UMYU Scientifica, Vol. 1 NO. 1, September 2022, Pp 127 – 131
https://doi.org/10.56919/usci.1122.017
ISSN: 2955 – 1145 (print); 2955 – 1153 (online)
*A periodical of the Faculty of Natural and Applied Sciences, UMYU, Katsina*

**ORIGINAL RESEARCH ARTICLE**

# Sine-Exponential Distribution: Its Mathematical Properties and Application to Real Dataset

[1]Isa Alhaji Modu, [2]Sule Omeiza Bashiru, [1]Ali Buhari Alhaji, [3]Akeem Ajibola Adepoju and [2]Ibrahim Ismaila Itopa

[1]Department of Mathematics and Computer Science, Borno State University, Maiduguri, Nigeria
[2]Department of Mathematical Sciences, Prince Abubakar Audu University, Anyigba, Kogi State, Nigeria.
[3]Department of Statistics, Kano University of Science and Technology, Wudil, Kano State, Nigeria

## ABSTRACT

To increase flexibility or to develop covariate models in a variety of ways, new parameters can be added to existing families of distributions or a new family of distributions can be compounded with well-known standard normal distribution. In this paper, a trigonometric-type distribution was developed in order to come up with flexible distribution without adding parameter, considering Exponential distribution as the baseline distribution and Sine-G as the generator. The proposed distribution is referred to as Sine Exponential Distribution. Statistical features including the moment, moment generating function, entropy, and order statistics were obtained. The proposed distribution's parameters were estimated using the Maximum Likelihood method. Using real datasets, the model's importance was demonstrated. The newly developed model was proven to be better than its competitors.

**ARTICLE HISTORY**

Received August 22, 2022
Accepted September 27, 2022
Published September 30, 2022

**KEYWORDS**

Sine G family, Exponential Distribution, Sine Exponential Distribution, Moment, Maximum likelihood estimators



## INTRODUCTION

Researchers in the field of statistics have shown a keen interest in developing flexible distributions using generalization or compounding methods. Over the years, many generalized or compounded distributions have been proposed, some of them are: Gompertz Inverse Rayleigh distribution by Halid and Sule (2022) which was developed by extending the inverse Rayleigh distribution using the Gompertz generated family of distribution; Sine BurrXII proposed by Isa *et al.,* (2022) was derived by extending the Burr XII distribution using the Sine G family; Exponentiated Odd Lomax Exponential Distribution by Dhugana and Khumar (2022) which was proposed by compounding an Exponential Odds Function and Lomax Generated family of distributions. Ogunsanya *et al.,* (2021) introduced the Weibull Inverse Rayleigh Distribution by compounding the Inverse Rayleigh Distribution with the Weibull Generated family of distributions. Ahmad *et al.*, (2021) developed the Topp-Leone Power Rayleigh distribution by substituting the Power Rayleigh Distribution into the Topp-Leone family of distribution. Oguntunde *et al.,* (2018) developed the "Gompertz Inverse Exponential Distribution" by compounding the inverse exponential distribution with the Gompertz G family of distribution.

Despite the development of these generalized distributions, there are emerging data of interest that exhibit non-normal features like very high skewness and kurtosis. Thus, there is the need to develop more generalized or compound distributions that will have the ability of handling these emerging data of recent times with the stated features. By compounding a well-known standard distribution with a generated family of distribution, the model will be more flexible with high level of skewness and kurtosis to enable the generalized or compounded model have the capability of properly modelling data sets that are heavy tailed or leptokurtic (these are data sets with kurtosis greater than 3).
Some of the methods used in developing these flexible models include: Exponentiated family of distribution proposed by Gupta et al. (1998), Beta generated families by Eugene et al.(2002), Transform Transformer (T-X)

**Correspondence:** Isa A. M. Department of Mathematics and Computer Science, Borno State University, Maiduguri, Nigeria. ✉ alhajimoduisa@bosu.edu.ng; Phone Number: 08066866166
**How to cite**: Isa A. M., Sule O. B., Ali B. A., Akeem A. A., and Ibrahim I. I. (2022). Sine-Exponential Distribution: Its Mathematical Properties and Application to Real Dataset. UMYU Scientifica, 1(1), 127 – 131. https://doi.org/10.56919/usci.1122.017

family by Alzaatreh et al. (2013) and Transmuted family by Shaw and Buckley (2009). Researchers moved farther to have more options in terms of the availability of distributions that focus on trigonometric functions due to the rise in interest in data analysis. However, there aren't many distributions that incorporate trigonometric functions. (Muhammad *et al.,* 2021). This study focuses on a novel compound distribution called the Sine Exponential Distribution and its application to real datasets.

The Sine G family of distribution is quite new and very adaptable. The Exponential distribution will be compounded with Sine generated family of distributions. The Exponential distribution is one of the most extensively used models that can be widely applied to various fields like engineering, medicine, agriculture and so on.

## MATERIALS AND METHODS

### The Sine G Family

Let H(x) represent the cumulative distribution function (cdf) of a univariate continuous distribution and h(x) represent the corresponding probability density function (pdf).Then, according to Kumar et al. (2015), the Sine-G family of probability distributions is given by:

$$F(x,\xi) = \int_0^{\frac{\pi}{2}H(x,\xi)} \cos t \, dt = \sin\left\{\frac{\pi}{2}H(x,\xi)\right\} \quad (1)$$

and the associated pdf of the Sine-Family of distributions is:

$$f(x,\xi) = \frac{\pi}{2}h(x,\xi)\cos\left\{\frac{\pi}{2}H(x,\xi)\right\} \quad (2)$$

where the cdf and pdf of any baseline distribution with vector parameter $\xi$ are $H(x,\xi)$ and $h(x,\xi)$.

### The Sine-Exponential Distribution

The newly developed Sine Exponential Distribution's PDF is given by:

$$f(x;\lambda) = \frac{\pi}{2}\,\lambda\, e^{-\lambda x}\cos\left\{\frac{\pi}{2}\left(1-e^{-\lambda x}\right)\right\} \quad (3)$$

with corresponding cdf given as:

$$F(x) = \sin\left\{\frac{\pi}{2}\left(1-e^{-\lambda x}\right)\right\}, \quad \lambda, x > 0 \quad (4)$$

Below are the quantile function Q(u), cumulative hazard function H(x), reverse hazard function r(x), survival function S(x), and hazard functions h(x) of the sine-exponential distribution.

$$S(x) = 1-\sin\left\{\frac{\pi}{2}\left(1-e^{-\lambda x}\right)\right\} \quad (5)$$

$$h(x) = \frac{\frac{\pi}{2}\,\lambda\, e^{-\lambda x}\cos\left\{\frac{\pi}{2}\left(1-e^{-\lambda x}\right)\right\}}{1-\sin\left\{\frac{\pi}{2}\left(1-e^{-\lambda x}\right)\right\}} \quad (6)$$

$$r(x) = \frac{\pi}{2}\,\lambda\, e^{-\lambda x}\cos\left\{\frac{\pi}{2}\left(1-e^{-\lambda x}\right)\right\}\cot\left\{\frac{\pi}{2}\left(1-e^{-\lambda x}\right)\right\} \quad (7)$$

$$H(x) = -\ln\left\{1-\sin\left[\frac{\pi}{2}\left(1-e^{-\lambda x}\right)\right]\right\} \quad (8)$$

$$Q(u) = F^{-1}\left\{\left[-\frac{1}{\lambda}\log\left(1-\frac{\sin^{-1}u}{\pi/2}\right)\right]\right\} \quad (9)$$

### Mixture Representation

Expansion of the pdf

$$f(x) = \frac{\pi}{2}\,\lambda e^{-\lambda x}\cos\left[\frac{\pi}{2}\left(1-e^{-\lambda x}\right)\right]$$

Expanding $e^{-\lambda x}$ gives:

$$e^{-\lambda x} = \sum_{i=0}^{\infty}\frac{(-1)^i}{i!}(\lambda x)^i$$

And also $\cos\left[\frac{\pi}{2}\left(1-e^{-\lambda x}\right)\right]$ can be expressed as follows:

$$\cos\left[\frac{\pi}{2}\left(1-e^{-\lambda x}\right)\right] = \sum_{j=0}^{\infty}\frac{(-1)^j}{(2j)!}\frac{\pi^{2j}}{2j}\left(1-e^{-\lambda x}\right)^j$$

$\left(1-e^{-\lambda x}\right)^j$ can also be expressed as:

$$\left(1-e^{-\lambda x}\right)^j = \sum_{k=0}^{\infty}(-1)^k\binom{j}{k}e^{-k\lambda x}$$

$$e^{-k\lambda x} = \sum_{l=0}^{\infty}\frac{(-1)^l}{l!}(k\lambda x)^l$$

$$f(x) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\frac{\pi^{2j+1}}{2^{2j1}}\frac{(-1)^{i+j+k+l}}{2j!\,i!\,l!}\,k\lambda^{i+l+1}\binom{j}{k}x^{i+l}$$

$$\Psi = \frac{\pi^{2j+1}}{2^{2j1}}\frac{(-1)^{i+j+k+l}}{2j!\,i!\,l!}\,k\lambda^{i+l+1}\binom{j}{k}$$

Therefore, the pdf can be expressed as

$$f(x) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\Psi\, x^{i+l} \quad (10)$$

The expansion of the cdf:

$$F(x) = \sin\left[\frac{\pi}{2}\left(1-e^{-\lambda x}\right)\right] = \sum_{p=0}^{\infty}\frac{(-1)^p}{(2p+1)!}\frac{\pi^{2p+1}}{2^{2p+1}}\left(1-e^{-\lambda x}\right)^{2p+1}$$

$$\left(1-e^{-\lambda x}\right)^{2p+1} = \sum_{q=0}^{\infty}(-1)^q\binom{2p+1}{q}e^{-q\lambda x}$$

$$e^{-q\lambda x} = \sum_{r=0}^{\infty} \frac{(-1)^r}{r!} (q\lambda x)^r$$

$$F(x) = \sum_{p=0}^{\infty}\sum_{q=0}^{\infty}\sum_{r=0}^{\infty} \frac{(-1)^{p+q+r}}{r!\,(2p+1)!} \frac{\pi^{2p+1}}{2^{2p+1}} \binom{2p+1}{q} (q\lambda)^r x^r$$

Let $\Omega = \frac{(-1)^{p+q+r}}{r!(2p+1)!} \frac{\pi^{2p+1}}{2^{2p+1}} \binom{2p+1}{q} (q\lambda)^r$

Therefore, the cdf can also be expressed as:

$$F(x) = \sum_{p=0}^{\infty}\sum_{q=0}^{\infty}\sum_{r=0}^{\infty} \Omega\, x^r \quad (11)$$

**Some Mathematical Properties**

The $r^{th}$ moment, moment generating function, entropy, and order statistics were some of the mathematical properties that were obtained, and they are given below:

**Moment**

The Sine-Exponential Distribution $r^{th}$ moment was derived in this section. The $r^{th}$ moment of the random variable *X* with pdf $f(x)$ can be obtained as follows:

$$\mu_r' = \int_{-\infty}^{\infty} x^r f(x)dx = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty} \Psi \int_0^{\infty} x^{i+l}\,dx$$

$$Let\ \rho = \int_0^{\infty} x^{i+l}\ dx$$

Therefore, the $r^{th}$ moment of the Sine-Exponential distribution is given by:

$$\mu_r' = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty} \rho\, \Psi \quad (12)$$

**Moment Generating Function**

The moment generating function for a random variable X is the expected value of $e^{tx}$. If a positive constant "a" exists such that M x (t) is finite for $\forall t \in [-a, a]$, then the moment generating function is said to exist.

$$M_x(t) = E(e^{tx}) = \int_{-\infty}^{\infty} e^{tx} f(x)dx$$

$$M_x(t) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty} \Psi \int_0^{\infty} e^{tx} x^{i+l}\,dx$$

$$Let\ \Upsilon = \int_0^{\infty} e^{tx} x^{i+l}\ dx$$

Therefore, the Moment Generating Function of the Sine-Exponential distribution is given by:

$$\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty} \Psi\, \Upsilon \quad (13)$$

**Entropy**

Entropy is a metric indicating how much information or uncertainty there is in a random observation of a population's real composition. The amount of uncertainty in the data will increase if entropy is large. Following is the formula for the Shannon entropy of a continuous random variable X:

$$I_\theta(x) = \frac{1}{1-\theta} \log \int_{-\infty}^{\infty} f(x)^\theta dx$$

$$f(x)^\theta = \left( \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty} \Psi\, x^{i+l} \right)^\theta$$

$$f(x)^\theta = \left( \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty} \Psi \right)^\theta \left(x^{i+l}\right)^\theta$$

$$Let\ \beta = \left(x^{i+l}\right)^\theta$$

Then,

$$I_\theta(x) = \frac{1}{1-\theta} \left[ \left( \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty} \Psi \right)^\theta \log \int_{-\infty}^{\infty} \beta^\theta dx \right] \quad (14)$$

**Order Statistics**

Let $X_1, X_2, \ldots, X_n$ be a random samples of size $n$ from a continuous population having pdf $f(x)$ and distribution function $(cdf)$ F(x), Let $X_{1:n} \leq X_{2:n} \leq X_{n:n}$ be the associated order statistics (OS). David (1970) defined the pdf of $X_{r:n}$ that is the $r^{th}$ OS by:

$$f_{r:n}(x) = \frac{\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty} \Psi\, x^{i+l}}{B(m, n-m+1)} \left( \sum_{p=0}^{\infty}\sum_{q=0}^{\infty}\sum_{r=0}^{\infty} \Omega\, x^r \right)^{m-1} \left( 1 - \sum_{p=0}^{\infty}\sum_{q=0}^{\infty}\sum_{r=0}^{\infty} \Omega\, x^r \right)^{n-m} \quad (15)$$

**Parameter Estimation**

The technique of maximum likelihood was explored to estimate the Sine-Exponential Distribution's unknown parameters. Maximum likelihood estimates (MLEs) are attractive because they are used to produce confidence intervals and offer straightforward approximations that perform well with finite samples. You can quickly work with approximations of MLE results in distribution theory using analytical or numerical techniques. Let $x_1, x_2, x_3, \ldots, x_n$ be a random samples of size $n$ from the Sine-Exponential Distribution. Then, the likelihood function of the Sine-Exponential Distribution is given as follows:

$$l = n\log\left(\frac{\pi}{2}\right) + n\log\lambda - \lambda\sum_{i=0}^{n} x + \sum_{i=0}^{n}\log\cos\left\{\frac{\pi}{2}\left[1-e^{-\lambda x}\right]\right\} \quad (16)$$

Differentiating equation (16) with respect of $\lambda$ gives the following expression:

$$\frac{\partial l}{\partial \lambda} = \frac{n}{\lambda} + \sum_{i=0}^{n} x - \sum_{i=0}^{n}\tan\left\{\frac{\pi}{2}\left[1-e^{-\lambda x}\right]\right\} \quad (17)$$

Equation (17) gives the estimate of the maximum likelihood of the parameter$\lambda$ of the Sine-Exponential Distribution.

## RESULTS

### Application

Breaking stress of carbon fibres data of sample 100 observations was considered to test for the efficiency of the model. This data set was used by Joseph *et al.,* (2020). The result of the analysis was compared with the baseline distribution (exponential) and Lomax distributions for illustrative purposes. We estimated the unknown parameters of each distribution by the maximum-likelihood method. The values of the Akaike information criterion (AIC), Bayesian information criterion (BIC) and consistent Akaike information criterion (CAIC) were also computed and compared with the newly developed model.

The data are as follows:

3.7, 2.74, 2.73, 2.5, 3.6, 3.11, 3.27, 2.87, 1.47, 3.11, 4.42, 2.41, 3.19, 3.22, 1.69, 3.28, 3.09, 1.87, 3.15, 4.9, 3.75, 2.43, 2.95, 2.97, 3.39, 2.96, 2.53, 2.67, 2.93, 3.22, 3.39, 2.81, 4.2, 3.33, 2.55, 3.31, 3.31, 2.85, 2.56, 3.56, 3.15, 2.35, 2.55, 2.59, 2.38, 2.81, 2.77, 2.17, 2.83, 1.92, 1.41, 3.68, 2.97, 1.36, 0.98, 2.76, 4.91, 3.68, 1.84, 1.59, 3.19, 1.57, 0.81, 5.56, 1.73, 1.59, 2, 1.22, 1.12, 1.71, 2.17, 1.17, 5.08, 2.48, 1.18, 3.51, 2.17, 1.69, 1.25, 4.38, 1.84, 0.39, 3.68, 2.48, 0.85, 1.61, 2.79, 4.7, 2.03, 1.8, 1.57, 1.08, 2.03, 1.61, 2.12, 1.89, 2.88, 2.82, 2.05, 3.65

Table 1 gives the summary statistics of the data set such as the mean, the median, the first and third quartile, the minimum and the maximum values while table 2 gives the values of the estimates of the parameters of the models, their AIC, CAIC, BIC and HQIC respectively.

**Table 1: Summary statistics of the datasets**

| Data | Minimum | $Q_1$ | Median | Mean | $Q_3$ | Maximum |
|---|---|---|---|---|---|---|
| Dataset | 0.390 | 1.840 | 2.700 | 2.621 | 3.220 | 5.560 |

**Table 2: MLE, AIC, CAIC, BIC, and HQIC of the dataset**

| Distribution | $\lambda$ | $\gamma$ | $MLE$ | $AIC$ | $CAIC$ | $BIC$ | $HQIC$ |
|---|---|---|---|---|---|---|---|
| Sexpo | 0.21884 | – | 191.2010 | 384.4021 | 384.4429 | 391.6124 | 385.4565 |
| Expo | 0.3814453 | | 196.3701 | 394.7417 | 394.7825 | 401.9520 | 398.8504 |
| Lomax | 3119957 | 8176557 | 196.3079 | 396.7417 | 396.8654 | 401.9521 | 398.8504 |

Table 2 gives the results of the analysis of the dataset. The results of the analysis of the Sine-Exponential was compared with Exponential and Lomax distribution. The proposed Sine-Exponential distribution proven to be the better model because it has the least AIC, CAIC, BIC and HQIC. In other words, the proposed model outperformed its competitors with minimum AIC, BIC, CAIC and HQIC.

## CONCLUSION

There has been increase of interest among statisticians and applied researchers in developing more flexible lifetime distributions for the furtherance of modeling survival data. In this paper, we introduced a one parameter Sine-Exponential distribution which was achieved by considering Exponential distribution as the baseline and Sine G family. We study some of its mathematical properties. Maximum likelihood Estimation was used in parameter estimation. The potential of the Sine Exponential distribution was demonstrated using real data, and comparisons with other competing distributions revealed that it could provide a better fit. We anticipate that the extended approach that has been provided will find more uses in survival analysis.

## ACKNOWLEDGEMENT

We gratefully acknowledge the effort of the editors and the reviewers for their excellent comments on this manuscript, which helped to improve it.

## REFERENCES

Ahmed, A., Ahmad, A. & Tripathi, R. (2021). "Topp-Leone Power Rayleigh Distribution with Properties and Application in Engineering Science". *Advances and Applications in Mathematical Sciences, 20*(11), 2852-2877.

Alzaatreh, A., Lee, C., & Famoye, F. (2013). "A new method for generating families of continuous distributions". *Metron*, *71*(1), 63–79. **[Crossref]**

Bhuamik D.K, Kapur K. & Gibbons R.D. (2009). "Testing Parameters of a Gamma Distribution for Small Samples". *Technometrics, 51*(3), 326 - 334. **[Crossref]**

Cordeiro, G. M., Ortega, E. M., & da Cunha, D. C. (2013). "The exponentiated generalized class of distributions". *Journal of data science*, *11*(1), 1–27. **[Crossref]**

David, H. A. (1970). Order Statistics, Wiley, New York.

Dhugana, G.P., & Kumar, V. (2022). "Exponentiated Odd Lomax Exponential Distribution with Application to Covid – 19 death cases of Nepal". *PloS ONE, 17*(6), 1-26. **[Crossref]**

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

Gupta, R. C., Gupta, P. L., & Gupta, R. D. (1998). "Modeling failure time data by Lehman alternatives". *Communications in Statistics-Theory and methods*, *27*(4), 887–904. **[Crossref]**

Halid, O.Y. & Sule, O. B. (2022). "Classical and Bayesian Estimation Techniques for Gompertz Inverse Rayleigh Distribution: Properties and Application". *Pakistan Journal of Statistics 38*(1), 49 – 76.

Isa, A. M., B. A., & U, Zannah, (2022). "Sine Burr XII Distribution: Properties and Application to Real Data Sets". *Arid Zone Journal of Basic and Applied Research, 1*(3), 48 – 58. **[Crossref]**

Kumar, D., Singh, U. & Singh, S. K. (2015). A New Distribution Using Sine Function. Its Application to Bladder Cancer Patients Data. Journal of Statistics and Applied Probabilit, 4(3), 417-427.

Muhammad, M., Alshanbari, H.M. Alanzi, A.R.A., Liu, L., Sami, W., Chesneau, C., Jamal, F. A. (2021). "New Generator of Probability Models: The Exponentiated Sine-G Family for Lifetime Studies". *Entropy, 23*, 1394.**[Crossref]**

Ogunsanya A. S., Yahaya, W. B., Adegoke, T. M., Iluno, C., Aderele, O. R. & Ekum, M. I. (2021). "A New Three – Parameter Warameter Inverse Rayleigh Distribution: Theoretical Development and Applications". *Mathematics and Statistics, 9*(3), 248-272. **[Crossref]**

Oguntunde, P. E., Khaleel, M. A., Adejumo, A. O., Okagbue, H. I., Opanuga, A. A. & Owolabi, F. O. (2018). "The Gompertz Inverse Exponential distribution with application". *Congent Mathematics & Statistics, 5,* 1507122: 1 - 11. **[Crossref]**

Shaw, W. T. & Buckley, I. R. (2009). "The Alchemy of probability distributions: Beyond Gramcharlier Expansions, and a Skew-Kurtotic-Normal Distribution from a Rank Transmutation Map". *arXiv preprint arXiv:0901.0434.*