x))\,dx = -\int_0^\infty f(x)\sum_{m=1}^{\infty}\sum_{n=0}^{m}(-1)^{n+1}\binom{m}{n}\frac{1}{m}f^n(x)dx$$

$$H(f) = \int_0^\infty\sum_{m=1}^{\infty}\sum_{n=0}^{m}(-1)^n\binom{m}{n}\frac{1}{m}f^{n+1}(x)dx \tag{23}$$

substituting from Eq.(17) in Eq.(23), to get

$$H(f) = \int_0^\infty\sum_{m=1}^{\infty}\sum_{n=0}^{m}(-1)^n\binom{m}{n}\frac{1}{m}b(\sigma,\mu,k,j,i,n+1,c_i)x^{(n+1)(k-1)}e^{\left(-\sigma(n+j+1)\left(\frac{x}{\mu}\right)^k\right)}dx$$

Then,

$$H(f) = \sum_{m=1}^{\infty}\sum_{n=0}^{m}(-1)^n\binom{m}{n}\frac{1}{m}b(\sigma,\mu,k,j,i,n+1,c_i)\frac{\left(\sigma(1+j+n)(\frac{1}{\mu})^k\right)^{-1+\left(-1+\frac{1}{k}\right)n}\Gamma[1+n-\frac{n}{k}]}{k}$$

5 Order Statistics

Let $X_1,X_2,...,X_n$ be a random sample of size n from the S-EP distribution with parameters $\mu > 0$, $\sigma > 0$, and $k > 0$, the PDF of the rth order statistics is obtain by

$$f_{X_{(r)}}(x) = n\binom{n-1}{r-1}f(x)[F(x)]^{r-1}[1-F(x)]^{n-r}. \tag{24}$$

Let X_r be the rth order statistic of $X \sim$ S-EP distribution (μ,σ,k) with $\mu \neq 0$, $\sigma \neq 0$ and k $\neq$0 Then the PDF of the rth order statistic is given by

$$f_{X_{(r)}}(x) = n\binom{n-1}{r-1}\frac{\pi\sigma k x^{k-1}}{2\mu^k}e^{\left(-\sigma\left(\frac{x}{\mu}\right)^k\right)}\cos\left[\frac{\pi}{2}[1-e^{\left(-\sigma\left(\frac{x}{\mu}\right)^k\right)}]\right]\times$$
$$\left(\sin\left[\frac{\pi}{2}[1-e^{\left(-\sigma\left(\frac{x}{\mu}\right)^k\right)}]\right]\right)^{r-1}\left(1-\sin\left[\frac{\pi}{2}[1-e^{\left(-\sigma\left(\frac{x}{\mu}\right)^k\right)}]\right]\right)^{n-r},$$

Then, the PDF of first-order statistic X_1 of S-EP distribution is given as

$$f_{X_{(1)}}(x) = n\frac{\pi\sigma k x^{k-1}}{2\mu^k}e^{\left(-\sigma\left(\frac{x}{\mu}\right)^k\right)}\cos\left[\frac{\pi}{2}[1-e^{\left(-\sigma\left(\frac{x}{\mu}\right)^k\right)}]\right]\times\left(1-\sin\left[\frac{\pi}{2}[1-e^{\left(-\sigma\left(\frac{x}{\mu}\right)^k\right)}]\right]\right)^{n-1}.$$

Therefore, of the largest order statistic $X_{(n)}$ of S-EP distribution is given by

$$f_{X_{(n)}}(x) = n\frac{\pi\sigma k x^{k-1}}{2\mu^k}e^{\left(-\sigma\left(\frac{x}{\mu}\right)^k\right)}\cos\left[\frac{\pi}{2}[1-e^{\left(-\sigma\left(\frac{x}{\mu}\right)^k\right)}]\right]\times\left(\sin\left[\frac{\pi}{2}[1-e^{\left(-\sigma\left(\frac{x}{\mu}\right)^k\right)}]\right]\right)^{n-1}.$$

6 Maximum Likelihood Estimation (MLE)

Assume $X_1, X_2, ..., X_n$ be a random sample of size n from S-EP distribution and Δ = (μ,σ,k)^T be the parameter vector, then the likelihood function can be written as

$$L(\Delta) = \prod_{i=1}^{n} f(x_i) = \frac{\pi^n\sigma^n k^n}{2^n\mu^{nk}}\prod_{i=1}^{n}x_i^{k-1}\, e^{\left(-\sum_{i=1}^{n}\sigma\left(\frac{x_i}{\mu}\right)^k\right)}\prod_{i=1}^{n}\cos\left[\frac{\pi}{2}[1-e^{\left(-\sigma\left(\frac{x_i}{\mu}\right)^k\right)}]\right].$$

The log-likelihood function is given by

$$l(\Delta) = \log L(\Delta),$$
$$= n\log\left(\frac{\pi}{2}\right) + n\log(\sigma) + n\log(k) - nk\log(\mu) + (k-1)\sum_{i=1}^{n}\log(x_i) - \sum_{i=1}^{n}\sigma\left(\frac{x_i}{\mu}\right)^k$$
$$+\sum_{i=1}^{n}\log\left(\cos\left[\frac{\pi}{2}[1-e^{\left(-\sigma\left(\frac{x_i}{\mu}\right)^k\right)}]\right]\right). \tag{25}$$

Equation (25) can be directly maximized using the R (optim function), SAS (PROC NLMIXED), Ox program (sub-routine MaxBFGS), or by solving the nonlinear likelihood equations derived by differentiating Eq.(25). The log-likelihood function in Eq.(25) is differentiated with respect to each parameter to obtain the score function, $U(\Delta) = (\frac{\partial l(\Delta)}{\partial\sigma}, \frac{\partial l(\Delta)}{\partial\mu}, \frac{\partial l(\Delta)}{\partial k})^T$.

$$\frac{\partial l(\Delta)}{\partial\mu} = \frac{-nk}{\mu} + \frac{k\sigma}{\mu}\sum_{i=1}^{n}\left(\frac{x_i}{\mu}\right)^k + \frac{\pi k\sigma}{2\mu}\sum_{i=1}^{n}\left(\frac{x_i}{\mu}\right)^k e^{\left(-\sigma\left(\frac{x_i}{\mu}\right)^k\right)}\tan[\frac{\pi}{2}(1-e^{\left(-\sigma\left(\frac{x_i}{\mu}\right)^k\right)})],$$

$$\frac{\partial l(\Delta)}{\partial\sigma} = \frac{n}{\sigma} - \sum_{i=1}^{n}\left(\frac{x_i}{\mu}\right)^k - \frac{\pi}{2}\sum_{i=1}^{n}\left(\frac{x_i}{\mu}\right)^k e^{\left(-\sigma\left(\frac{x_i}{\mu}\right)^k\right)}\tan[\frac{\pi}{2}(1-e^{\left(-\sigma\left(\frac{x_i}{\mu}\right)^k\right)})],$$

$$\frac{\partial l(\Delta)}{\partial k} = n\left(\frac{1}{k} - \log(\mu)\right) + \sum_{i=1}^{n}\log(x_i) - \sum_{i=1}^{n}\sigma\left(\frac{x_i}{\mu}\right)^k\log\left(\frac{x_i}{\mu}\right) - \frac{\pi\sigma}{2}\sum_{i=1}^{n}\left(\frac{x_i}{\mu}\right)^k e^{\left(-\sigma\left(\frac{x_i}{\mu}\right)^k\right)}$$
$$\times\log\left(\frac{x_i}{\mu}\right)\tan[\frac{\pi}{2}(1-e^{\left(-\sigma\left(\frac{x_i}{\mu}\right)^k\right)})].$$

The maximum likelihood estimators (MLEs) are the solution of the nonlinear equations denoted by $\hat{\Delta}$ is obtained by solving the nonlinear equation $(\frac{\partial l(\Delta)}{\partial \sigma}, \frac{\partial l(\Delta)}{\partial \mu}, \frac{\partial l(\Delta)}{\partial k})^T = 0$, using a numerical method such as Newton-Raphson procedure. The Fisher information matrix is given by $I(\Delta) = [I_{\theta_i,\theta_j}]_{3\times 3} = E(-\frac{\partial^2 l}{\partial \theta_i \partial \theta_j}), i,j = 1,2,3$. The total Fisher information matrix nI(Δ) can be approximated by

$$J_n(\hat{\Delta}) \approx \left[-\frac{\partial^2 l}{\partial \theta_i \partial \theta_j}|_{\Delta=\hat{\Delta}}\right]_{3\times 3}, \qquad i,j = 1,2,3. \tag{26}$$

The matrix given in Eq.(26) is obtained after the Newton-Raphson process has converged for a given set of observations. Expectations of Fisher's information matrix (FIM) can be obtained numerically. Let $\hat{\Delta} = (\hat{\sigma}, \hat{\mu}, \hat{k})$ be the maximum likelihood estimate of $\Delta = (\sigma,\mu,$k). Under the normal regularity conditions and that the parameters are in the interior of the parameter space, but not on the boundary, we have:$\sqrt{n}(\hat{\Delta} - \Delta) \xrightarrow{d} N_3(\underline{0}, I^{-1}(\Delta))$, where I($\Delta$) is the expected Fisher information matrix. The asymptotic behavior is still valid if I(Δ) is replaced by the observed information matrix evaluated at $\hat{\Delta}$, that is $J(\hat{\Delta})$. The multivariate normal distribution $N_3(\underline{0}, J^{-1}(\Delta))$, where the mean vector $\underline{0} = (0,0,0)^T$, and $J^{-1}(\hat{\Delta})$ is the observed Fisher information matrix evaluated at $\hat{\Delta}$, can be used to construct confidence intervals and confidence regions for the individual model parameters and for the survival and hazard rate functions. That is, the approximate $100(1-\varphi)\%$ two-sided confidence intervals for σ, μ and k are $\hat{\sigma} \pm Z_{\frac{\varphi}{2}} \sqrt{I^{-1}_{\sigma\sigma}(\hat{\Delta})}$, $\hat{\mu} \pm Z_{\frac{\varphi}{2}} \sqrt{I^{-1}_{\mu\mu}(\hat{\Delta})}$, and $\hat{k} \pm Z_{\frac{\varphi}{2}} \sqrt{I^{-1}_{kk}(\hat{\Delta})}$, respectively, where $I^{-1}_{\sigma\sigma}(\hat{\Delta})$, $I^{-1}_{\mu\mu}(\hat{\Delta})$ and $I^{-1}_{kk}(\hat{\Delta})$, are the diagonal elements of $I_n^{-1}(\hat{\Delta}) = (nI(\hat{\Delta}))^{-1}$, and $Z_{\frac{\varphi}{2}}$ is the upper $(\frac{\varphi}{2})^{th}$ percentile of a standard normal distribution.

7 Simulation Study

This section will present simulation results for n = 30, 50, 100, 300, 400, 600, and 800 samples to assess the accuracy and consistency of MLEs for each parameter of the distribution of S-EP using Monte Carlo simulation. The R programming language with the argument method "L-BFGS-B" is used to run simulations. The actual values of the parameters as follows: (σ,μ,k) = (0.2, 0.5, 3), (0.7, 0.5, 4),(0.1,3.5,0.5) and (1.5, 3, 0.5). Repeated simulations of N = 1000 times were performed, evaluating the mean estimates, average bias (Abias) and root mean squared error (RMSE). For consistent MLEs, it is expected that as the sample size n increases, the mean estimates gets closer to the true parameters and, the RSMEs and Abias also decay to zero. Tables 3 and 4 show the mean estimates together with their respective RSMEs and Abias. The Abias and RMSEs for the estimated parameter, say, $\widehat{\theta}$ are respectively given as:

$$\text{Abias}\left(\widehat{\theta}\right) = \frac{\sum_{i=1}^{N} \widehat{\theta}_i}{N} - \theta \text{ and RMSE } \left(\widehat{\theta}\right) = \sqrt{\frac{\sum_{i=1}^{N} \left(\widehat{\theta}_i - \theta\right)^2}{N}}$$

From the results in Table 3 and 4 we conclude that estimation method is adequate as the simulated estimates are closed to the true values of parameters. We also observed that estimated root mean square errors (RMSEs) consistently decreases with increasing sample size and the average bias decreases as the sample size n increases.

8 Applications

Four real data applications are shown in this section to demonstrate the flexibility of S-EP distribution compared with other existing distributions.The new S-EP distribution was compared to Exponential Pareto (EP) distribution, New Weibull-Pareto (NWP) distribution, sine Exponential (S-E) distribution, Exponential (E) distribution, Sine Half-Logistic Inverse Rayleigh (S-HLIR) distribution, Sine Lindley (S-L) distribution and Sine Modified Lindley (S-ML) distribution. With the following PDF
The Exponential Pareto (EP) distribution

$$f(x;\sigma,\mu) = \frac{\sigma k x^{k-1}}{\mu^k} e^{\left(-\sigma\left(\frac{x}{\mu}\right)^k\right)}$$

The new Weibull-Pareto (NWP) distribution [24]

		(0.2,0.5,3)			(0.7,0.5,4)		
Parameter	Size	Mean	RMSE	Abias	Mean	RMSE	Abias
σ	30	0.2450911	0.07167169	0.045091062	1.0340214	0.6055962	0.3340215
	50	0.2361332	0.05872181	0.036133179	0.9583074	0.4510855	0.2583074
	100	0.2261253	0.04010630	0.026125342	0.8888830	0.3319789	0.1888830
	300	0.2163265	0.02614697	0.016326468	0.8153148	0.1982788	0.1153148
	400	0.2157181	0.02350279	0.015718108	0.8005841	0.1716874	0.1005841
	600	0.2122519	0.01881200	0.012251862	0.7817111	0.1429706	0.0817111
	800	0.2115609	0.01710020	0.011560870	0.7761191	0.1351243	0.0761191
μ	30	0.5343876	0.05558445	0.034387641	0.5357069	0.0520179	0.0357069
	50	0.5272853	0.04390810	0.027285319	0.5311114	0.0460026	0.0311114
	100	0.5212184	0.03302626	0.021218416	0.5249413	0.0367815	0.0249413
	300	0.5141166	0.02297605	0.014116641	0.5174863	0.0261347	0.0174863
	400	0.5127097	0.02049589	0.012709730	0.5154899	0.0233200	0.0154899
	600	0.5095906	0.01631827	0.009590613	0.5126437	0.0198133	0.0126437
	800	0.5091206	0.01495196	0.009120570	0.5121953	0.0191907	0.0121953
k	30	3.1496487	0.49504710	0.149648700	4.1879039	0.6708503	0.1879039
	50	3.0796932	0.35328182	0.079693187	4.0983997	0.5024887	0.0983997
	100	3.0527250	0.24778273	0.052725000	4.0599054	0.3258629	0.0599054
	300	3.0200105	0.14097969	0.020010546	4.0233247	0.1849521	0.0233247
	400	3.0126430	0.12143441	0.012642977	4.0051984	0.1609323	0.0051984
	600	3.0052480	0.09701817	0.005247959	4.0111684	0.1314587	0.0111684
	800	3.0051492	0.08891517	0.005149168	4.0016558	0.1199355	0.0016558

Table 3: simulation results from the S-EP distribution

		(0.1,3.5,0.5)			(1.5, 3, 0.5)		
Parameter	Size	Mean	RMSE	Abias	Mean	RMSE	Abias
σ	30	0.1045513	0.03777537	0.00455127	1.6286118	0.3552606	0.1286118
	50	0.1051240	0.03051245	0.00512396	1.5847376	0.2446279	0.0847376
	100	0.1026655	0.01978007	0.00266553	1.5540661	0.1700121	0.0540660
	300	0.1022413	0.01279401	0.00224129	1.5181415	0.0882373	0.0181415
	400	0.1022445	0.01071934	0.00224449	1.5252732	0.0783169	0.0252732
	600	0.1017690	0.00850948	0.00176901	1.5140017	0.0634229	0.0140017
	800	0.1015443	0.00786234	0.00154429	1.5088827	0.0641809	0.0088827
μ	30	3.4982486	0.06700761	-0.0017514	2.9812699	0.1333470	-0.018730
	50	3.4954476	0.08147722	-0.0045524	2.9928362	0.1011073	-0.007164
	100	3.5004942	0.08877552	0.0004942	3.0010284	0.0858786	0.0010284
	300	3.4990742	0.07276937	-0.0009258	3.0037909	0.0487038	0.0037909
	400	3.5052354	0.08164894	0.00523539	3.0018513	0.0355535	0.0018513
	600	3.5063684	0.07444357	0.00636839	3.0102425	0.0604417	0.0102425
	800	3.5071606	0.07377153	0.00716059	3.0029058	0.0261768	0.0029058
k	30	0.5213434	0.07914699	0.02134342	0.5396932	0.087118	0.0396932
	50	0.5075420	0.05780914	0.00754201	0.5268334	0.065129	0.0268334
	100	0.5015904	0.03759643	0.00159044	0.5196839	0.043301	0.0196839
	300	0.4955081	0.02049841	0.00449191	0.5142159	0.025631	0.0142159
	400	0.4943441	0.01783521	-0.00565594	0.5100009	0.022199	0.0100009
	600	0.4958552	0.01391259	-0.00414477	0.5108516	0.0201986	0.0108517
	800	0.4954825	0.01297103	-0.00451749	0.5124547	0.0172848	0.0124547

Table 4: simulation results from the S-EP distribution

$$f(x;\beta,\mu,\alpha) = \frac{\beta\mu}{\alpha}(\frac{x}{\alpha})^{\beta-1}exp(-\mu(\frac{x}{\alpha})^{\beta}$$

The Sine Exponential (S-E) distribution

$$f(x;\sigma) = \frac{\pi}{2}\sigma\, e^{-\sigma x}\sin\left[\frac{\pi}{2}e^{-\sigma x}\right]$$

The Exponential (E) distribution

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

Sine Half-Logistic Inverse Rayleigh (S-HLIR) distribution[28]

$$f(x;\mu,\alpha) = \frac{2\pi\mu\alpha^2 x^{-3} e^{-(\frac{\alpha}{x})^2}[1-e^{-(\frac{\alpha}{x})^2}]^{\mu-1}}{\left(1+\left[1-e^{-(\frac{\alpha}{x})^2}\right]^{\mu}\right)^2} cos\left[\frac{\pi}{2}\left(\frac{1-\left[1-e^{-(\frac{\alpha}{x})^2}\right]^{\mu}}{1+\left[1-e^{-(\frac{\alpha}{x})^2}\right]^{\mu}}\right)\right], \quad x,\mu,\alpha > 0$$

Sine Lindley (S-L) distribution [19]

$$f(x;\delta) = \frac{\pi}{2}\frac{\delta^2}{1+\delta}(1+x)e^{-\delta x}\sin\left(\frac{\pi}{2}\left(1+\frac{x\delta}{1+\delta}\right)e^{-\delta x}\right), \quad x,\delta > 0$$

Sine Modified Lindley (S-ML) distribution [29]

$$f(x;\delta) = \frac{\pi}{2}\frac{\delta}{1+\delta}e^{-2\delta x}[(1+\delta)e^{\delta x}+2x\delta-1]\sin\left(\frac{\pi}{2}\left(1+e^{-\delta x}\frac{x\delta}{1+\delta}\right)e^{-\delta x}\right), \quad x,\delta > 0$$

R software was used to execute model parameter estimations and goodness-of-fit measures, and the results were used to compare the S-EP distribution with other existing models. Tables 5, 7, 9 and 11 provide an overview of descriptive statistics measures, and Figures 4, 9, 14 and 19 show box and TTT plots.The box plot gives a visual representation of the descriptive measures of the data and the TTT plot, proved useful for gaining information about the hazard form of the data. In many real-world situations, there is qualitative information about the shape of the failure rate function that might help in the selection of a particular distribution. The TTT plot has a convex shape for decreasing HRF and a concave shape for increasing HRF. Fitted densities and empirical CDF plots in Figures 5, 10, 15 and 20. Figures 6, 11, 16 and 21 show Kaplan-Meier and HRF plot. And probability plot in Figures 8, 13, 18 and 23. The estimated values of the model parameters along with the corresponding standard errors (SE) and the goodness of fit (GOF) of the S-E, EP, E, S-HLIR, S-PL and S-ML, have been introduced using different comparison measures as we consider some criteria including Bayesian information criterion (BIC), Akaike Information criterion (AIC), along with Anderson Darling statistic (A^*), Cramér-von Mises statistic (w^*) and Kolmgrov–Smirnov statistic (D_n) with its correspondig p-value for the blood cancer, 1.5 cm glass fibres dataset , Bladder cancer dataset and the polyester fibers dataset in tables 6, 8, 10 and 12.

Data Set 1: Blood cancer dataset

The life time (in years) of 40 patients with blood cancer (leukemia) from one of the Saudi Arabian Ministry of Health institutions is what makes up this data.[17]. This actual data are: 0.315, 0.496, 0.616, 1.145,1.208, 1.263, 1.414, 2.025, 2.036, 2.162, 2.211, 2.370, 2.532, 2.693, 2.805, 2.910, 2.912, 3.192, 3.263, 3.348, 3.348, 3.427, 3.499, 3.534, 3.767, 3.751, 3.858, 3.986, 4.049, 4.244, 4.323, 4.381, 4.392, 4.397, 4.647, 4.753, 4.929, 4.973, 5.074, 5.381.

Figures 5, 6 and 8 illustrates how best the S-EP distribution fits the blood cancer data and Figure 7(a), 7(b) and 7(c) illustrates profile plots of the MLEs of σ, μ and k. It can be observed that the parameters for the blood cancer data attained their absolute maximum. The fitted density shows that S-EP distribution can accommodate skewed data. The estimated variance-covariance matrix for S-EP model in blood cancer data is given by

$$\begin{bmatrix} 6.51\times10^{-4} & -0.0076 & -7.53\times10^{-5} \\ -7.61\times10^{-3} & 0.10692 & 8.79\times10^{-4} \\ -7.53\times10^{-5} & 0.00088 & 8.71\times10^{-6} \end{bmatrix}$$

and the 95% confidence intervals for the model parameters are given by $\sigma \in [\,0.0718 \mp 0.050]$, $\mu \in [\,1.491 \mp 0.00578]$ and $k \in [2.398 \mp 0.64089]$

mean	Median	Skewness	kurtosis
3.141	3.348	-0.4331	2.586

Table 5: Descriptive statistics of the blood cancer dataset.

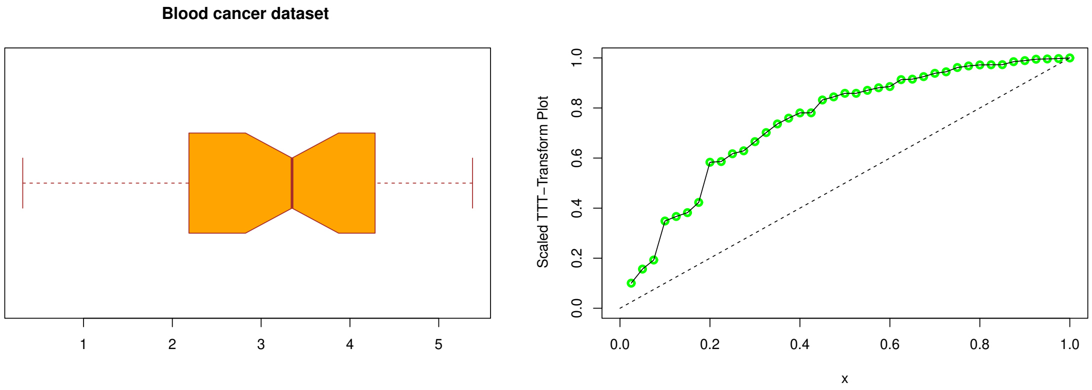

Fig. 4: Box plot and TTT plot for the blood cancer dataset.

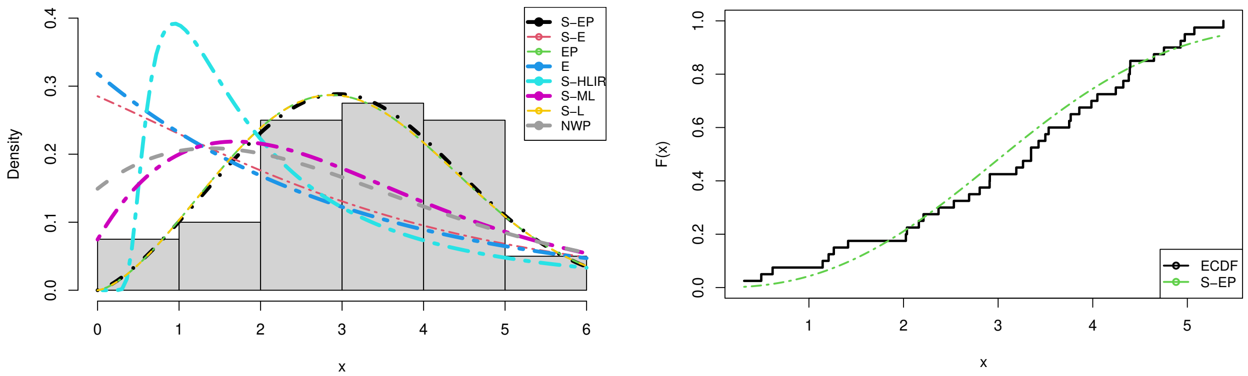

Fig. 5: Fitted densities and empirical CDF plots for the blood cancer data set.

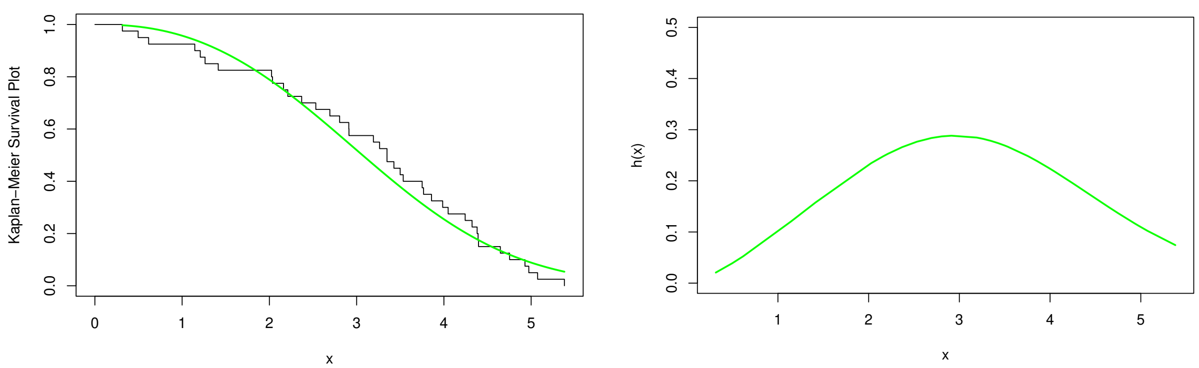

Fig. 6: Kaplan-Meierand and hrf plot for the blood cancer data set.

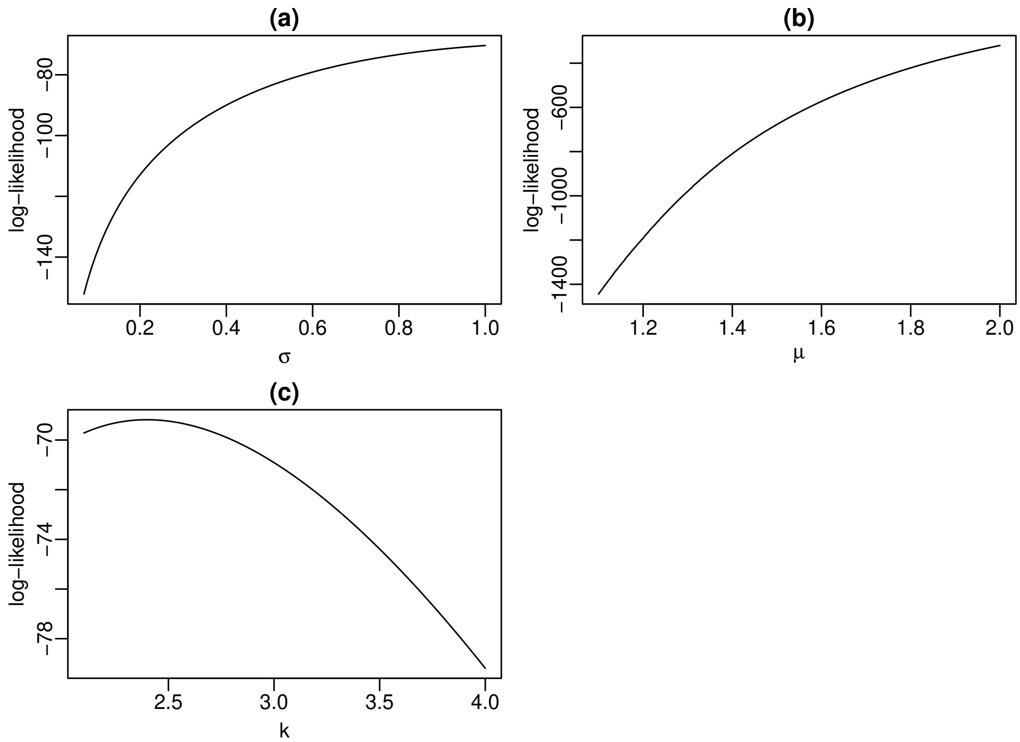

Fig. 7: Profile log-likelihood functions for the blood cancer data (a–c).

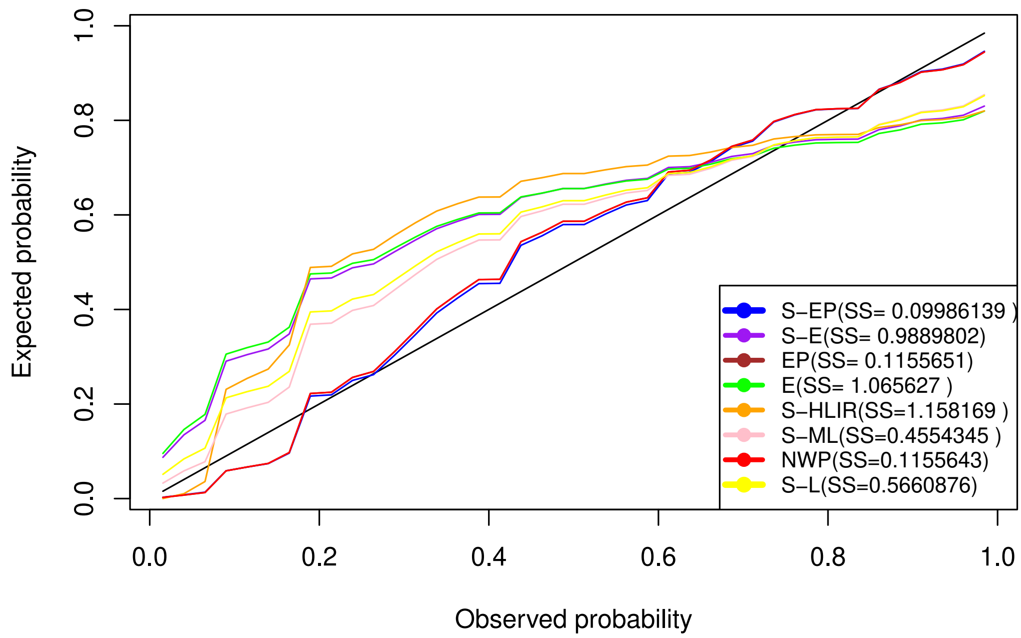

Fig. 8: Probability plot for the blood cancer dataset.

Distribution	Parameter	Estimate	SE	-2logl	AIC	BIC	A*	W*	D_n	P-Value
S-EP	σ	0.07187	0.02552	138.37	144.4	149.43	0.705	0.1072	0.1107	0.7112
	k	2.39819	0.32699							
	μ	1.49089	0.00295							
S-E	σ	0.18156	0.02652	167.89	169.9	171.58	1.407	0.2276	0.2897	0.0024
EP	σ	0.42244	0.08103	139.12	145.1	150.2	0.773	0.1187	0.1184	0.629
	k	2.49949	0.33706							
	μ	2.49246	0.03433							
E	σ	0.31839	0.05034	171.56	173.56	175.25	1.496	0.2434	0.3002	.00148
NWP	β	2.4994896	0.3370598	139.12	145.12	150.18	0.7729	0.1187	0.1184	0.6291
	μ	0.0791531	0.0323587							
	α	1.2754098	0.0050198							
S-HLIR	μ	0.40112	0.06296	183.73	187.7	191.1	4.227	0.7629	0.3138	0.0008
	α	0.91847	0.13216							
S-L	δ	0.359289	0.035991	155.165	157.17	158.85	1.1451	0.18193	0.2198	0.04199
S-ML	δ	0.2428	0.0253	152.63	154.6	156.32	1.249	0.1998	0.1938	0.0993

Table 6: MLE's of the parameters, SE and GOF metrics of the blood cancer dataset.

The TTT plot of the blood cancer dataset is displayed in Figure 4. It shows an increasing HRF plot. Moreover, analysis of the data set implies that the (S-EP) distribution is the best model among the other competitive models, when statistical GOF criteria and the increasing HRF are considered. From the results in Table 6, S-EP model performed better than any other model. It had the lowest values for -2logl, AIC, BIC, A^*, W^* and D_n, as well as the highest p-value when compared to competing models across for the blood cancer data.

Data Set 2: 1.5 cm glass fibres dataset
The 1.5 cm glass fibre strength data obtained by the workers of the National Physical Laboratory of the United Kingdom are now being used. They were previously analyzed by [20]. The data are: 0.55, 0.74, 0.77, 0.81, 0.84, 1.24, 0.93, 1.04, 1.11, 1.13, 1.30,1.25, 1.27, 1.28, 1.29, 1.48, 1.36, 1.39, 1.42, 1.48, 1.51, 1.49, 1.49, 1.50, 1.50, 1.55, 1.52, 1.53, 1.54,1.55, 1.61, 1.58, 1.59, 1.60, 1.61, 1.63, 1.61, 1.61, 1.62, 1.62, 1.67, 1.64, 1.66, 1.66, 1.66, 1.70, 1.68,1.68, 1.69, 1.70, 1.78, 1.73, 1.76, 1.76, 1.77, 1.89, 1.81, 1.82, 1.84, 1.84, 2.00, 2.01, 2.24.

Figures 10, 11 and 13 illustrates how best the S-EP distribution fits the 1.5 cm glass fibres data and Figure 12(a), 12(b) and 12(c) illustrates profile plots of the MLEs of σ, μ and k. It can be observed that the parameters for 1.5 cm glass fibres data reached absolute maximum. The fitted density shows that S-EP distribution can accommodate skewed data. The estimated variance-covariance matrix for S-EP model in 1.5 cm glass fibres data is given by

$$\begin{bmatrix} 4.499*10^{-5} & 0.003389 & -0.000935 \\ 3.389*10^{-3} & 0.310236 & -0.070431 \\ -9.35*10^{-4} & -0.070431 & 0.0194345 \end{bmatrix}$$

and the 95% confidence intervals for the model parameters are given by $\sigma \in [10.34 \mp 0.01315]$, $\mu \in [2.75 \mp 0.27324]$and $k \in [5.53 \mp 1.091698]$

mean	Median	Skewness	kurtosis
1.507	1.59	-0.89993	0.92376

Table 7: Descriptive statistics of 1.5 cm glass fibres dataset.

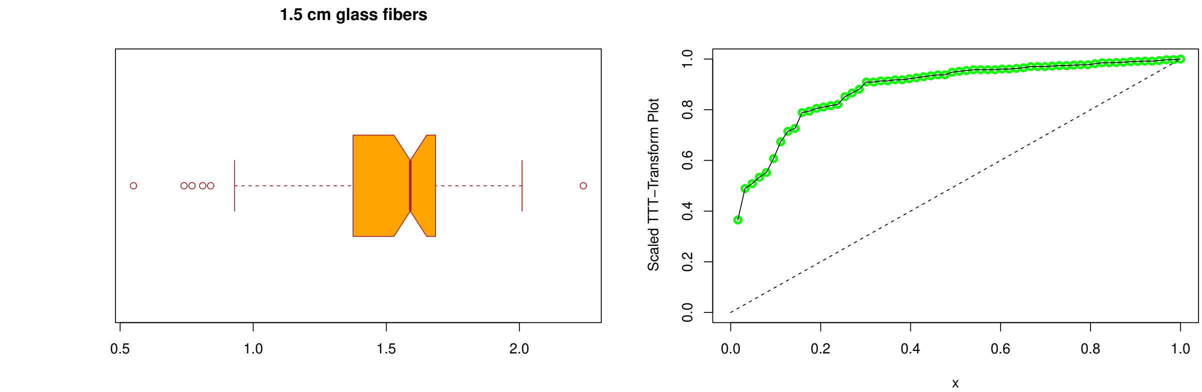

Fig. 9: Box plot and TTT plot for the 1.5 cm glass fibers dataset.

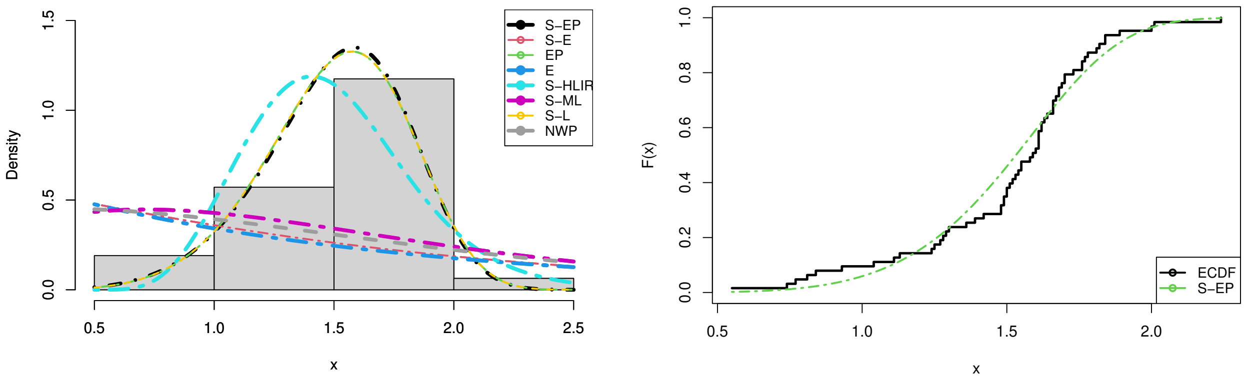

Fig. 10: Fitted densities and empirical CDF plots for the 1.5 cm glass fibres dataset.

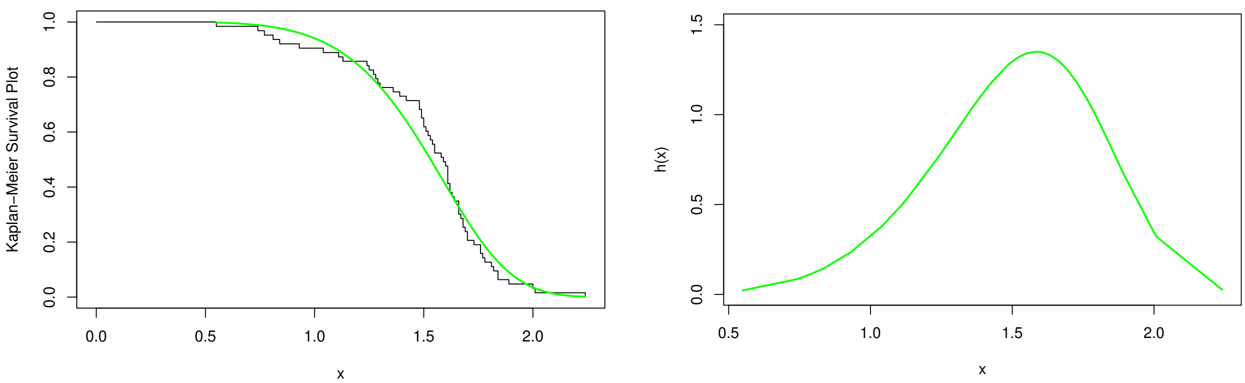

Fig. 11: Kaplan-Meier and hrf plot for the 1.5 cm glass fibres dataset.

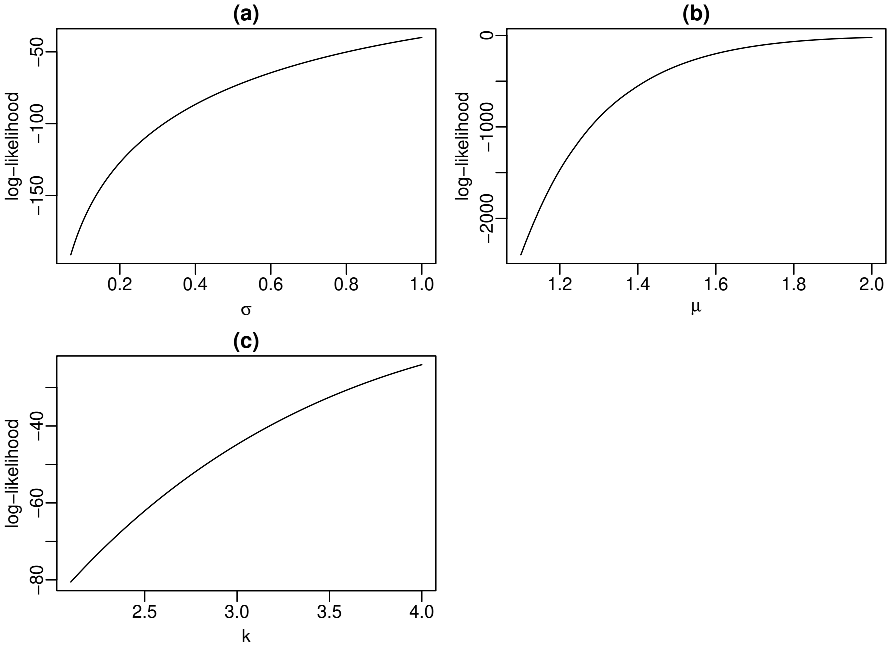

Fig. 12: Profile log-likelihood functions for the 1.5 cm glass fibres data (a–c).

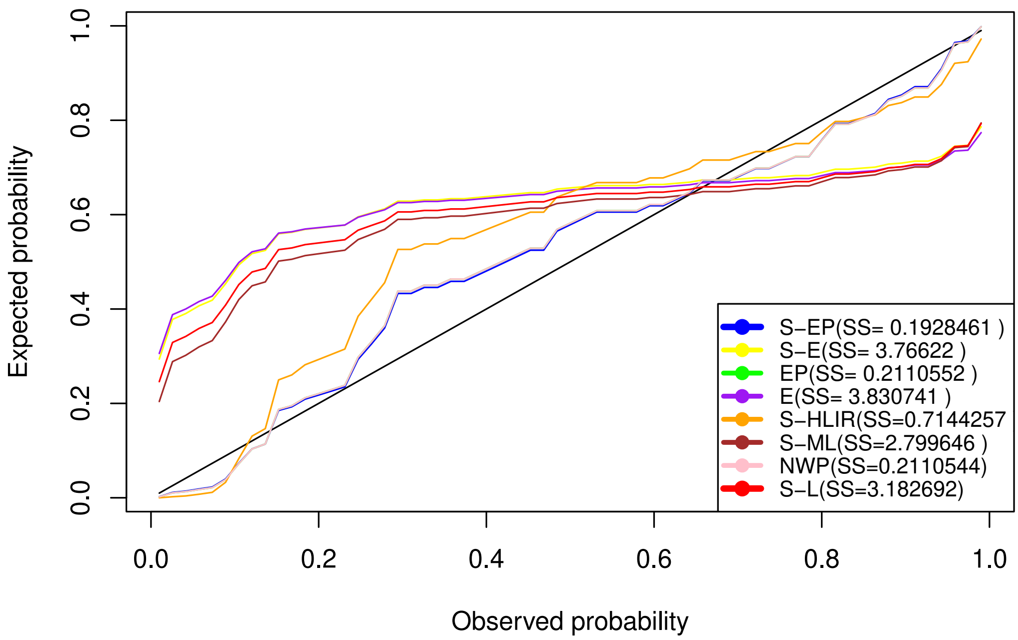

Fig. 13: Probability plot for the 1.5 cm glass fibres dataset.

Distribution	Parameter	Estimate	SE	-2logl	AIC	BIC	A*	W*	D_n	P-Value
S-EP	σ	10.34147	0.006708	29.404	35.40	41.834	1.197	0.2175	0.1473	0.1301
	k	5.529111	0.556989							
	μ	2.751241	0.139408							
S-E	σ	0.38293	0.04427	169.96	171.9	174.10	3.047	0.5555	0.4166	$6.3*10^{-10}$
EP	σ	2.01165	0.00636	30.414	36.41	42.84	1.304	0.2372	0.1522	0.1078
	k	5.78070	0.576094							
	μ	1.837368	0.040255							
E	σ	0.663657	0.083612	177.66	179.7	181.80	3.127	0.5702	0.418	$5.5*10^{-10}$
NWP	β	5.780701	0.576094	30.414	36.41	42.84	1.3037	0.2372	0.15224	0.1078
	μ	0.296058	0.022440							
	α	1.318981	0.029116							
S-HLIR	μ	4.17499	1.03164	55.035	59.03	63.321	3.576	0.6569	0.2406	0.00136
	α	1.99707	0.14092							
S-L	δ	0.678144	0.055365	151.21	153.21	155.36	2.837	0.5169	0.38305	$1.9*10^{-8}$
S-ML	δ	0.49009	0.04122	139.61	141.6	143.8	2.897	0.5276	0.3586	$1.8*10^{-7}$

Table 8: MLE's of the parameters, SE and GOF metrics measures of the 1.5 cm glass fibres dataset.

The TTT plot of the 1.5 cm glass fibres data set is displayed in Figure 9. It shows an increasing HRF plot. In addition, analysis of the data set shows that the evaluated model (S-EP) is the best model throughout all elements of the model selection criteria, such as the increasing hazard function. We can observe from Table 8, that the S-EP distribution has minimum values for the test statistics with a higher p-value and least values for GOF metrics.

Data Set 3: Bladder cancer dataset

In 36 cases of bladder cancer, a series of data on the duration (months) of remission was reported in [14] given by: 0.08, 0.2, 0.4, 0.5, 0.51, 0.81, 0.87, 0.9, 1.05, 1.19, 1.26, 1.35, 1.4, 1.46, 1.76, 2.02, 2.02,2.07, 2.09, 2.23, 2.26, 2.46, 2.54, 2.62, 2.64, 2.69, 2.69, 2.75, 2.83, 2.87, 3.02, 3.02, 3.25, 3.31, 3.36, 3.36.

Figures 15, 16 and 18 illustrates how best the S-EP distribution fits the blood cancer dataand Figure 17(a), 17(b) and 17(c) illustrates profile plots of the MLEs of σ, μ and k. It is clear that the parameters have reached their absolute maximum for the bladder cancer data. The fitted density shows that S-EP distribution can accommodate skewed data. The estimated variance-covariance matrix for S-EP model in blood cancer data is given by

$$\begin{bmatrix} 0.0032542 & -0.010226 & -0.001212 \\ -0.010226 & 0.0753008 & 0.0038089 \\ -0.001212 & 0.0038089 & 0.0004514 \end{bmatrix}$$

and the 95% confidence intervals for the model parameters are given by $\sigma \in [0.3151 \mp 0.1118091]$**,** $\mu \in [1.5912 \mp 0.0416439]$and $k \in [1.8808 \mp 0.5378436]$

mean	**Median**	**Skewness**	**kurtosis**
1.94	2.08	-0.309	2.147

Table 9: Descriptive statistics of the bladder cancer dataset.

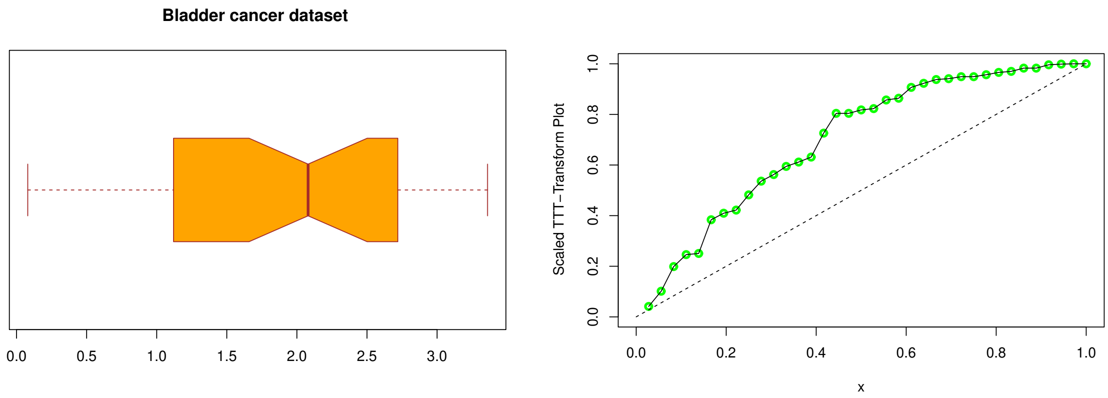

Fig. 14: Box plot and TTT plot for the bladder cancer dataset.

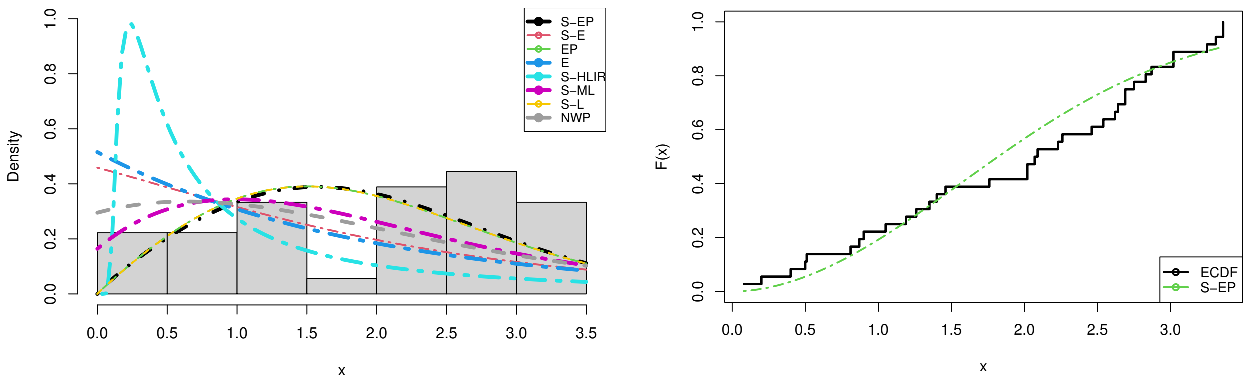

Fig. 15: Fitted densities and empirical CDF plots for the bladder cancer dataset.

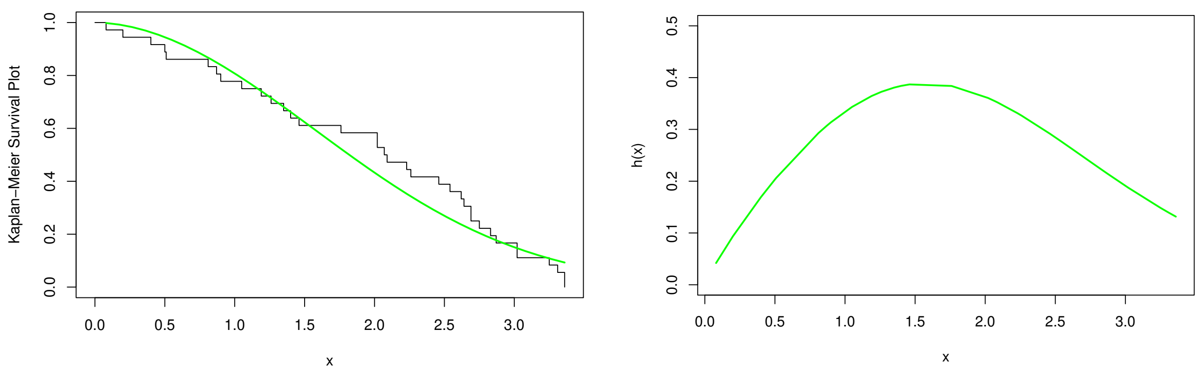

Fig. 16: Kaplan-Meier and hrf plot for the bladder cancer dataset.

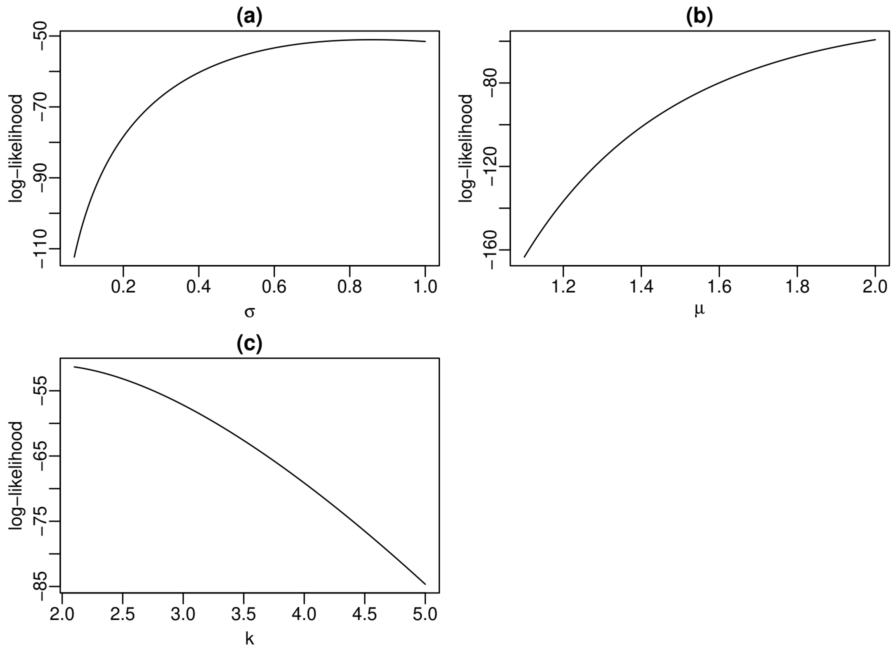

Fig. 17: Profile log-likelihood functions for the bladder cancer data (a–c).

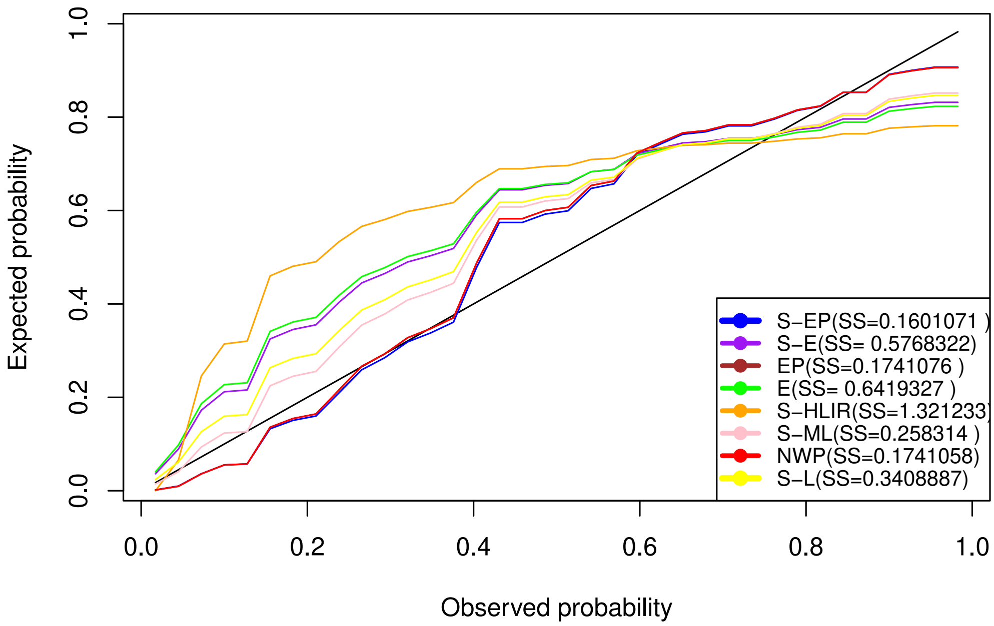

Fig. 18: Probability plot of the bladder cancer dataset.

Distribution	Parameter	Estimate	SE	-2logl	AIC	BIC	A*	W*	D_n	P-Value
S-EP	σ	0.315112	0.0571	102.09	108.09	112.84	0.915	0.1508	0.1578	0.332
	k	1.88087	0.2744							
	μ	1.591223	0.021247							
S-E	σ	0.29213	0.0451	116.98	118.98	120.56	1.301	0.2151	0.2277	0.0479
EP	σ	0.86694	0.0913	102.77	108.99	113.74	0.937	0.1545	0.1864	0.1637
	k	1.95698	0.2816							
	μ	2.01217	0.0769							
E	σ	0.51546	0.0859	119.71	121.71	123.29	1.371	0.2268	0.2303	0.0439
NWP	β	1.95698	0.281594	102.77	108.77	113.52	0.966	0.1592	0.1659	0.2752
	μ	0.19627	0.055135							
	α	0.941901	0.022483							
S-HLIR	μ	0.22064	0.04024	153.77	157.77	160.93	4.189	0.7492	0.3213	0.0012
	α	0.23836	0.03558							
S-L	δ	0.537571	0.057979	109.64	111.64	113.22	1.1135	0.1842	0.20116	0.1086
S-ML	δ	0.37974	0.04267	107.83	109.83	111.14	1.161	0.1928	0.1909	0.1448

Table 10: MLE's of the parameters, SE and GOF metrics measures of the of the bladder cancer dataset.

The TTT plot of the bladder cancer data set in Figure 14 displays increasing HRF that indicates the appropriateness of the S-EP distribution to fit the data sets. In Table 10, we compare the S-EP model with the S-E, EP,E, NWP, S-HLIR, S-L and S-ML distributions. Its noted that the proposed model has the lowest values for the AIC, A^*, W^* and D_n statistics among all fitted models (except BIC for the S-ML), as well as the highest p- value. So, the S-EP can be chosen as the best model among the competing distributions studied in this article.

Data Set 4: polyester fibers
The fourth data set includes 30 tensile strength measurements of polyester fibers, which has been discussed by [21]. The data are: 0.023, 0.032, 0.054, 0.069, 0.081, 0.094, 0.105, 0.127, 0.148, 0.169, 0.188, 0.216, 0.255, 0.277, 0.311, 0.361, 0.376, 0.395, 0.432, 0.463, 0.481, 0.519, 0.529, 0.567, 0.642, 0.674, 0.752, 0.823,0.887, and 0.926.

Figures 20, 21 and 23 illustrate how best the S-EP distribution fits the the polyester fibers datase data and Figure 22(a), 22(b) and 22(c) illustrates profile plots of the MLEs of σ, μ and k. It is clear that the parameters for polyester fibers data attained the absolute maximum. The fitted density shows that S-EP distribution can accommodate skewed data. The estimated variance-covariance matrix for S-EP model in the polyester fibers data is given by

$$\begin{bmatrix} 0.02582 & 0.02548 & -0.07127 \\ 0.02548 & 0.03663 & -0.07033 \\ -0.07127 & -0.07033 & 0.196757 \end{bmatrix}$$

and the 95% confidence intervals for the model parameters are given by $\sigma \in [4.512 \mp 0.3149]$, $\mu \in [2.062 \mp 0.8694]$and $k \in [1.262 \mp 0.3751]$

mean	**Median**	**Skewness**	**kurtosis**
0.3659	0.336	0.5756	2.651

Table 11: Descriptive statistics of the polyester fibers dataset.

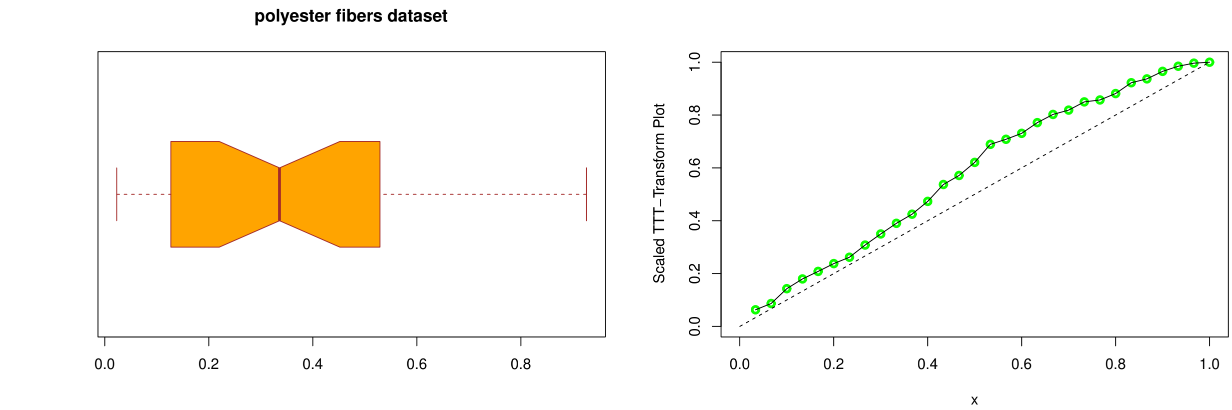

Fig. 19: Box plot and TTT plot of the polyester fibers dataset

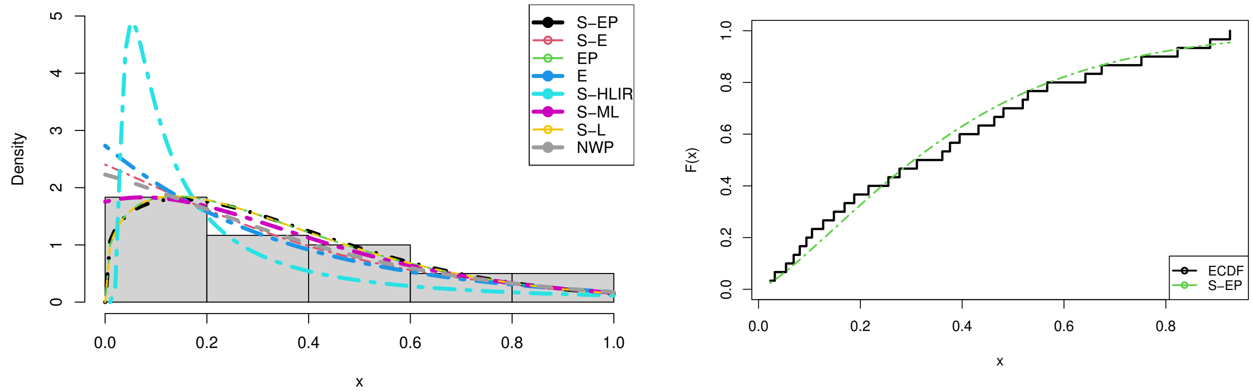

Fig. 20: Fitted densities and empirical CDF plots of the polyester fibers dataset.

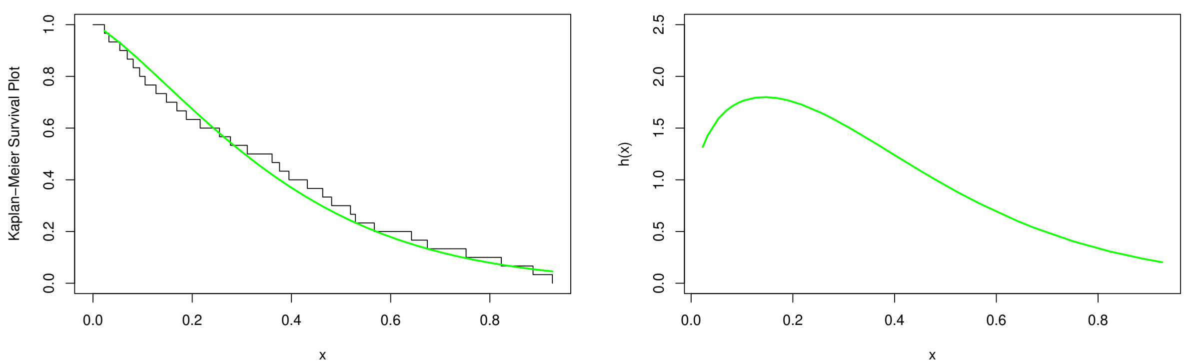

Fig. 21: Kaplan-Meier and hrf plot of the polyester fibers dataset.

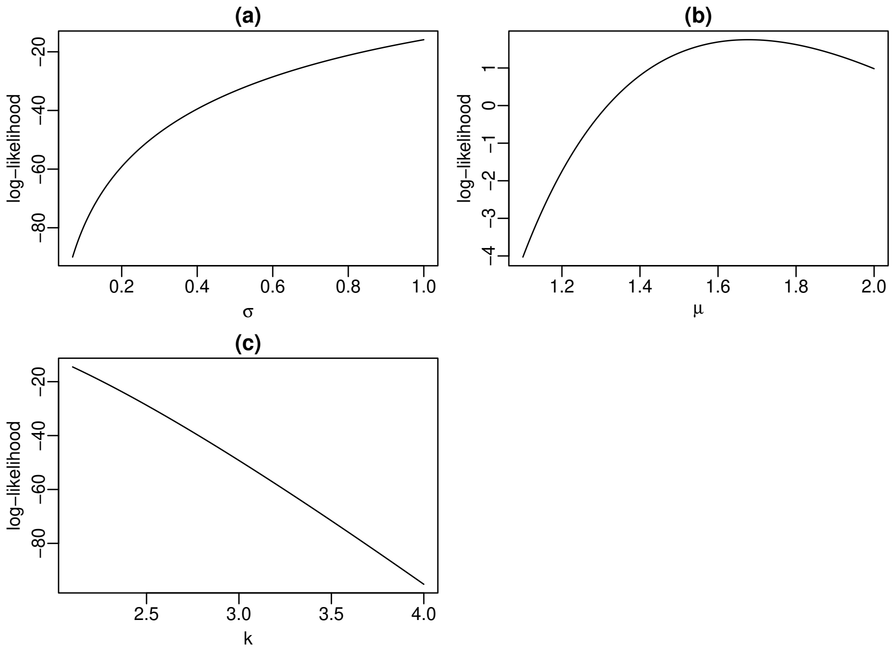

Fig. 22: Profile log-likelihood functions of the polyester fibers dataset (a–c).

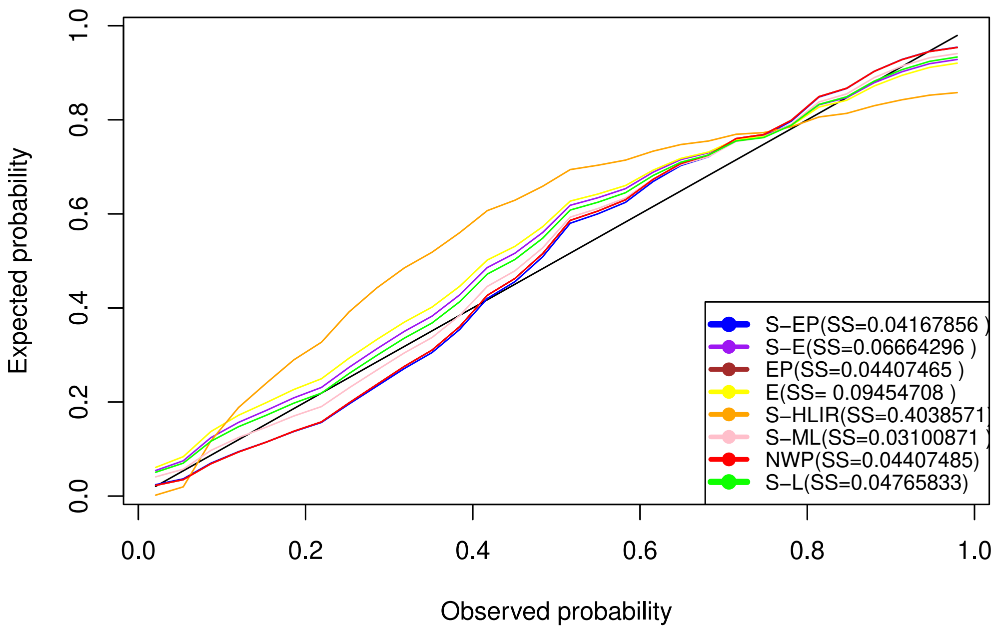

Fig. 23: Probability plot of the polyester fibers dataset.

Distribution	Parameter	Estimate	SE	-2logl	AIC	BIC	A*	W*	D_n	P-Value
S-EP	σ	4.51183	0.16068	-3.505	2.4955	6.6991	0.23056	0.03467	0.08011	0.9822
	k	1.26153	0.19140							
	μ	2.06181	0.44357							
S-E	σ	1.52980	0.26149	-1.371	0.62899	2.03019	0.26984	0.04188	0.1183	0.7516
EP	σ	7.707762	0.074957	-3.503	2.4969	6.7005	0.2358	0.03586	0.08619	0.965
	k	1.32536	0.19747							
	μ	1.85236	0.41338							
E	σ	2.73324	0.49902	-0.3292	1.6708	3.0720	0.2873	0.04507	0.12719	0.6701
NWP	β	1.32536	0.19747	-3.503	2.49685	6.7005	0.2358	0.03587	0.08619	0.965
	μ	0.89579	0.01530							
	α	0.36515	0.04975							
S-HLIR	μ	0.26899	0.04441	13.528	17.528	20.330	1.5162	0.24495	0.20688	0.1327
	α	0.04981	0.00943							
S-L	δ	2.0964	0.2828	-2.110	-0.1103	1.2909	0.24889	0.03813	0.10813	0.8376
S-ML	δ	1.7535	0.24548	-2.723	-0.7231	0.67806	0.2504	0.03793	0.09309	0.9358

Table 12: MLE’s of the parameters ,SE and GOF metrics measures of the polyester fibers dataset.

Figure 19 shows the TTT plot of this data set. It illustrates an increasing HRF plot. In Table 12, we compare the S-EP model with the S-E, EP,E, NWP, S-HLIR, S-L and S-ML distributions. Its noted that the proposed model has the lowest values for the A^*, W^* and D_n statistics among all fitted models (except AIC and BIC for the S-ML), as well as the highest p value. So, the S-EP can be chosen as the best model among the competing distributions studied in this article.

9 Concluding

In this paper, we proposed a new model, which is called the S-EP model. Some basic statistical properties of the S-EP model are studided such as moments,moment generating function, quantile functions, entropies and order statistics. The maximum likelihood method is used to estimate the distribution.The performance of the S-EP distribution was examined by conducting Monte Carlo simulations for different sizes. Finally, The S-EP model surpasses other well-known competitor models like the S-E, EP,NWP ,E, S-HLIR, S-L and S-ML models in terms of fit, according to applications to four real datasets.

Declarations

Competing interests: The author declare no Competing of interest.
Authors’ contributions: The author wrote the entire article alone without any input from anyone else.
Funding: This research did not receive any funding
Availability of data and materials: Data and materials included in article
Acknowledgments: The author is thankful to the editorial team and referees for helpful comments that improved this article.

References

[1] A.A. Al-Babtain, I. Elbatal, C. Chesneau, M. Elgarhy, *Sine Topp-Leone-G family of distributions: Theory and applications*, OpenPhys, **18**(**1**)(2020), 574–593.

[2] K.A Al-Kadim,& M.A Boshi, *Exponential pareto distribution* , Mathematical Theory and Modeling, **3**(**5**)(2013), 135-146.

[3] M.A Aldahlan, F. Jamal, Ch. Chesneau, M. Elgarhy, I. Elbatal, *The truncated Cauchy power family of distributions with inference and applications*, Entropy, **22**(**3**)(2020), 1-25.

[4] A. Algarni, A. M Almarashi, I. Elbatal, A.S Hassan, E. M Almetwally, A.M Daghistani, and M. Elgarhy, *Type I half logistic Burr X-G family: properties, Bayesian, and non-Bayesian estimation under censored samples and applications to COVID-19 data*, Mathematical Problems in Engineering,(2021), 1-21.

[5] R. Amirov, A. Ergun, and S. Durak, *Half-inverse problems for the quadratic pencil of the Sturm-Liouville equations with impulse*, Numerical Methods for Partial Differential Equations, **37**(2021), 915-924.

[6] R.K Amirov, & Ergün, *Half inverse problems for the impulsive singular diffusion operator*, Turkish Journal of Science, **5(3)**(2020), 186-198.
[7] M. Bourguignon, R.B Silva,& G.M Cordeiro, *The Weibull-G family of probability distributions*, Journal of Data Science, **12(1)**(2014), 53-68.
[8] S.M Burroughs,& S.F Tebbens, *Upper-truncated power law distributions*, Fractals,**(2)9**(2001), 209-222.
[9] G.M Cordeiro, M. Alizadeh, T.G Ramires ,& E.MM Ortega, *The generalized odd half-Cauchy family of distributions: properties and applications*, Communications in Statistics,theory and Methods, **46(11)**(2017), 5685-5705.
[10] G.M Cordeiro,& M.de Castro, *A new family of generalized distributions*, Journal of Statistical Computation and Simulation,**81(7)**(2011), 883-898.
[11] A. Ergun, *A half-inverse problem for the singular diffusion operator with jump conditions*, Miskolc Mathematical Notes, **21(2)**(2020), 805-821.
[12] M.A Haq,& M. Elgarhy, *The odd Fréchet-G family of probability distributions*, Journal of Statistics Applications & Probability, **7(1)**(2018), 189-203.
[13] J.Havrda,& F. Charv ´at, *Quantification method of classification processes, concept of structural a-entropy*, Journal of Kybernetika, **3(1)**(1967), 30-35.
[14] R. Hibatullah, Y .Widyaningsih, & S. Abdullah, *Marshall-Olkin extended power Lindley distribution with application*, Jurnal Riset dan Aplikasi Matematika (JRAM),**2(2)**(2018), 84-92.
[15] F. Jamal ,C. Chesneau ,DL. Bouali , & M. Ul Hassan, *Beyond the Sin-G family: The transformed Sin-G family*, Plosone. **16(5)**2021 .
[16] JF. Kenney, *Mathematics of statistics*. D. Van Nostrand; (1939).
[17] H.S Klakattawi, *Survival analysis of cancer patients using a new extended Weibull distribution*,Plos one, **17(2)**, (2022).
[18] D. Kumar, U. Singh, & S.K Singh, *A new distribution using sine function- its application to bladder cancer patients data*, Journal of Statistics Applications & Probability, **4(3)**(2015), 417-427.
[19] D. Kumar, U. Singh, S. K. Singh, and P. K. Chaurasia, *Statistical properties and application of a Lifetime model using sine function*, Int. J. Creat. Res. Thoughts, **6**(2018), 933-1002.
[20] M. Mansour, H.M Yousof, W.A Shehata, M. A Ibrahim, *New two parameter Burr XII distribution: properties, copula, different estimation methods and modeling acute bone cancer data*,Journal of Nonlinear Science and Applications, **13(5)**(2020), 223-238.
[21] J.Mazucheli, A.F Menezes, and S. Dey, *Unit-Gompertz distribution with applications*, Statistica,**79(1)**(2019), 25–43.
[22] J.J Moors, *A quantile alternative for kurtosis*, Journal of the Royal Statistical Society: Series D, **37(1)**(1988), 25-32.
[23] M. Muhammad, H. M Alshanbari, A. R Alanzi, L. Liu, W. Sami, C. Chesneau, and F. Jamal, *A new generator of probability models: the exponentiated sine-G family for lifetime studies*, Entropy, **23(11)**(2021), 1394.
[24] S. Nasiru, & A. Luguterah, *The new weibull-pareto distribution*, Pakistan Journal of Statistics and Operation Research,(2015), 103-114.
[25] V. Pareto, *Cours d'economie politique,Librairie Droz, **1**, (1964).*
*[26] A. Rényi, On measures of entropy and information, In Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, **4**(1961), 547-561.*
*[27] C.E Shannon, A mathematical theory of communication, The Bell system technical journal, **27(3)**(1948), 379-423.*
[28] M. Shrahili, I. Elbatal, and M. Elgarhy, Sine Half-Logistic Inverse Rayleigh Distribution: Properties, Estimation, and Applications in Biomedical Data, Journal of Mathematics,(2021), 1-10.
*[29] L. Tomy, G. Veena and C. Chesneau, Applications of the Sine Modified Lindley Distribution to Biomedical Data, Mathematical and Computational Applications, **27(3)**(2022), 43-59.*
*[30] A.Ullah, Entropy, divergence and distance measures with econometric applications, Journal of Statistical Planning and Inference, **49(1)**(1996), 137-162.*
*[31] H.M Yousof, A.Z Afify, G.G Hamedani, & G. Aryal, The burr X generator of distributions for lifetime data, Journal of Statistical ,Theory and Applications, **16(3)**(2017), 288-305.*
*[32] D. Zwillinger , & A. Jeffrey , Table of Integrals, Series and Products, Elsevier,**7** (2007).*