Volume 65, Issue 5, 2021
Journal of Scientific Research
of
The Banaras Hindu University

A novel family of generating distributions based on trigonometric function with an application to exponential distribution.

Anwar Hassan[1], I. H. Dar*[2] and M. A. Lone[3]

[1]Department of Statistics, University of Kashmir, Srinagar, India. Anwar.hassan5@gmail.com
*[2]Department of Statistics, University of Kashmir, Srinagar, India. Ishfaqqh@gmail.com
[3]Department of Statistics, University of Kashmir, Srinagar, India. Murtazastat@gmail.com

***Abstract*—In this manuscript, we introduce a noval procedure for generating distributions based on the sine trigonometric function, and we called this procedure as the Sine Exponentiated Transformation (SET). The SET procedure is then specialized on exponential distribution and a new distribution, namely, Sine Exponentiated Exponential (SEE) distribution is acquired. The introduced model is quite flexible in terms of density and hazard rate functions. Besides flexibility, several other well known properites including moments, moment generating function, mean residual life, mean waiting time, stress strength parameter and order statistics has been highlighted. Simulation study has been carried out to assess the performance of all the estimators. Lastly the applicability of the distribution is discussed on three different real data sets.**

***Index Terms*—Sine exponentiated transformation; Exponential distribution; Hazard rate function; Moments; maximum likelihood estimation.**

I. Introduction

The statistical literature incorporate a plethora of probability models for modeling different real life random phenomenon in various areas such as insurance, actuarial, demography, economics, medical sciences, finance and engineering. Since there is no particular distribution sufficient for modeling every phenomenon, a number of new models with a high degree of flexibility is increasing year by year. So the researchers have switched their attention to set up new family of distributions and proposed a variety of new families of distributions so that real life data can be better assessed and investigated in different applied areas. Among them Marshall and Olkin (1997) proposed a new method for generating family of distributions. Eugene et al. (2002) proposed the beta generalized method. Mudholkar and Srivastava (1993) proposed a method to introduce an extra parameter to the Weibull distribution. Aldeni et al. (2017) developed a new family of distribution arising from the inverse cumulative distribution function of the generalized lambda distribution. Alzaatreh et al. (2014) acted on the family of T-normal distributions.Alzaatreh et al. (2016) suggested the family of generalized Cauchy distributions. Mahdavi and Kundu (2017) focused on the Alpha Power Transformation (APT) family of distributions. Cordeiro et al. (2017) put forward the half Cauchy family of distributions. Ijaz et al. (2020) acted on the Gull Alpha Power Weibull distribution. Nassar et al. (2018) suggested a method based on the idea of Alpha power transformation. Recently, Ijaz et al. (2021) proposed class of New Alpha Power Transformed family (NAPT) of distributions. They employed exponential distribution in NAPT family and derived a new distribution called New Alpha Power Transformed exponential (NAPTE) distribution. In fact, the statistical literature has a scarcity of distributions which are based on trigonometric functions, most of them are based on algebraic functions. Among trigonometric distribution functions, let us point out Mahmood and Chesneau (2019) who introduced a new sine-G family of distributions. Chesneau et al. (2019) proposed a new class of probability distributions via cosine and sine functions. Chesneau and Jamal (2019) introduced the Sine Kumaraswamy-G Family of distributions. Recently, Al-Babtain et al. (2020) proposed the Sine Topp-Leone-G family of distributions.

Based on the motivations above, we introduce a noval family of distributions which is based on trigonomentric function that brings more flexibility to the given family. Then we give a comprehensive account of its general mathematical properties, such as shapes of probability density and hazard rate functions, useful expansions, moments and moment generating function. After providing a comprehensive treatment of its mathematical properties, we focus our attention on a special member of this family, defined with the exponential distribution as baseline. It is named as the SEE distribution.

The rest of the paper is sorted as follows: In section 2 a novel family of probability distributions called SET has been focused and some well known properties of this family have been

DOI: 10.37398/JSR.2021.650519

discussed. In section 3, SEE distribution has been examined, its structural properties have been discussed. In section 4, Maximum likelihood estimators of unknown parameter as well as simulation study have been carried out. In secton 5, three real life data sets have been analyzed to demonstrate the efficacy of the proposed model. Finally, the paper is concluded in section 6.

II. Sine Exponentiated Trasformation (SET)and its properties

Let $g(y)$ and $G(y)$ be the probability density function (pdf) and cumulative distribution function (cdf) of any continuous random variable Y, respectively. Then the cdf $F(y)$ of the SET family of distributions is defined as

$$F(y) = G(y)\sin\left(\frac{\pi}{2}G^{\alpha}(y)\right) \ ; \ y\in\mathbb{R},\ \alpha\geq 0 \tag{1}$$

the corresponding pdf is

$$f(y) = g(y)\left(\frac{\alpha\pi}{2}G^{\alpha}(y)\cos(\frac{\pi}{2}G^{\alpha}(y)) + \sin(\frac{\pi}{2}G^{\alpha}(y))\right) \ ; \quad y\in\mathbb{R},\ \alpha\geq 0.$$

The survival function $S(y)$ for SET family of distributions is given by

$$S(y) = 1 - G(y)\sin\left(\frac{\pi}{2}G^{\alpha}(y)\right) \ ; \ y\in\mathbb{R},\ \alpha\geq 0.$$

The hazard rate function, $\lambda(y)$ is given by

$$\lambda(y) = \frac{g(y)\left(\frac{\alpha\pi}{2}G^{\alpha}(y)\cos(\frac{\pi}{2}G^{\alpha}(y)) + \sin(\frac{\pi}{2}G^{\alpha}(y))\right)}{1 - G(y)\sin(\frac{\pi}{2}G^{\alpha}(y))} \ ; \quad y\in\mathbb{R},\ \alpha\geq 0.$$

The following series representations will be used throughout the paper.

$$(1-x)^n = \sum_{k=0}^{n}\binom{n}{k}(-1)^k x^k \ ; \qquad |x|<1 \tag{2}$$

$$\sin(x) = \sum_{k=0}^{\infty}(-1)^k\frac{x^{2k+1}}{(2k+1)!} \ ; \qquad x\in\mathbb{R} \tag{3}$$

$$\cos(x) = \sum_{k=0}^{\infty}(-1)^k\frac{x^{2k}}{(2k)!} \ ; \qquad x\in\mathbb{R} \tag{4}$$

$$e^{x} = \sum_{k=0}^{\infty}\frac{x^k}{k!} \ ; \qquad x\in\mathbb{R} \tag{5}$$

$$(\alpha)^{x} = \sum_{k=0}^{\infty}\frac{(log\alpha)^k}{k!}x^k \tag{6}$$

using (2), (3) and (4) the power series expansion for the cdf and pdf of SET family of distributions in terms survial function $S(y)$ are respectively given by

$$F(y) = \sum_{j=0}^{\infty}\sum_{k=0}^{\alpha(2j+1)+1} a_{jk}\ S^k(y)$$

where

$$a_{jk} = \frac{(-1)^{j+k}(\frac{\pi}{2})^{2j+1}}{(2j+1)!}\binom{\alpha(2j+1)+1}{k}$$

and

$$f(y) = g(y)\sum_{j=0}^{\infty}\sum_{j=0}^{\alpha(2j+1)} a_{jl}\ S^k(y)$$

where

$$a_{jl} = \frac{(-1)^{j+k}(\frac{\pi}{2})^{2j+1}}{(2j)!}\binom{\alpha(2j+1)}{l}(\alpha + \frac{1}{2j+1})$$

III. Sine Exponentiated Exponential (SEE) distribution and its properties

Let Y be a random variable follows the exponential distribution with cdf $G(y) = 1 - e^{-\theta y}$; $y, \theta > 0$ then the cdf of the SEE distribution is defined as

$$F(y) = (1-e^{-\theta y})\sin\left(\frac{\pi}{2}(1-e^{-\theta y})^{\alpha}\right) \ ; \ y\in\mathbb{R}^{+},\ \alpha\geq 0,$$

the corresponding pdf is

$$f(y) = \theta e^{-\theta y}\left(\frac{\alpha\pi}{2}(1-e^{-\theta y})^{\alpha}\cos(\frac{\pi}{2}(1-e^{-\theta y})^{\alpha}) + \sin(\frac{\pi}{2}(1-e^{-\theta y})^{\alpha})\right) \ ; \ y\in\mathbb{R}^{+},\ \alpha\geq 0$$

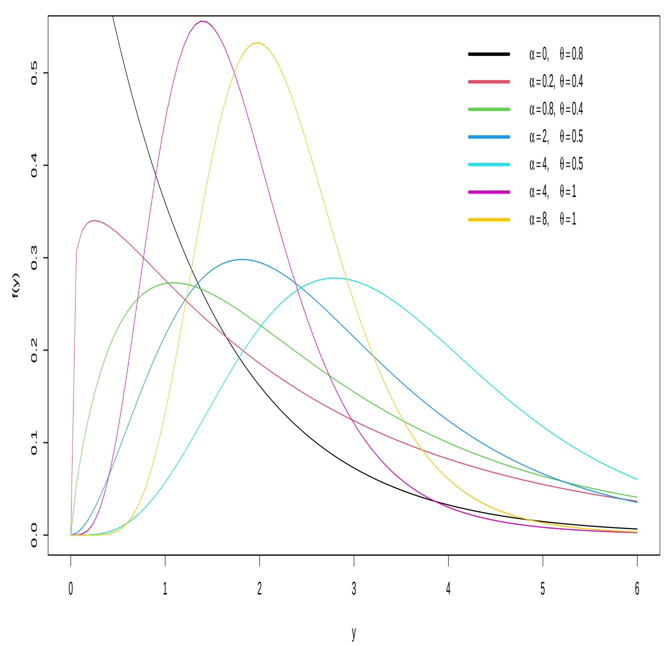

Fig. 1. Plots of the SEE density for different values of α and θ.

The survival and hazard rate functions are, respectively , given by

$$S(y) = 1 - (1 - e^{-\theta y}) \sin\left(\frac{\pi}{2}(1 - e^{-\theta y})^{\alpha}\right) \; ; \quad y \in \mathbb{R}^{+},\; \alpha \geq 0$$

and

$$\lambda(y) = \frac{\theta e^{-\theta y}\left(\frac{\alpha\pi}{2}(1 - e^{-\theta y})^{\alpha}\cos(\frac{\pi}{2}(1 - e^{-\theta y})^{\alpha}) + \sin(\frac{\pi}{2}(1 - e^{-\theta y})^{\alpha})\right)}{1 - (1 - e^{-\theta y})\sin(\frac{\pi}{2}(1 - e^{-\theta y})^{\alpha})}; \quad y \in \mathbb{R}^{+},\; \alpha \geq 0$$

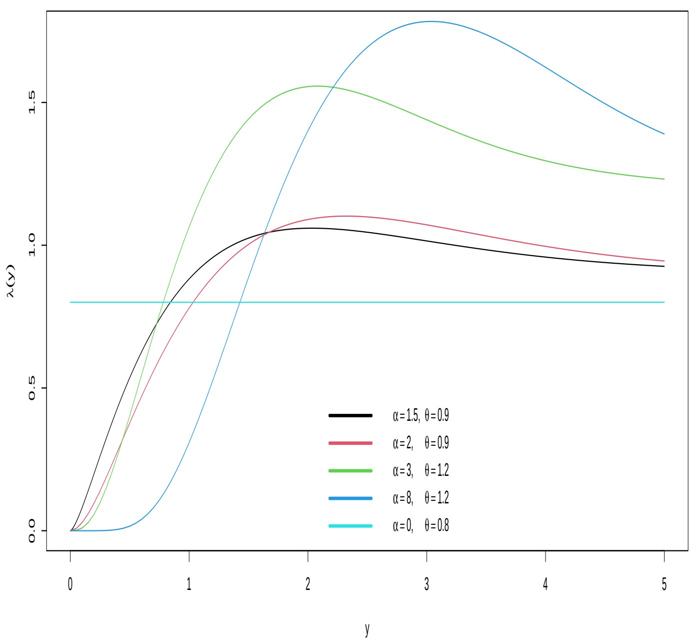

Fig. 2. Plots of the SEE hazard rate function for different values of α and θ.

A. Moments

The r^{th} moment of the SEE distribution is given by

$$E(Y^r) = \int_0^{\infty} y^r f(y)dy$$
$$= \frac{1}{\theta}\sum_{j=0}^{\infty}\frac{(-1)^j r!(\frac{\pi}{2})^{2j+1}}{(2j)!}\left[\alpha\sum_{i=0}^{\alpha}\sum_{k=0}^{2\alpha j}\frac{(-1)^{i+k}\binom{\alpha}{i}\binom{2\alpha j}{k}}{(i+k)^r} + \sum_{l=0}^{\alpha(2j+1)}\frac{(-1)^l\binom{\alpha(2j+1)}{l}}{(2j+1)l^r}\right]$$

B. Moment generating function

The moment generating function of SEE distribution is given by

$$M_Y(t) = \int_0^{\infty} e^{ty} f(y)dy$$
$$= \frac{1}{\theta}\sum_{j=0}^{\infty}\sum_{r=0}^{\infty}\frac{(-1)^j t^r(\frac{\pi}{2})^{2j+1}}{(2j)!}\left[\alpha\sum_{i=0}^{\alpha}\sum_{k=0}^{2\alpha j}\frac{(-1)^{i+k}\binom{\alpha}{i}\binom{2\alpha j}{k}}{(i+k)^r} + \sum_{l=0}^{\alpha(2j+1)}\frac{(-1)^l\binom{\alpha(2j+1)}{l}}{(2j+1)l^r}\right]; \quad t < \theta$$

C. Mean residual life and mean waiting time

Suppose that Y is a continuous random variable with survivial function $S(y)$ then, the mean residual life function, say $\mu(t)$, is defined by

$$\mu(t) = \frac{1}{S(t)}\left(E(t) - \int_0^{t} yf(y)dy\right) - t$$

The mean residual life of SEE distribution is given by

$$\mu(t) = \frac{1}{\theta}\sum_{j=0}^{\infty}\frac{(-1)^j(\frac{\pi}{2})^{2j+1}}{(2j)!}\left[\alpha\sum_{i=0}^{\alpha}\sum_{k=0}^{2\alpha j}\frac{(-1)^{i+k}\binom{\alpha}{i}\binom{2\alpha j}{k}}{(i+k)}A + \sum_{l=0}^{\alpha(2j+1)}\frac{(-1)^l\binom{\alpha(2j+1)}{l}}{(2j+1)l}B\right] - t$$

where

$$A = \left(1 - e^{-\theta(i+k)t}(\theta t + 1)\right)$$

and

$$B = \left(1 - e^{-\theta l t}(\theta t + 1)\right)$$

The mean waiting time of Y, say $\bar{\mu}(t)$, is given by

$$\bar{\mu}(t) = t - \frac{1}{F(t)}\int_0^{t} yf(y)dy$$
$$\bar{\mu}(t) = t - \frac{1}{F(t)}\left\{\frac{1}{\theta}\sum_{j=0}^{\infty}\frac{(-1)^j(\frac{\pi}{2})^{2j+1}}{(2j)!}\left[\alpha\sum_{i=0}^{\alpha}\sum_{k=0}^{2\alpha j}\frac{(-1)^{i+k}\binom{\alpha}{i}\binom{2\alpha j}{k}}{(i+k)}A + \sum_{l=0}^{\alpha(2j+1)}\frac{(-1)^l\binom{\alpha(2j+1)}{l}}{(2j+1)l}B\right]\right\}$$

where

$$A = \left(1 - e^{-\theta(i+k)t}(\theta t + 1)\right)$$

and

$$B = \left(1 - e^{-\theta l t}(\theta t + 1)\right)$$

D. Order statistics

Here, we focus on the order statistics related to the sine exponentiated transformation family of distributions. Let $Y_1, Y_2, ..., Y_n$ be a random sample of size n, and let $Y_{i:n}$ denote the i^{th} order statistic, then the pdf of $Y_{i:n}$, say $f_{i:n}(y)$ is given by

$$f_{i:n}(y) = \frac{n!}{(i-1)!(n-i)!} F(y)^{i-1} f(y)(1-F(y))^{n-i} \quad (7)$$

By using (2), (3), (4), (5) and (6) we get the i^{th} order statistic of SEE distribution as

$$f_{i:n}(y) = \frac{n!}{(i-1)!(n-i)!} \theta e^{-\theta y} \quad (8)$$

$$\times \sum_{k=0}^{\infty} \frac{(-1)^k (\frac{\pi}{2})^{2k+1}}{(2k)!} [a_{mp} + b_q] c_{rs} \quad (9)$$

where

$$a_{mp} = \alpha \sum_{m=0}^{\alpha} \sum_{p=0}^{2k\alpha} (-1)^{m+p} \binom{\alpha}{m} \binom{2k\alpha}{p} e^{-\theta(m+p)y}$$

$$b_q = \sum_{q=0}^{\alpha(2k+1)} \frac{(-1)^q \binom{\alpha(2k+1)}{q} e^{-\theta q y}}{(2k+1)}$$

and

$$c_{rs} = \sum_{r=0}^{n-i} \sum_{s=0}^{r} (-1)^{r+s} \binom{n-i}{r} \binom{r}{s} e^{-\lambda s y} \left(sin \left(\frac{\pi}{2} (1 - e^{-\theta y})^{\alpha} \right) \right)^r$$

E. Stress strength parameter

Suppose Y_1 and Y_2 be independent strength and stress random variables respectively, where $Y_1 \sim SEE(\alpha_1, \theta)$ and $Y_2 \sim SEE(\alpha_2, \theta)$, then the stress strength parameter $\mathbb{P}(Y_1 > Y_2)$, say R, is defined as

$$R = \int_{-\infty}^{\infty} f_1(y) F_2(y) dy,$$

and is given by

$$R = \theta \left(\sum_{j=0}^{\infty} \frac{(-1)^j (\frac{\pi}{2})^{2j+1}}{(2j)!} \right)^2 \sum_{m=0}^{\alpha_2} \sum_{n=0}^{\alpha_2(2j+1)} (-1)^{m+n} \binom{\alpha_2}{m} \times \binom{\alpha_2(2j+1)}{n} [C + D]$$

where

$$C = \alpha_1 \sum_{i=0}^{\alpha_1} \sum_{k=0}^{2\alpha_1 j} \frac{(-1)^{i+k} \binom{\alpha_1}{i} \binom{2\alpha_1 j}{k}}{\theta(i+k+m+n+1)}$$

and

$$D = \sum_{l=0}^{\alpha_1(2j+1)} \frac{(-1)^l \binom{\alpha_1(2j+1)}{l}}{\theta(l+m+n+1)}$$

TABLE I
AVERAGE VALUES OF MLES THEIR CORRESPONDING MSES AND BIAS (N=50).

Parameter		MLE		MSE		Bias	
α	θ	$\hat{\alpha}$	$\hat{\theta}$	$\hat{\alpha}$	$\hat{\theta}$	$\hat{\alpha}$	$\hat{\theta}$
0.5	0.5	0.57134	0.52205	0.08633	0.00652	0.07134	0.02205
	1	0.55059	1.03528	0.08402	0.02709	0.05059	0.03528
	1.5	0.54335	1.55186	0.07599	0.06009	0.04335	0.05186
	2	0.57550	2.08942	0.09867	0.09950	0.07550	0.08942
	3	0.55956	3.10780	0.08314	0.23320	0.05956	0.10780
1	0.5	1.10631	0.51406	0.14425	0.00513	0.10631	0.01406
	1	1.09188	1.03370	0.14567	0.02435	0.09188	0.03370
	1.5	1.07865	1.54974	0.14357	0.05318	0.07865	0.04974
	2	1.10072	2.07756	0.12611	0.09244	0.10072	0.07756
	3	1.10125	3.08374	0.17917	0.22898	0.10124	0.08374
1.5	0.5	1.62299	0.51058	0.29192	0.00464	0.12299	0.01058
	1	1.68653	1.04376	0.29055	0.02256	0.18653	0.04376
	1.5	1.63695	1.54223	0.25616	0.04723	0.13694	0.04223
	2	1.63100	2.07223	0.24397	0.08516	0.13100	0.07223
	3	1.61684	3.08658	0.27250	0.20119	0.11684	0.08658
2	0.5	2.19155	0.51765	0.45862	0.00582	0.19155	0.01765
	1	2.20198	1.02842	0.44930	0.02257	0.20198	0.02842
	1.5	2.17571	1.54804	0.48812	0.04791	0.17571	0.04804
	2	2.16670	2.06759	0.43514	0.09086	0.16670	0.06758
	3	2.14301	3.05925	0.42633	0.17063	0.14301	0.05925
3	0.5	3.21627	0.51368	0.97494	0.00511	0.21627	0.01368
	1	3.34516	1.03478	1.05264	0.01954	0.34516	0.03478
	1.5	3.20841	1.53981	0.89062	0.04742	0.20841	0.03981
	2	3.20939	2.04707	0.85746	0.07370	0.20939	0.04707
	3	3.27796	3.08750	1.16273	0.18555	0.27796	0.08750

TABLE II
AVERAGE VALUES OF MLES THEIR CORRESPONDING MSES AND BIAS (N=100).

Parameter		MLE		MSE		Bias	
α	θ	$\hat{\alpha}$	$\hat{\theta}$	$\hat{\alpha}$	$\hat{\theta}$	$\hat{\alpha}$	$\hat{\theta}$
0.5	0.5	0.51387	0.50797	0.03745	0.00312	0.01387	0.00797
	1	0.53577	1.02252	0.03693	0.01068	0.03577	0.02252
	1.5	0.51881	1.53341	0.03299	0.02613	0.01881	0.03341
	2	0.51376	2.03174	0.03739	0.05001	0.01376	0.03174
	3	0.52029	3.06187	0.03304	0.10124	0.02029	0.06187
1	0.5	1.04931	0.50710	0.06118	0.00274	0.04931	0.00710
	1	1.03798	1.01716	0.06440	0.01006	0.03798	0.01716
	1.5	1.03728	1.52403	0.05765	0.02356	0.03728	0.02403
	2	1.03812	2.02885	0.06048	0.04616	0.03812	0.02885
	3	1.04487	3.03736	0.07504	0.09693	0.04487	0.03736
1.5	0.5	1.58102	0.50790	0.11685	0.00215	0.08102	0.00790
	1	1.57586	1.02099	0.11108	0.00957	0.07586	0.02099
	1.5	1.56230	1.52513	0.10804	0.02354	0.06230	0.02513
	2	1.53990	2.02573	0.11185	0.04467	0.03990	0.02573
	3	1.55084	3.06152	0.09967	0.09573	0.05084	0.06152
2	0.5	2.06847	0.50904	0.16163	0.00244	0.06847	0.00904
	1	2.09184	1.01562	0.17683	0.00998	0.09184	0.01562
	1.5	2.07191	1.51987	0.17464	0.02352	0.07191	0.01987
	2	2.07935	2.03059	0.18402	0.03948	0.07935	0.03059
	3	2.08810	3.04104	0.17481	0.08929	0.08810	0.04104
3	0.5	3.15415	0.50689	0.40534	0.00243	0.15415	0.00689
	1	3.11076	1.01398	0.35868	0.00872	0.11076	0.01398
	1.5	3.11446	1.51956	0.38501	0.01971	0.11446	0.01956
	2	3.13329	2.03309	0.40414	0.03598	0.13329	0.03309
	3	3.14568	3.05367	0.41051	0.08683	0.14568	0.05367

TABLE III
MLEs (STANDARD ERRORS IN PARENTHESES), K-S STATISTIC, AND P-VALUES FOR THE DATA SET I.

Model	Estimates $\hat{\alpha}$	$\hat{\theta}$	$\hat{\theta}$	Statistics K-S	p-value
WE	3.13216 (0.13308)	0.01807 (0.01229)	0.80693 (0.13308)	0.06202	0.83650
MW	0.10124 (0.01012)	0.00100 (0.00703)	0.02597 (0.01287)	0.17369	0.00479
SEE	1.21399 (0.26669)	0.12592 (0.01236)	-	0.03715	0.99910
APE	21.16923 (14.16432)	0.18307 (0.0.01974)	-	0.05285	0.94270
NAPTE	2.32487 (1.75530)	0.07530 (0.01570)	-	0.04024	0.99690
W	0.11171 (0.00925)	0.90032 (0.11114)	-	0.05745	0.89610
G	2.00884 (0.26389)	0.20338 (0.03032)	-	0.04252	0.99360
E	0.10124 (0.01012)	-	-	0.17301	0.00502
SE	0.05722 (0.00534)	-	-	0.15587	0.01552

TABLE IV
$-2l(\hat{\theta})$, AIC, AICC, BIC FOR THE DATA SET I.

Model	$-2l(\hat{\theta})$	AIC	AICC	BIC
WE	641.6886	647.6886	647.9386	655.5410
MW	658.2426	664.2426	664.4926	672.0581
SEE	634.2848	638.2848	638.4085	643.4952
APE	638.0736	642.0736	642.1973	647.2840
NAPTE	634.8204	638.8204	638.9441	644.0307
W	637.4892	641.4893	641.6130	446.6996
G	634.6002	638.6002	638.7240	643.8106
E	658.0418	660.0418	660.0826	662.6469
SE	653.0729	655.0729	655.1137	657.6780

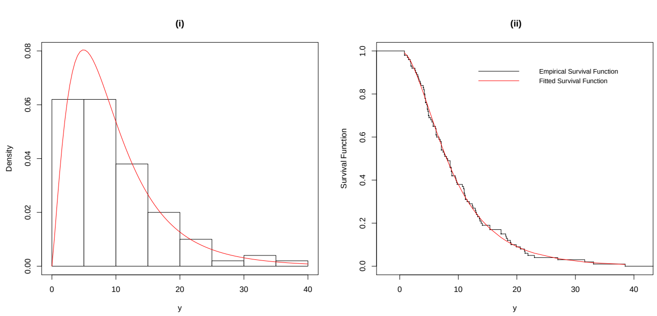

Fig. 3. (i) The relative histogram and the fitted SEE distribution. (ii) The fitted SEE survival function and empirical survival function for data set I.

IV. STATISTICAL INFERENCE

A. Maximum likelihood estimators

Let $y_1, y_2, ..., y_n$ be a random sample from SEE(α, θ) distribution with parameter vector Θ=(α, θ) then the logarithm

TABLE V
MLEs (STANDARD ERRORS IN PARENTHESES), K-S STATISTIC, AND P-VALUES FOR THE DATA SET II.

Model	Estimates $\hat{\alpha}$	$\hat{\theta}$	$\hat{\lambda}$	Statistics K-S	p-value
WE	3.95810 (1.21408)	0.01796 (0.00466)	0.85819 (0.05928)	0.07458	0.47470
MW	0.10518 (0.05587)	0.00100 (0.03494)	1.16895 (0.98153)	0.08387	0.32890
SEE	0.43482 (0.15363)	0.11365 (0.01031)	-	0.06377	0.67520
APE	1.17446 (0.15363)	0.11134 (0.01031)	-	0.07932	0.39630
NAPTE	3.39395 (0.51098)	0.12091 (0.01359)	-	0.07250	0.51140
W	0.09438 (0.01912)	1.04576 (0.06742)	-	0.07000	0.55720
G	1.17255 (0.00043)	0.12520	-	0.07329	0.49750
E	0.10676 (0.00943)	-	-	0.08463	0.31830
SE	0.05974 (0.00498)	-	-	0.07126	0.53390

TABLE VI
$-2l(\hat{\theta})$, AIC, AICC, BIC FOR THE DATA SET II.

Model	$-2l(\hat{\theta})$	AIC	AICC	BIC
WE	838.7996	845.7996	845.9931	854.3557
MW	828.6628	834.6628	834.8564	843.2189
SEE	824.0739	828.0739	828.1699	833.7780
APE	828.6364	832.6364	832.7324	838.3404
NAPTE	826.1586	830.1586	830.2546	835.8627
W	828.1748	832.1747	832.2707	837.8788
G	826.7356	826.7356	830.8316	836.4396
E	828.6838	830.6838	830.7155	833.5358
SE	828.6652	830.6652	830.6969	833.5172

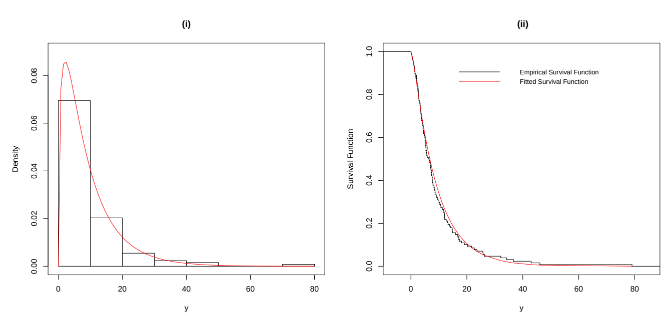

Fig. 4. (i) The relative histogram and the fitted SEE distribution. (ii) The fitted SEE survival function and empirical survival function for data set II.

of the likelihood function can be expressed as

$$l(\Theta) = -nlog\theta - \theta\sum_{i=1}^{n} y_i - \sum_{i=1}^{n} log\left(\frac{\alpha\pi}{2}(1-e^{-\theta y_i})^{\alpha} \times cos(\frac{\pi}{2}(1-e^{-\theta y_i})^{\alpha} + \sin(\frac{\pi}{2}(1-e^{-\theta y_i})^{\alpha})\right)$$

TABLE VII
MLES (STANDARD ERRORS IN PARENTHESES), K-S STATISTIC, AND P-VALUES FOR THE DATA SET III.

Model	Estimates $\hat{\alpha}$	$\hat{\theta}$	$\hat{\lambda}$	Statistics K-S	p-value
WE	0.00881 (0.00405)	0.94623 (0.97387)	1.48174 (1.57994)	0.13208	0.22180
MW	0.32687 (0.04119)	0.01191 (0.29582)	0.00100 0.08371	0.09427	0.63000
SEE	65.75017 (25.39974)	1.37387 (0.13296)	-	0.08654	0.73290
APE	6.68161 (1.67772)	1.04561 (4.81109)	-	0.18082	0.03250
NAPTE	5.56317 (1.18632)	1.04092 (4.70195)	-	0.16544	0.06358
W	4.85846 (0.24854)	0.00337 (0.00105)	-	0.11335	0.39320
G	25.59175 (0.26389)	8.36522 (0.03032)	-	0.08980	0.70120
E	0.32687 (0.04118)	-	-	0.48600	2.378e-13
SE	0.00881 (0.02182)	-	-	0.48152	4.103e-13

TABLE VIII
$-2l(\hat{\theta})$, AIC, AICC, BIC FOR THE DATA SET III.

Model	$-2l(\hat{\theta})$	AIC	AICC	BIC
WE	138.4807	144.4807	144.8875	150.9102
MW	115.6353	119.6353	119.8353	123.9216
SEE	113.0204	117.0204	117.4272	121.3067
APE	147.2311	151.2311	151.4311	155.5173
NAPTE	144.4528	148.4528	148.6528	152.7391
W	124.9071	128.9071	129.1071	133.1934
G	113.7575	117.7575	118.1643	118.1643
E	266.8915	268.8915	268.9571	271.0347
SE	259.0475	261.0475	261.4543	263.1907

The maximum likelihood estimators of Θ can be obtained by solving the non linear normal equations $\frac{\partial l}{\partial \Theta}(\Theta) = 0$. These equations cannot be solved analytically, so in order to get MLEs of parameters we use R software.

B. Simulation study

In this subsection, a Monte Carlo simulation study has been carried out by using R software to attest the consistency of the MLEs. This study is replicated 500 times each with sample sizes (n=50, n=100) with different values of parameters $\alpha = (0.5, 1, 1.5, 3)$, $\theta = (0.5, 1, 1.5, 2, 3)$ were generated from SEE. In each case, the average values of MLEs and the corresponding empirical mean squared errors (MSEs) and bias were attained. The simulation results are presented in table I and table II. In particular, with repect to the theory, we observe that the MSEs and biases decrease with increasing sample size.

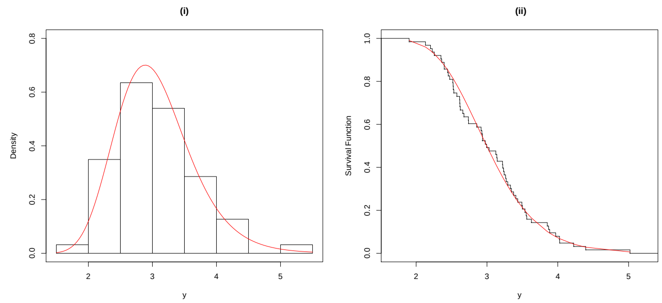

Fig. 5. (i) The relative histogram and the fitted SEE distribution. (ii) The fitted SEE survival function and empirical survival function for data set III.

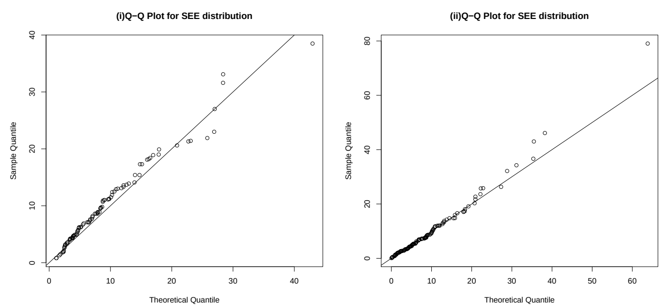

Fig. 6. Q-Q plot for the SEE distribution for data set I and data set II, respectively.

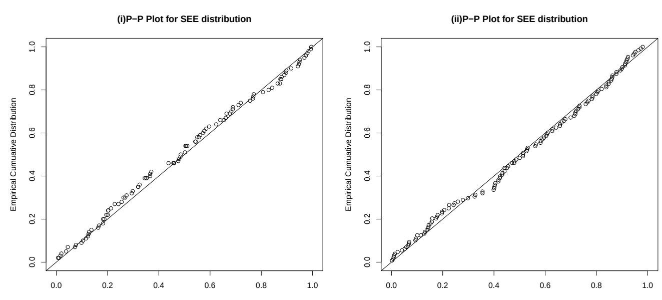

Fig. 7. P-P plot for the SEE distribution for data set I and data set II, respectively.

V. APPLICATION

To test the applicability of the SEE distribution, three real data sets were analyzed. The data set I corresponds to the waiting time (in minutes) of 100 bank customers. The data were taken from Ghitany et al. (2008) and also reported by Bhat et al. (2018).

The data set II corresponds to the remission time in months of 128 bladder cancer patients. The data were taken from Aldeni et al. (2017) and was recently reported by Ijaz et al. (2021). The data set III which is related to engineering field consists of 63 observations of the gauge length of 10mm taken from Kundu and Raqab (2009).

For comparison purpose, we have fitted the proposed SEE with several other models, namely Weibull exponential (WE) Oguntunde et al. (2015), modified Weibull (MW) Sarhan

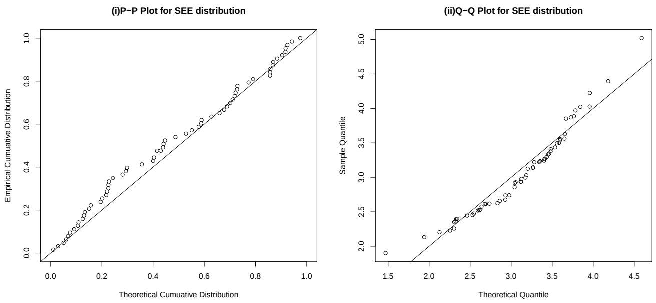

Fig. 8. P-P and Q-Q plots for the SEE distribution for data set III.

and Zaindin (2009), alpha power exponential (APE) Mahdavi and Kundu (2017), noval alpha power transformed exponential (NAPTE) Ijaz et al. (2021), Weibull (W), gamma (G), exponential (E) and sine exponential (SE) Kumar et al. (2015) distributions, their corresponding density functions for $y > 0$ are as follows

$$\text{WE} \quad f(y) = \alpha\beta\theta(1-e^{-\theta y})^{\beta-1}e^{\theta\beta y-\alpha(e^{\theta y}-1)^{\beta}}$$

$$\text{MW} \quad f(y) = (\alpha+\theta\beta y^{\beta-1})e^{-\alpha y-\theta y^{\beta}}$$

$$\text{APE} \quad f(y) = \frac{log\alpha}{\alpha-1}\theta e^{-\theta y}\alpha^{1-e^{-\theta y}}$$

$$\text{NAPTE} \quad f(y) = \theta log(\alpha)\frac{\alpha^{log(1-e^{-\theta y})}}{e^{\theta y}-1}$$

$$\text{SE} \quad f(y) = \frac{\pi}{2}\theta e^{-\theta y}cos\left(\frac{\pi}{2}(1-e^{-\theta y})\right)$$

From Table III, Table IV, Table V, Table VI, Table VII and Table VIII it is apparent that SEE distribution has lowest $-2l(\hat{\theta})$, AIC, AICC, BIC, K-S statistic and highest p-value among all the other competitive models. Hence the introduced model offers the better fit than the other models for the given data sets.

The relative histogram and the fitted SEE distribution of the data set I, II and III are displayed in Figures 3(i), 4(i) and 5(i) respectively. The plots of the fitted SEE survival function and empirical survival function of the data set I, II and III are displayed in Figures 3(ii), 4(ii) and 5(ii) respectively. The Q-Q plots for data set I and II are displayed in Figure 6(i) and 6(ii) respectively. Also, the P-P plots for data set I and II are displayed in Figure 7(i) and 7(ii) respectively, for data set III the P-P and Q-Q plots are displayed in figure 8 that permits us to make a comparison between the empirical distribution of the data with the SEE distribution. These graphical goodness of fit measures undoubtedly support the results given in Tables III, Table IV, Table V, Table VI, Table VII and Table VIII

VI. Conclusion

A noval family of distributions called SET has been introduced. SET family has been specialized on the exponential distribution and a new two-parameter SEE distribution has been obtained. Various mathematical properties of SEE distribution were highlighted. It has been noticed that the two-parameter SEE distribution has more flexibility in terms of the hazard rate and density functions. The potentiality of the proposed model is compared with other existing models by using goodness of fit measures. The model has been fitted to three different real data sets, the figures display that the proposed model provides reasonable fit for all the three data sets in comparison to all other competitive models.

References

Al-Babtain, A. A., Elbatal, I., Chesneau, C., and Elgarhy, M. (2020). Sine topp-leone-g family of distributions: Theory and applications. *Open Physics*, 18(1):574–593.

Aldeni, M., Lee, C., and Famoye, F. (2017). Families of distributions arising from the quantile of generalized lambda distribution. *Journal of Statistical Distributions and Applications*, 4(1):1–18.

Alzaatreh, A., Lee, C., and Famoye, F. (2014). T-normal family of distributions: A new approach to generalize the normal distribution. *Journal of Statistical Distributions and Applications*, 1(1):1–18.

Alzaatreh, A., Lee, C., Famoye, F., and Ghosh, I. (2016). The generalized cauchy family of distributions with applications. *Journal of Statistical Distributions and Applications*, 3(1):1–16.

Bhat, A., Mudasir, S., and Ahmad, S. (2018). Mixture of exponential and weighted exponential distribution: Properties and applications. *Int. J. Sci. Res. in Mathematical and Statistical Sciences Vol*, 5:6.

Chesneau, C., Bakouch, H. S., and Hussain, T. (2019). A new class of probability distributions via cosine and sine functions with applications. *Communications in Statistics-Simulation and Computation*, 48(8):2287–2300.

Chesneau, C. and Jamal, F. (2019). The sine kumaraswamy-g family of distributions.

Cordeiro, G. M., Alizadeh, M., Ramires, T. G., and Ortega, E. M. (2017). The generalized odd half-cauchy family of distributions: properties and applications. *Communications in Statistics-Theory and Methods*, 46(11):5685–5705.

Eugene, N., Lee, C., and Famoye, F. (2002). Beta-normal distribution and its applications. *Communications in Statistics-Theory and methods*, 31(4):497–512.

Ghitany, M. E., Atieh, B., and Nadarajah, S. (2008). Lindley distribution and its application. *Mathematics and computers in simulation*, 78(4):493–506.

Ijaz, M., Asim, S. M., Farooq, M., Khan, S. A., and Manzoor, S. (2020). A gull alpha power weibull distribution with applications to real and simulated data. *PloS one*, 15(6):e0233080.

Ijaz, M., Mashwani, W. K., Göktaş, A., and Unvan, Y. A. (2021). A novel alpha power transformed exponential distribution with real-life applications. *Journal of Applied Statistics*, pages 1–16.

Kumar, D., Singh, U., and Singh, S. K. (2015). A new distribution using sine function-its application to bladder cancer patients data. *Journal of Statistics Applications & Probability*, 4(3):417.

Kundu, D. and Raqab, M. Z. (2009). Estimation of r= p (y¡ x) for three-parameter weibull distribution. *Statistics & Probability Letters*, 79(17):1839–1846.

Mahdavi, A. and Kundu, D. (2017). A new method for generating distributions with an application to exponential distribution. *Communications in Statistics-Theory and Methods*, 46(13):6543–6557.

Mahmood, Z. and Chesneau, C. (2019). A new sine-g family of distributions: properties and applications.

Marshall, A. W. and Olkin, I. (1997). A new method for adding a parameter to a family of distributions with application to the exponential and weibull families. *Biometrika*, 84(3):641–652.

Mudholkar, G. S. and Srivastava, D. K. (1993). Exponentiated weibull family for analyzing bathtub failure-rate data. *IEEE transactions on reliability*, 42(2):299–302.

Nassar, M., Alzaatreh, A., Abo-Kasem, O., Mead, M., and Mansoor, M. (2018). A new family of generalized distributions based on alpha power transformation with application to cancer data. *Annals of Data Science*, 5(3):421–436.

Oguntunde, P., Balogun, O., Okagbue, H., and Bishop, S. (2015). The weibull-exponential distribution: Its properties and applications. *Journal of Applied Sciences*, 15(11):1305–1311.

Sarhan, A. M. and Zaindin, M. (2009). Modified weibull distribution. *APPS. Applied Sciences*, 11:123–136.