Statistical Properties and Application of a Lifetime Model Using Sine Function

[1]Dinesh Kumar, [2]Sanjay Kumar Singh, [3]Umesh Singh, [4]Pradip Kumar and [5]Prashant Kumar Chaurasia
[1]Assistant Professor, [2] Professor, [3]Professor, [4]Research Scholar, [5]Research Scholar
Department of Statistics, Banaras Hindu University, Varanasi-221005

Abstract
In this paper, a new lifetime distribution is introduced on the basis of SS transformation as suggested by Kumar et al. (2015(a)). The considered baseline distribution is Lindley(θ) −distribution. Some of the statistical properties of this new distribution such as MGF, Mean, Median, Skewness and Kurtosis have been studied. A real dataset has been considered and AIC, BIC, K-S test value with its p-value are calculated for this new distribution and for some other distribution too in order to show its application and superiority as compared to some other distributions.

Keywords: Life Time Distribution, Trigonometric Function and Reliability Analysis.

1. Introduction

In day to day life of human, new technologies are being designed or developed, in order to have better item to fulfill need, satisfaction etc. The industries are, therefore, grows as fungus and it is responsibilities of the manufacturer to produce good qualities of their product in terms of their average life or expiry date. As the lifetime of any item is random, therefore any statistical distribution over positive support may be used to model it. In literature, a huge number of such distributions are available. Initially, Exponential, Gamma, Weibull, Log-Normal etc. were used due to their simplicity, flexibility or capability of closed form solution of the estimators of their parameter(s). But, till the date no one distribution is available that can fits all kinds of data. Therefore, several authors have paid attention on this and developed several techniques to generalized or transform any available distribution, called baseline distribution. For examples, (see, Gupta et al.(1998)) proposed the cumulative distribution function (cdf) $G_1(x)$ of new distribution corresponding to the cdf$F_1(x)$ of a baseline distribution as,

$$G_1(x) = \left(F_1(x)\right)^{\alpha}$$

where,$\alpha > 0$is the shape parameter of the proposed distribution. For$\alpha = 1$, the new distribution and the baseline distributions are the same.

Another idea of generalizing a baseline distribution is to transmute it by using the quadratic rank transmutation map (QRTM) (see, Shaw and Buckley (2007)). If $G_2(x)$ be the cdf of transmuted distribution then,

$$G_2(x) = (1+\lambda)F_2(x) - \lambda(F_2(x))^2$$

Where $F_2(x)$ is the cdf of some baseline distribution and $|\lambda| \le 1$. For $\lambda = 0$, the new distribution is same as the baseline distribution.

The one explored by Kumar et al. (2015b) provides a modern alternative. They introduced a new class of distributions in terms of a trigonometric transformation and they call it as SS transformation. If F(x) be the cdf of some baseline distribution, then using SS transformation as suggested by them, the cdf G(x) of new distribution is given by $G(x) = \sin(\frac{\pi}{2}F(x))$. The generalization of SS transformation is introduced by Hussain et al. (2018) and they have used 4 additional parameters. But, additional parameters may lead computational errors, if the closed form estimators may not be obtained. So motivated with this, we are using transformation of baseline distribution.

In the fields of medical sciences, engineering and biological sciences, Lindley distribution has been widely used. It was introduced by Lindley (1958).The probability density function (pdf) and cumulative distribution function (cdf) of a Lindley distribution with the shape parameter θ are given by,

$$f(x) = \frac{\theta^2}{\theta+1}(1+x)e^{-\theta x} \quad ; \; x > 0, \; \theta > 0 \tag{1}$$

and

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

respectively.

Merovci and Elbatal [13] had used the QRTM in order to generate a flexible family of probability distributions taking Lindley-geometric distribution as the baseline distribution that would offer more distributional flexibility and called it transmuted Lindley-geometric distribution. Bakouch et al. [9] obtained an extended Lindley distribution and discussed its various properties and applications. Ghitany et al. [6] developed a two-parameter weighted Lindley distribution and discussed its applications to survival data. A new generalization of the Lindley distribution is recently proposed by Ghitany et al. [4], called as the power Lindley distribution. Another generalization of the Lindley distribution was introduced by Nadarajah et al. [7], named as the generalized Lindley distribution. Oluyede and Yang [14] proposed a new four-parameter class of generalized Lindley (GLD) distribution and called it as beta-generalized Lindley distribution (BGLD). Asgharzadeh et al. [15] introduced a general family of continuous lifetime distributions by compounding any continuous distribution and the Poisson-Lindley distribution.

In the present paper, we have derived a new lifetime distribution using SS transformation and the considered baseline distribution is Lindley distribution having pdf (1). We will use the abbreviation $SS_L(\theta) - distribution$ to denote this new distribution. The cdf and pdf of $SS_L(\theta) - distribution$are given by;

$$G(x) = \cos(\frac{\pi}{2}\left(1 + \frac{\theta x}{\theta+1}\right) e^{-\theta x}) \; ; \;\; x > 0, \theta > 0 \tag{3}$$

and

$$g(x) = \frac{\pi}{2}\frac{\theta^2}{\theta+1}(1+x)e^{-\theta x} \sin\left(\frac{\pi}{2}\left(1 + \frac{\theta x}{\theta+1}\right) e^{-\theta x}\right); \;\; x > 0, \theta > 0 \tag{4}$$

$respectively.$

The shapes of the cdf and pdf of $SS_L(\theta) - distribution$are shown in Figures 1 and 2 respectively for different value of θ.

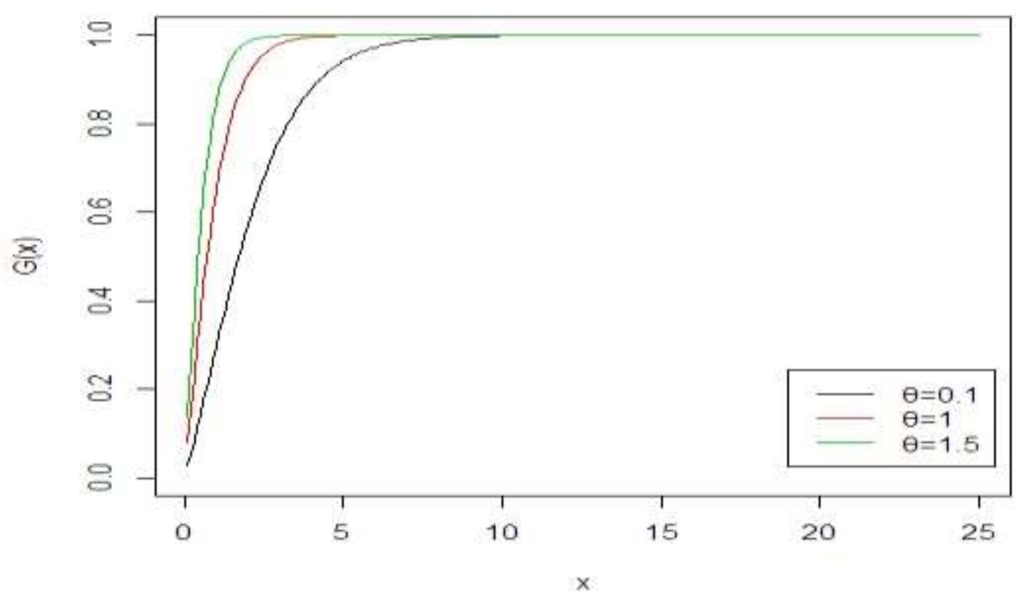


Figure 1: Plots of The cdf of the $SS_L(\theta) - distribution$for different values of θ.

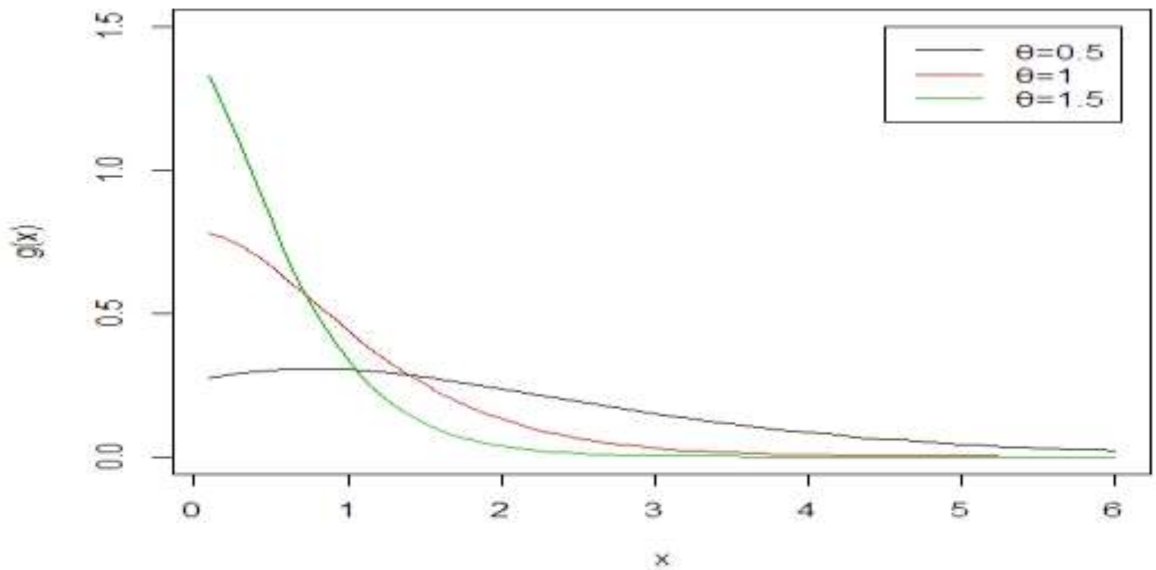


Figure 2: Plots of The pdf of the $SS_L(\theta) - distribution$for different values of θ.

2. RELIABILITY ANALYSIS

In this section, we present survival function, hazard rate function and cumulative hazard rate function of $SS_L(\theta) - distribution$

2.1 Survival Function

The survival function R(x), which is the probability that an item is survived at least x units of time, is defined by R(x)= 1-G(x) and hence the survival function of $SS_{L(\theta)} - distribution$is obtained as follows,

$$R(x) = 1 - \cos\left(\frac{\pi}{2}\left(1 + \frac{\theta x}{\theta + 1}\right)e^{-\theta x}\right); x > 0 \tag{5}$$

The survival behavior of the survival function R(x) for different values of θ is shown in Figure 3.

2.2 Hazard Rate Function

The hazard rate function at time x is the instantaneous failure per unit time for the event to occur, given that the unit has survived up to time x. The hazard rate function h(x) is given as

$$h(x) = \lim_{\Delta \to 0} \frac{P(x < X < x + \Delta x | X > x)}{\Delta x}$$

which provides information about a small interval after time x , (x + Δx). The hazard rate function for a $SS_{L(\theta)} - distribution$can be shown to be

$$h(x) = \frac{\pi}{2}\frac{\theta^2}{\theta+1}(1 + x)e^{-\theta x}\cot\left(\frac{\pi}{2}\left(1 + \frac{\theta x}{\theta+1}\right)e^{-\theta x}\right) ; x > 0 \tag{6}$$

Figure 4, shows that$SS_L(\theta) - distribution$has decreasing hazard rate function for different considered valuesof θ.

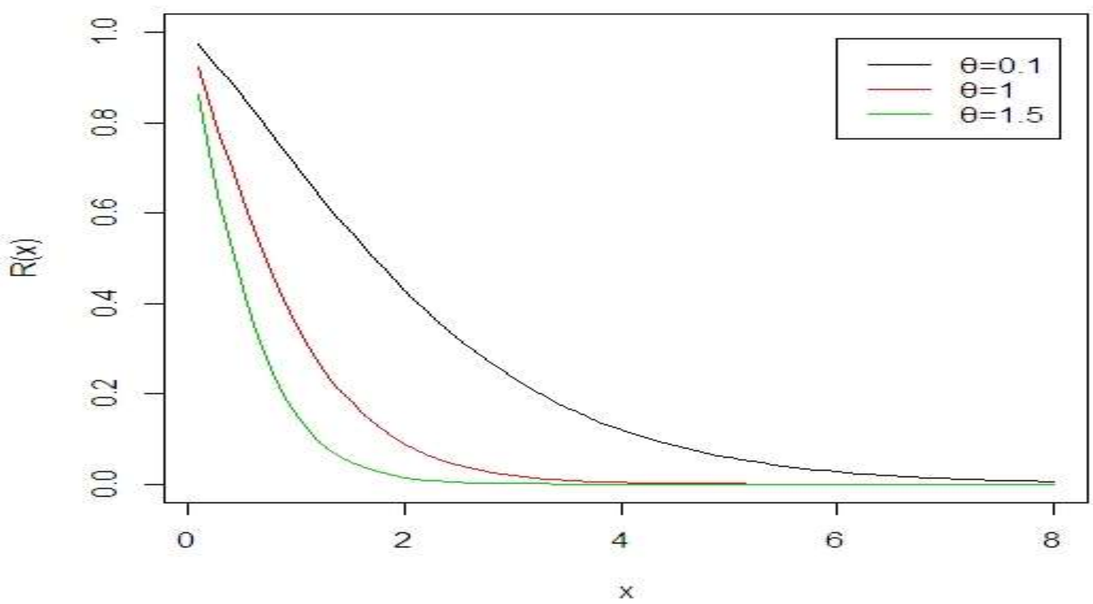


Figure 3: Plots of The Reliability Function of the $SS_{L(\theta)} - distribution$ for different values of θ.

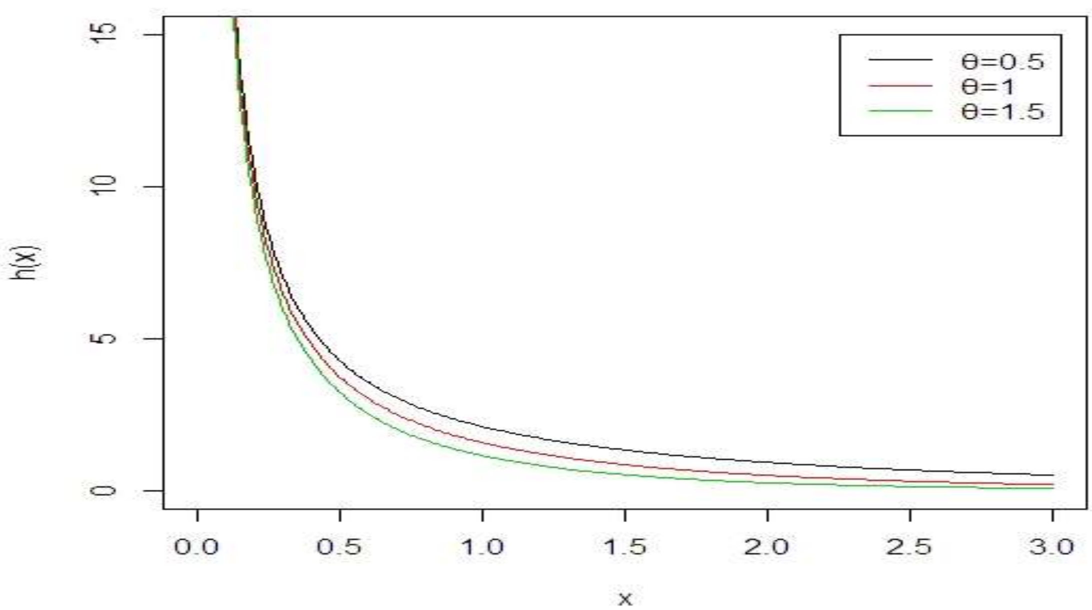


Figure 4: Plots of Hazard Rate function of $SS_L(\theta) - distribution$for different values of θ

2.3 Cumulative Hazard Rate Function

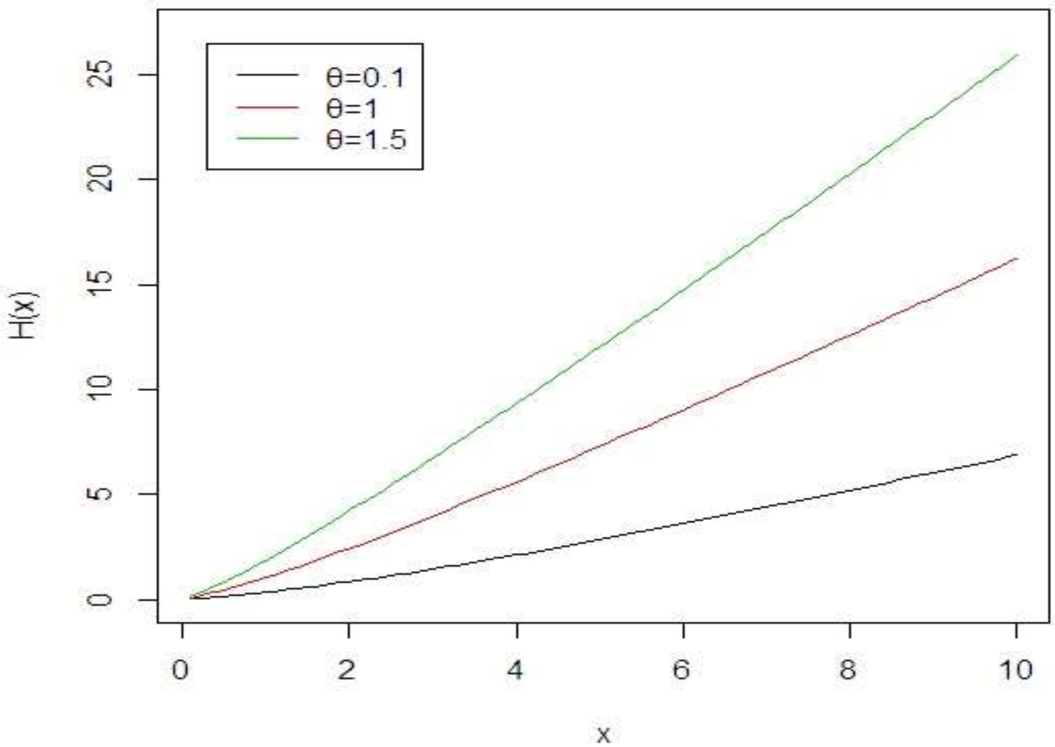
Many generalized models have been proposed in reliability literature through the relationship between the reliability function R(x) and its cumulative hazard rate function H(x) given by H(x) = -lnR(x). The cumulative hazard rate function of the $SS_{L(\theta)} - distribution$is given by

$$H(x) = -\ln\left\{1 - cos\left(\frac{\pi}{2}\left(1+\frac{\theta x}{\theta+1}\right)e^{-\theta x}\right)\right\} ; x > 0 \tag{7}$$

where H(x) is the total number of failure or deaths over an interval of time, and H(x) is a non-decreasing function of x satisfying;

(a) H(0)=0, (b) $\lim_{x\to\infty} H(x) = \infty$

Figure 5 illustrates the behavior of the cumulative hazard rate function of $SS_L(\theta) - distribution$ for different values of the parameter



θ.

Figure 5: Plots of Cumulative Hazard Rate Function of $SS_L(\theta) - distribution$ for different values of θ

3. Statistical Properties

3.1 Moment Generating Function

The Moment Generating Function (MGF) of $SS_L(\theta) - distribution$ having pdf (3) is obtained as follows,

$$M_x(t) = \frac{\pi}{2}\frac{\theta^2}{\theta+1}\sum_{w=1}^{\infty}\frac{(-1)^{w+1}\left(\frac{\pi}{2}\right)^{2w-1}}{(2w-1)!}\sum_{k=0}^{2w+1}\left(\frac{\theta}{\theta+1}\right)^{2w-k+1}\binom{2w-1}{k}\frac{2w(\theta+1)-t-k}{(2w\theta-t)^{2w-k+1}}\Gamma(2w-k) \tag{8}$$

provided $2w\theta > t$.

And we obtain r^th^ moment about origin (i.e. raw moment) of $SS_L(\theta) - distribution$ as follows,

$$\mu_r{}' = \left[\frac{\delta^r M_X(t)}{\delta t^r}\right]_{t=0}$$

3.2 Mean of $SS_L(\theta) - distribution$

If μ is the mean $of\ SS_L(\theta) - distribution$, we have

$$\mu = E(X) = \int_0^\infty x g(x) dx$$

$$\mu = \frac{\pi}{2}\frac{\theta^2}{\theta+1}\int_0^\infty x(x+1)\, e^{-\theta x} \sin\left(\frac{\pi}{2}\left(1+\frac{\theta x}{\theta+1}\right)e^{-\theta x}\right)dx \tag{9}$$

The above integral is not solvable analytically. To solve it numerically for given value θ, one have to use some numerical integration technique such as Gauss-Lagurre quadrature formula or Monte-Carlo simulation or some other methods may be used.

3.3 Median of $SS_L(\theta) - distribution$

If M is the median of the $SS_L(\theta) - distribution$, we have

$$\int_0^\infty g(x) dx = \frac{1}{2}$$

$$\frac{\pi}{2}\frac{\theta^2}{\theta+1}\int_0^\infty (x+1)\, e^{-\theta x} \sin\left(\frac{\pi}{2}\left(1+\frac{\theta x}{\theta+1}\right)e^{-\theta x}\right)dx = \frac{1}{2}$$

After simplification, it reduces to,

$$\cos\left(\frac{\pi}{2}\left(1+\frac{\theta M}{\theta+1}\right)e^{-\theta M}\right) = \frac{1}{2} \tag{10}$$

which is not solvable analytically, some numerical iteration technique will be used for its numerical solution for any specific value of θ.

3.4 Skewness and Kurtosis of $SS_L(\theta) - distribution$

The coefficient of Skewness is a measure of the degree of symmetry of probability distribution (see, Sheskin (2011)). It come in the form of negative skewness or positive skewness, depending on whether data points are skewed to the left or to the right of the data average.

The coefficient of Kurtosis is a measure for the degree of tailendness in the probability distribution (see, Westfall (2014)). There are three categories of kurtosis that can be displayed by the data. First category is a mesokurtic distribution. Second is leptokurtic and third one is platykurtic distribution. The measure of skewness and kurtosis can be calculated using the following expressions

$$\beta_1 = \frac{\left(\mu_3' - 3\mu_2'\mu_1' + 2{\mu_1'}^3\right)^2}{\left(\mu_2' - {\mu_1'}^2\right)^3} \tag{11}$$

$$\beta_2 = \frac{\mu_4' - 4\mu_3'\mu_1' + 6\mu_2'{\mu_1'}^2 - 3{\mu_1'}^4}{\left(\mu_2' - {\mu_1'}^2\right)^2} \tag{12}$$

The values of β_1 and β_2 have been calculated for different values of θand for all considered values of θ, we get $\beta_1 > 0$ and $\beta_2 > 3$. Thus we may conclude that our proposed distribution is positively skewed and leptokurtic. The graphs of values of β_1 and β_2 for different values of θ are shown in figures6 and 7 respectively.

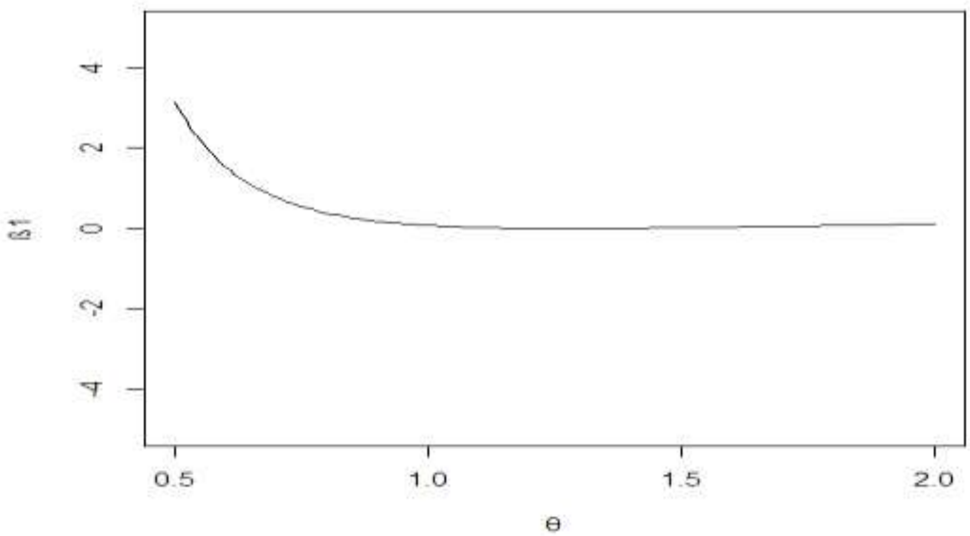


Figure 6: The values of β_1 for different values of θ

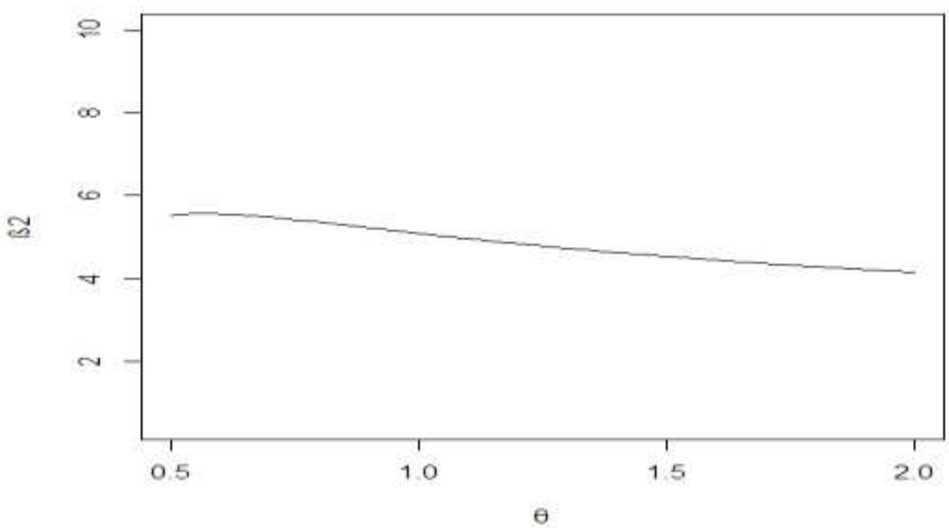


Figure 7: The values of β_2 for different values of θ

4. Estimation of the parameter $\boldsymbol{\theta of}$ $\boldsymbol{SS_L(\theta) - distribution}$

Maximum Likelihood Estimator

Let x=(x$_1$, x$_2$, ...,x$_n$) a sample of size n from $SS_L(\theta) - distribution$having pdf (4). The likelihood function is given by,

$$L = \prod_{i=1}^{n} g(x_i)$$

$$L = \left(\frac{\pi}{2}\right)^n \left(\frac{\theta^2}{\theta+1}\right)^n \prod_{i=1}^{n}(1+x_i)e^{-\theta x_i} \prod_{i=1}^{n} \sin\left(\frac{\pi}{2}\left(1+\frac{\theta x_i}{\theta+1}\right)e^{-\theta x_i}\right)$$

And hence the log likelihood function is,

$$\ln L = n ln\left(\frac{\pi}{2}\right) + 2n\ln(\theta) - n\ln(\theta+1) + \sum_{i=1}^{n}\ln(1+x_i) - \theta\sum_{i=1}^{n}x_i + \sum_{i=1}^{n}\ln\left\{sin\left(\frac{\pi}{2}\left(1+\frac{\theta x_i}{\theta+1}\right)e^{-\theta x_i}\right)\right\} \qquad (13)$$

Now, the log likelihood equation for estimating θ is given by,

$$\frac{\partial \ln L}{\partial \theta} = 0 \qquad (14)$$

This is not solvable analytically for θ and thus we have used nlm, an R-code to solve it analytically on R-software

.

5. Real data application

In this section, a real data set has been considered for checking applicability, suitability and superiority of the proposed distribution over on some existing distributions such as $SS_{EXP}(\theta) - distribution$, $DUS_{Lindley}(\theta)$-distribution and Lindley Distribution. Consider following real data which represents the number of million revolutions before failure for each of 23 ball bearings in a life testing experiment, extracted from Lawless (2011).

17.88, 28.92, 33, 41.52, 42.12, 45.6, 48.4, 51.84, 51.96, 54.12, 55.56, 67.8, 68.64, 68.64, 68.88, 84.12, 93.12, 98.64, 105.12, 105.84, 127.92, 128.04, 173.4.

To check the validity of the considered model for the above data set, we have calculated various statistical tools such as Kolmogrov-Smirnov (K-S) test statistic, Akaike information criterion (AIC) and Bayesian information criterion (BIC). These criterions are defined as follows

$$AIC = -2\ln(\hat{L}) + 2p$$
$$BIC = -2\ln(\hat{L}) + p\ln(n)$$
$$D_n = \sup|F(x) - F_n(x)|$$

where n is the sample size, p is the no. of unknown parameters in the model, $\hat{L}$ is the maximized value of the likelihood function and $F_n(x)$ (x) is empirical distribution function.

Table 1: The values of AIC, BIC, K-S test statistic with its p-value for considered real data set

	Estimate	AIC	BIC	KS	P-value
$SS_{EXP}(\theta)$	0.007911357	242.8583	243.9938	0.29532	0.0362
$SS_{Lindley}(\theta)$	0.0187719	231.4266	232.5621	0.16423	0.5644
$Lindley(\theta)$	0.02732466	233.4707	234.6062	0.19286	0.3593
$DUS_{Lindley}(\theta)$	0.01725388	239.1521	240.2875	0.93389	2.20E-16

We have computed MLE of the parameter θ of $SS_L(\theta) - distribution$having pdf (4) for above data set and found it as 0.0187719 and consequently, AIC, BIC and K-S test value with its p- value for $SS_L(\theta) - distribution$are calculated and their values are shown in Table 1. For other distributions, viz $SS_{EXP}(\theta)$, $DUS_{Lindley}(\theta)$and Lindley(θ) distributions, the value of AIC, BIC and K-S test value with its p-value are calculated for the above data set and are shown in comparative Table 1.

From table 1, it is seen that the criterion values AIC, BIC and also K-S test value of our proposed distribution $SS_L(\theta) - distribution$is smallest as compared to those of $SS_{EXP}(\theta)$, $DUS_{Lindley}(\theta)$, and Lindley(θ) distributions and has largest p-value too. Thus we may conclude that $SS_L(\theta) - distribution$ is well fitted to the considered data set as compared to $SS_{Exp}(\theta), DUS_{Lindley}\,(\theta)$and Lindley(θ) distributions.

The plots of empirical cdf Fn and fitted cdf G(x) of $SS_L(\theta)$ −distribution having pdf (4) for above data set shown in Figure 8.

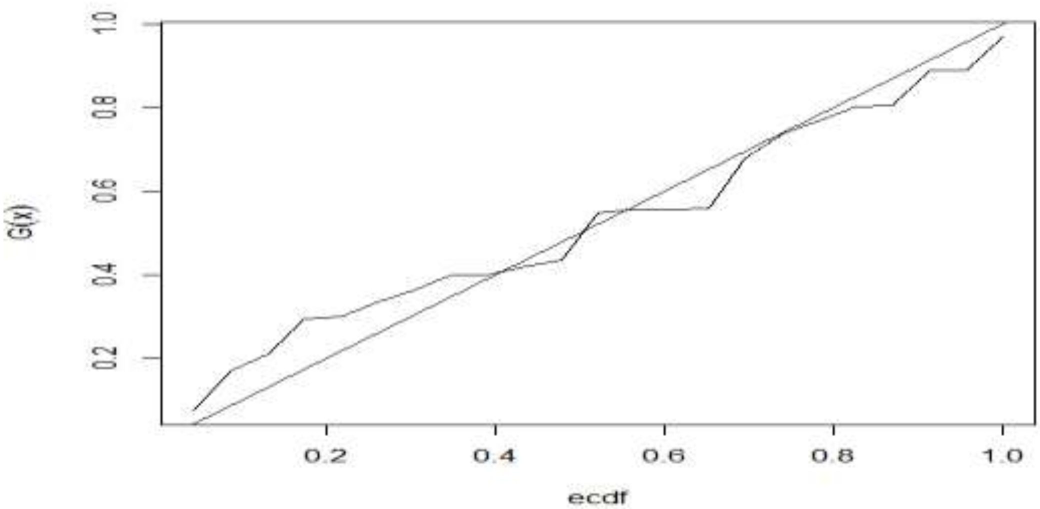


Figure 8: Plot of ecdf of $SS_L(\theta) - distribution$ for given data set

6. Simulation Study

A simulation study is carried out to study the performance of MLEs of θ of proposed model. We consider mean square errors (MSEs) of MLEs of the parameter θ, reliability R≡R(t) and hazard rate h≡h(t) for a specific value of t and asymptotic CI of the parameter θ.

Table 2: Mean Square Error of MLEs of parameter θ, Reliability, Hazard Rate Function at t=0.5 and Confidence Interval for parameter θ

n	$MSE(\hat{\theta})$	$MSE(\hat{R})$		$MSE(\hat{h})$		95 % CI of θ
		t=0.5	t=1	t=0.5	t=1	
10	0.001700994	0.000371909	0.0009872	0.003178	0.00441178	(0.3207599 ,0.7617263)
15	0.001011944	0.000217979	0.0006159	0.001771	0.00248071	(0.3551504 , 0.7084717)
20	0.000803559	0.00017035	0.0004986	0.001347	0.00189602	(0.3764446, 0.6802496)
30	0.000603145	0.000125231	0.0003804	0.000963	0.00136155	(0.4014849, 0.6476331)
80	0.000409562	8.20E-05	0.0002637	0.000604	0.0008594	(0.4455370 , 0.5949382)
140	0.000340706	6.75E-05	0.0002204	0.000492	0.00070106	(0.4621882 ,0.5747282)
200	0.000334482	6.69E-05	0.0002203	0.000491	0.00070135	(0.4714691,0.5656514)

Table 3: Mean Square Error of MLEs of parameter θ, Reliability, Hazard Rate Function at t=1 and Confidence Interval for parameter θ

n	$MSE(\hat{\theta})$	$MSE(\hat{R})$		$MSE(\hat{h})$		95 % CI of θ
		t=0.5	t=1	t=0.5	t=1	
10	0.005138153	0.000691209	0.0003998	0.017109	0.0197296	(0.6114234 1.5319384)
15	0.002542872	0.000367103	0.0002407	0.008273	0.00966946	(0.6830099, 1.4178439)
20	0.002033988	0.000315523	0.0002557	0.006332	0.00753446	(0.7288118,1.3613877)
30	0.001155612	0.000186872	0.0001676	0.003512	0.00422427	(0.7787723 1.2892163)
80	0.000672169	0.000120234	0.0001452	0.001877	0.00233	(0.8710001 1.1808524)
140	0.000500426	9.19E-05	0.0001197	0.001362	0.00170659	(0.905702 , 1.139038)
200	0.000439904	8.18978e-05	0.0001107	0.001182	0.00148761	(0.9235093 ,1.1184385)

Tables 2 and 3 shows that the performance of the estimators for the proposed model$SS_L(\theta) - distribution$. In this regard, we have calculated the values of mean square error (MSE) for the parameterθ, reliability and hazard rate function (for t=0.5 and 1) and 95% Confidence interval for varying of sample size n=(10,15,20,30,80,140,200) for the true parameter $\theta = 0.5$ and 1 respectively and it is observed that MSEs of the MLEs of the parameter θ, Reliability and Hazard as well as the width of CI of the parameter θ are decreases as sample size increases.

7. Conclusion

In the present piece of work, a new distribution using SS transformation as suggested by Kumar et al. (2015) is introduced and the considered baseline distribution is Lindley(θ)-distribution. Some statistical properties of this new distribution has been studied and it is observed that it is suitable for data having decreasing failure rate. A real data application shows that it fits well as compared to $SS_E(\theta) - distribution$, $DUS_L(\theta) - distribution$ and Lindley(θ) distribution and the real data set represents the no. of revaluations before failure rate of each of 23 ball bearings. Thus we may recommend further application of our proposed distribution to draw exact and valid inferences in engineering, science and technology.

References

[1] Lindley, D. (1958): Fiducial distributions and Bayes theorem .J.R.Stat.Soc.Ser.B 20:102107.
[2] Kumar, D., Singh, S.K. and Singh, U. (2017): Life Time Distributions: Derived from some Minimum Guarantee Distribution. Sohag J. Math. 4, No. 1, 7-11.
[3] Kumar et al. (2015 (a)). A New Distribution Using Sine Function-Its Application To Bladder Cancer Patients Data. Journal of Statistics Applications and Probability, 4(3), 417.
[4] Lawless, J. F. (2011). Statistical models and methods for lifetime data (Vol. 362).John Wiley and Sons.
[5] Kumar et al. (2015 (b)). A method of proposing new distribution and its application to Bladder cancer patients data. J. Stat. Appl. Pro. Lett, 2(3), 235-245.
[6] Ghitany, M. E., Al-Mutairi, D. K. and Nadarajah, S. (2008): Zero-truncated poisson-lindley distribution and its application, Mathematics and Computers in Simulation 79(3): 279287.
[7] Ghitany, M. E., Atieh, B. and Nadarajah, S. (2008): Lindley distribution and its application, Mathematics and computers in simulation 78(4): 493-506.
[8] Chesneau, C., Bakouch, H. S., and Hussain, T. (2018). A new class of probability distributions via cosine and sine functions with applications. Communications in Statistics-Simulation and Computation, 1-14.
[9] Nadarajah, S., Bakouch, H. S. and Tahmasbi, R. (2011): A generalized lindley distribution, Sankhya B 73(2): 331-359.
[10] Potdar, K. G. (2016). Reliability estimation for the generalized inverted Rayleigh distribution. Electronic Journal of Applied Statistical Analysis, 9(1), 154-168.
[11] Gupta, R. C., Gupta, R. D. and Gupta, P. L. (1998): Modeling failure time data by Lehman alternatives. Communication in Statistics-Theory and Methods, 27 (4), 887-904.
[12] Shaw, W. T. and Buckley, I. R. (2007): The alchemy of probability distributions: beyond Gram-Charlier expansions and a skew-kurtotic-normal distribution from a rank transmutation map. Research report.
[13] Merovci, F., and Elbatal, I. (2013). Transmuted Lindley-geometric distribution and its applications. arXiv preprint arXiv:1309.3774.
[14] Oluyede, B. O., and Yang, T. (2015). A new class of generalized Lindley distributions with applications. Journal of Statistical Computation and Simulation, 85(10), 2072-2100.
[15] Asgharzadeh, A., Bakouch, H. S., and Esmaeili, L. (2013). Pareto PoissonLindley distribution with applications. Journal of Applied Statistics, 40(8), 1717-1734.
[16] Al-Babtain, A., Eid, A. M., Ahmed, A. H. N., and Merovci, F. (2015). The five parameter Lindley distribution. Pakistan Journal of Statistics, 31(4).
[17] Crawley, M. J. (2012). The R book. John Wiley & Sons.