Sine New Weighted Exponential Distribution, Properties and Applications to Datasets from Rainfall and Breast Cancer Studies

Umar Kabir Abdullahi[1], Jamilu Yunusa Falgore[2], Ibrahim Yusuf Ibrahim[1], and Saddam Adams Damisa[2].

[1]*Department of Mathematical Sciences, Bayero University, Kano*

[2]*Department of Statistics, Ahmadu Bello University, Zaria*

[*]*Corresponding author:* ukabdullahi2@gmail.com

Abstract

The exponential distribution, defined by its constant failure rate, was first expanded into a two-parameter generalised-exponential form through the Lehmann type I alternative, and later, a technique originally used to introduce skewness to the normal distribution was adapted to create a class of weighted exponential distributions capable of modelling positively skewed and hidden truncated data. Although a subsequent version of this weighted distribution was made mathematically simpler and easier to handle, the foundations laid by these developments ultimately led to the development of the Sine New Weighted Exponential Distribution (SNWED), which specifically adopted the methodology of Sin-G in its construction. Some of the properties of the SNWED, such as mean, quantile function, skewness, kurtosis, and median, were derived. The estimation of the parameters of the SNWED was done using the method of maximum likelihood. SNWED was fitted to breast cancer survival times and rainfall datasets alongside comparators, yet SNWED outperformed all of them. The superiority of the SNWED was clearly established when it outperformed existing competing distributions across two tested datasets, demonstrating its practical usefulness and supporting its recommendation for fitting datasets of a similar kind. The study suggests developing Bayesian inference procedures for the SNWED parameters, exploring other methods of parameter estimation (e.g., ordinary least squares, weighted least squares), as well as extending the SNWED to accommodate censored data and covariates (regression framework).

Keywords: Distribution, Estimation, Likelihood, Sine G**,** Weighted-Exponential

1. Introduction

The time between series of events in a Poisson process is described by the well-known single-parameter exponential distribution, which only has a constant failure rate (Lawless, 2003). It has also been employed extensively in many different contexts. An expansion of the exponential distribution based on the Lehmann type I alternative is the two-parameter generalised-exponential (GE) distribution developed by Gupta and Kundu (1999). The skew-normal distribution was developed by Azzalini (1985) using a novel technique for adding a skewness parameter to the normal distribution based on a weighted function. Many skew symmetric distributions, such as the Nadarajah skew-logistic distribution, have been produced using Azzalini's theory applied to other symmetric distributions. To develop a new class of weighted exponential distributions, which are generalisations of the exponential distribution that resemble the Weibull, gamma, and exponentiated exponential distributions, Gupta and Kundu (2009) followed a methodology similar to that of Azzalini (1985) for adding a skewness parameter to the exponential distribution. The study showed that the weighted exponential distribution can be used to analyse positively skewed data, like the distributions mentioned above, and data coming from hidden truncated models.

Bashir and Naqvi (2016) employed Azzalini's Methodology to develop a new weighted exponential distribution; in comparison to the weighted exponential distribution created by Gupta and Kundu (2009), it is more straightforward and easier to handle mathematically.

A random variable X is said to follow a New Weighted Exponential distribution by Bashir and Naqvi (2016) with parameters α *and* β if its probability density function (*pdf*) is given as

$$g(x;\alpha,\beta)=(\alpha-\beta)e^{-(\alpha-\beta)x} \qquad (1).$$

with corresponding cumulative distribution function (*cdf*) as

$$G(x;\alpha,\beta)=1-e^{-(\alpha-\beta)x} \qquad (2).$$

for $x>0,\alpha>0,0<\beta<1$ where α *and* β are the rate and weight parameters, respectively.

Since its introduction, the NWED has served as the baseline for several further generalisations aimed at improving its flexibility for skewed and heavy-tailed data. Abdullahi and Abba (2017) extended the NWED using the odd generalised exponential family to produce the Odd Generalised Exponential-New Weighted Exponential distribution, while Umar *et al.* (2018) incorporated the exponentiated Weibull generator to obtain the Exponentiated New Weighted Weibull distribution. More recently, Abdullahi *et al.* (2023) applied a transmutation approach to the NWED to develop the Transmuted New Weighted Exponential distribution (TNWED), which is used later in this study as one of the comparator models. These developments illustrate the continued relevance of the NWED as a flexible baseline for constructing new lifetime distributions, and motivate its combination with the Sin-G family in the present study.

Some extensions of the Exponential distribution and transmuted distributions include the odd generalized exponential-inverse lomax by Falgore *et al.* (2018), the Lomax-exponential by Ieren and Kuhe (2018), the transmuted odd generalized exponential-exponential by Abdullahi *et al*. (2018), generalized transmuted-inverted exponential by Usman *et al.* (2019), the odd generalized exponential-exponential by Maiti and Pramanik (2015), exponential new weighted Weibull by Umar et al. (2018), transmuted generalized linear exponential by Elbatel *et al.* (2013), transmuted exponential Lomax by Abdullahi and Ieren (2018), Odd Generalized Exponential-New Weighted Exponential by Abdullahi and Abba (2017) and transmuted Rayleigh by Merovci (2013). Khan *et al*. (2017) presented a three-parameter transmuted generalised exponential distribution and tested the potentiality of the distribution using survival data. The transmuted linear exponential distribution was introduced by Tian *et al.* (2014), which is a generalisation of the linear exponential distribution.

The development of trigonometric families of distributions has attracted considerable attention as a means of introducing additional flexibility into baseline models without increasing the number of parameters. Souza (2015) first proposed the use of the sine function as a generator of new distributions, an idea later formalised by Souza *et al.* (2019) as the Sin-G class, whose cumulative distribution function is obtained by composing the sine function with the cumulative distribution function of a baseline distribution. Because the sine function maps the interval [0, π/2] onto [0, 1] in the same manner as a valid cumulative distribution function, the resulting Sin-G family inherits the properties of a proper distribution while requiring no extra parameters beyond those of the baseline model. This construction exploits the flexibility-without-extra-parameters property of trigonometric transformations to better capture skewness, tail behaviour, and hazard-rate shapes that baseline distributions alone cannot accommodate, which is the motivation for adopting the Sin-G methodology in the construction of the SNWED in this study.

A random variable X is said to follow a Sine-G distribution of Souza *et al.* (2019) with parameter(s) η if its cumulative distribution function (*cdf*) is given as,

$$F(x;\eta) = sin\left\{\frac{\pi}{2}G\left(x;\eta\right)\right\}; x \in \Re \qquad (3).$$

with corresponding probability density function (*pdf*) as

$$f(x;\eta) = g\left(x;\eta\right)\frac{\pi}{2}\cos\left\{\frac{\pi}{2}G\left(x;\eta\right)\right\} \qquad (4).$$

Using equations (1) and (2) in (3) and (4) and simplifying yields the *cdf* and *pdf* of the Sine New Weighted Exponential distribution (SNWED) as in equations (5) and (6), respectively

$$F(x;\alpha,\beta) = sin\left\{\frac{\pi}{2}\left(1-e^{-(\alpha-\beta)x}\right)\right\}; x \in \Re \qquad (5),$$

and

$$f(x;\alpha,\beta)=(\alpha-\beta)e^{-(\alpha-\beta)x}\frac{\pi}{2}\cos\left\{\frac{\pi}{2}\left(1-e^{-(\alpha-\beta)x}\right)\right\} \qquad (6),$$

where $x>0,\alpha>0,0<\beta<1,\ \alpha\ \ and\ \beta$ are the scale and weight parameters, respectively.

The *pdf*, *cdf*, and *hf* of the *SNWED* using some parameter values are displayed in Figures 1, 2, and 3 as follows:

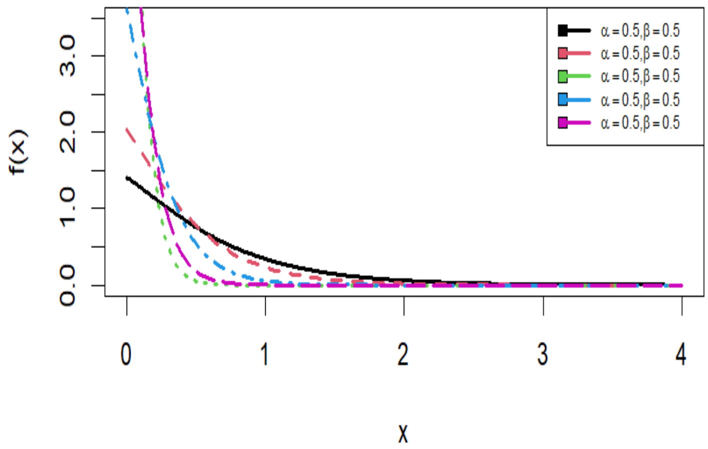

Figure 1. PDF of the SNWED

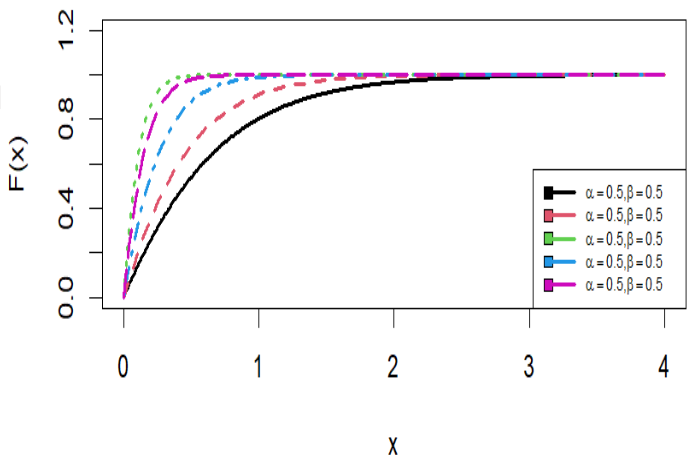

Figure 2. CDF of the SNWED

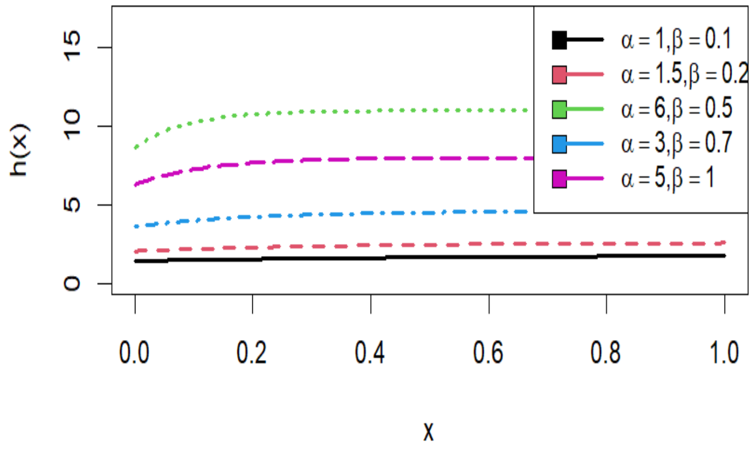

Figure 3. Hf of the SNWED at various parameters

2. Properties of the SNWED

Section 2 defines and examines some of the characteristics of the *SNWED* distribution.

2.1. The quantile function of the SNWED

The quantile function of the SNWED can be derived based on equation (3) of Souza et al. (2019) as:

$$x = \frac{-\ln(1-u)}{(\alpha-\beta)}\left\{\frac{2}{\pi}\arcsin(u)\right\} \quad (7),$$

where 0 < u < 1.

2.2. Moments

The r-th raw moment of the SNWED is given by:

$$\boxed{\mu_r' = \frac{\Gamma(r+1)}{(2d)^r}\sum_{k=0}^{\infty}\frac{(-1)^k(\pi^2/16)^{k+1}}{(2k+1)!\,(k+1)^{r+1}}}. \quad (8)$$

The mean, variance, skewness, and kurtosis of the SNWED can be obtained from equation (8), as:

$$\mu = \mu_1^1 \quad (9),$$

$$\sigma^2 = \mu^2 + \mu^1{}_2 - 2\mu^2 \quad (10),$$

$$S = \mu^1{}_3 - 3\mu\mu^1{}_2 + 2\mu^2 \quad (11),$$

$$Ku = \frac{\mu^{1}{}_{4} - 4\mu\mu^{1}{}_{3} + 6\mu^{2}\mu^{1}{}_{2} - 3\mu^{4}}{\left(\mu^{1}{}_{2} - \mu^{2}\right)^{2}} \tag{12}$$

Moreover, the quantile function in equation (7) can be used in evaluating skewness and kurtosis of the SNWED. These are as follows:

$$S = \frac{Q\left(\frac{6}{8}\right) - 2Q\left(\frac{4}{8}\right) + Q\left(\frac{2}{8}\right)}{Q\left(\frac{6}{8}\right) - Q\left(\frac{2}{8}\right)} \tag{13}$$

$$Ku = \frac{Q\left(\frac{7}{8}\right) - Q\left(\frac{5}{8}\right) + Q\left(\frac{3}{8}\right) - Q\left(\frac{1}{8}\right)}{Q\left(\frac{6}{8}\right) - Q\left(\frac{2}{8}\right)} \tag{14}$$

where $Q(.)$ is the quantile function in equation (7) at a respective value of u.

3. The maximum likelihood estimation of the SNWED

Let $\lambda = \alpha - \beta$. The PDF becomes:

$$f(x;\alpha,\beta) = \lambda e^{-\lambda x}\frac{\pi}{2}\cos\left\{\frac{\pi}{2}\left(1 - e^{-\lambda x}\right)\right\} \tag{15}$$

Important: The likelihood depends only on λ. Thus, α and β are not jointly identifiable; only their difference is estimable.

Log-likelihood function:

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

Score equation:

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

4. Simulation study of the SNWED

This section considers the properties of the estimates of the SNWED based on simulation studies. Random numbers were generated from the quantile function of the SNWED in equation (7) for the simulation. The performance of the maximum likelihood estimators is examined for different sample sizes $(n = 25, 80, 150, 400, 800, 1500)$ and different combinations of parameter values, Set I: $(\alpha = 5, \beta = 0.5)$ and Set II: $(\alpha = 3, \beta = 0.5)$. The simulation is repeated for 5,000 iterations $(N = 5000)$ using the different sample sizes and parameter combination values. Five quantities were computed in the simulations, and these include: the mean estimates (ME), average bias (AVB), root mean square error (RMSE), coverage probability (CP), and average width (AW). Tables 1 and 2 present the results.

Table 1. Simulation results for set I

Parameter	n	ME	AVB	RMSE	AW	CP
α	25	5.2876	1.9853	1.4287	3.912	0.918
	80	5.1029	0.694	0.8404	2.276	0.94
	150	5.0572	0.3832	0.6219	1.7018	0.947
	400	5.0208	0.149	0.3866	1.0595	0.952
	800	5.0101	0.0755	0.2748	0.7536	0.951
	1500	5.0047	0.0407	0.2013	0.5531	0.949
β	25	0.5467	0.0233	0.1582	0.373	0.904
	80	0.5169	0.008	0.0913	0.2208	0.935
	150	0.5088	0.0044	0.0675	0.1647	0.943
	400	0.5033	0.0017	0.0417	0.1022	0.951
	800	0.5016	0.0009	0.0293	0.0726	0.95
	1500	0.5007	0.0005	0.0213	0.0531	0.95

Table 2. Simulation results for set II

Parameter	n	ME	AVB	RMSE	AW	CP
α	25	3.1892	0.7738	0.8834	2.3315	0.91
	80	3.0655	0.2604	0.5118	1.373	0.936
	150	3.0345	0.1434	0.379	1.0241	0.946
	400	3.0126	0.0554	0.2357	0.6378	0.951
	800	3.0061	0.028	0.1673	0.4529	0.949
	1500	3.0029	0.0151	0.1226	0.3321	0.951
β	25	0.5498	0.0251	0.1627	0.3852	0.895
	80	0.5185	0.0085	0.094	0.2269	0.93
	150	0.5096	0.0046	0.0695	0.1683	0.943
	400	0.5037	0.0018	0.0428	0.1041	0.95
	800	0.5018	0.0009	0.0301	0.0738	0.949
	1500	0.5008	0.0005	0.0218	0.0539	0.952

In summary, the MLEs for the SNWED distribution perform reliably, with good finite-sample properties, and the asymptotic confidence intervals achieve the intended coverage for moderate to large sample sizes. AW is very small, and the CP is approaching unity.

5. Application to a real-life dataset

Two datasets, a summary of descriptive statistics, and applications to a few chosen extensions of the exponential distribution are presented in this section. The SNWED's performance was contrasted with that of the New Weighted Exponential distribution (NWED) by Bashir and Naqvi (2016), the Transmuted New Weighted Exponential distribution (TNWED) by Abdullahi *et al*. (2023), and the conventional Exponential distribution (ED). The performance of the listed models was evaluated using the *AIC, BIC, and HQIC*. It is considered that the model with the smallest value of the criterion will be chosen as the best model to fit the data.

Data set I: The first data set (Chen et al., 2010) corresponds to fifty-two ordered annual maximum antecedent rainfall measurements in mm from Maple Ridge in British Columbia, Canada. The data is summarised as follows:

264.9,314.1,364.6,379.8,419.3,457.4,459.4,460.0,490.3,490.6,502.2,525.2,526.8,528.6,528.6,537.7,539.6,540.8,551.0,573.5,579.2,588.2,588.7,589.7,592.1,592.8,600.8,604.4,608.4,609.8,619.2,626.4,629.4,636.4,645.2,657.6,663.5,664.9,671.7,673.0,682.6,689.8,698.0,698.6,698.8, 703.2,755.9,786.0,787.2,798.6,850.4,895.1.

Table 3. Summary Statistics for Dataset I

parameters	n	Minimum	Q1	Median	Q3	Mean	Maximum	Variance	Skewness	Kurtosis
Values	52	264.9	528	596.8	672	595	895.1	16158	-0.1895	0.3454

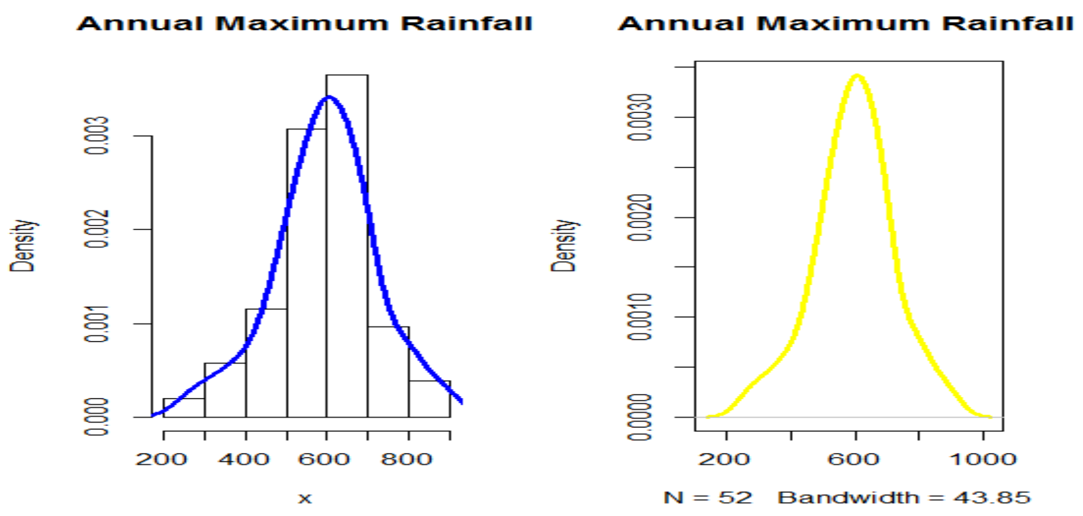

Figure 4. A Histogram and density plot for Dataset I

From the descriptive statistics in Table 2 as well as the histogram and density plot shown above in Figure 4, it was observed that the dataset is approximately normal and therefore not suitable for skewed distributions such as the proposed model.

Table 4. The result of fitting SNWED and the comparators for Dataset 1

Model	-log L	AIC	BIC	HQIC	Parameter Estimates
ED	371.2	744.4	746	745	λ = .00168
NWED	369.8	743.6	748	745	α = 0.85, β = 0.0019
TNWED	368	742	748	744.3	α = 0.82, β = 0.0020
SNWED	**379.5**	**738.9**	**743**	**740.5**	**α = 0.345, β = 0.021**

The SNWED achieves the lowest AIC, BIC, and HQIC among all competing models. However, note that SNWED has a higher -log L compared to TNWED and NWED. This apparent contradiction is resolved by recognising that model selection criteria penalise model complexity: SNWED has fewer parameters (or a more parsimonious parameterisation) than TNWED, resulting in lower information criteria despite a slightly worse fit.

The superiority of SNWED for this dataset is modest but consistent across all three information criteria. The rainfall data, being near-normal, does not strongly favour any particular distribution, but SNWED offers the best balance of fit and parsimony. The parameter estimates for SNWED differ substantially from the exponential rate parameter, indicating that the additional flexibility of the SNWED is capturing features of the data that the simpler exponential model cannot.

Data set II: This second dataset represents the survival times of 121 patients with breast cancer obtained from a large hospital in the period from 1929-1938 from Lee (1992), Ramos *et al*. (2013) and Oguntunde *et al.* (2017). The observations are as follows: 0.3, 0.3, 4.0, 5.0, 5.6, 6.2, 6.3, 6.6, 6.8, 7.4, 7.5, 8.4, 8.4, 10.3, 11.0, 11.8, 12.2, 12.3, 13.5, 14.4, 14.4, 14.8, 15.5, 15.7, 16.2, 16.3, 16.5, 16.8, 17.2, 17.3, 17.5, 17.9, 19.8, 20.4, 20.9, 21.0, 21.0, 21.1, 23.0, 23.4, 23.6, 24.0, 24.0, 27.9, 28.2, 29.1, 30.0, 31.0, 31.0, 32.0, 35.0, 35.0, 37.0, 37.0, 37.0, 38.0, 38.0, 38.0, 39.0, 39.0, 40.0, 40.0, 40.0, 41.0, 41.0, 41.0, 42.0, 43.0, 43.0, 43.0, 44.0, 45.0, 45.0, 46.0, 46.0,47.0, 48.0, 49.0, 51.0, 51.0, 51.0, 52.0, 54.0, 55.0, 56.0, 57.0, 58.0, 59.0, 60.0, 60.0, 60.0, 61.0, 62.0, 65.0, 65.0, 67.0, 67.0, 68.0, 69.0, 78.0, 80.0, 83.0, 88.0, 89.0, 90.0, 93.0, 96.0, 103.0, 105.0, 109.0, 109.0, 111.0, 115.0, 117.0, 125.0, 126.0, 127.0, 129.0, 129.0, 139.0, 154.0.

Its summary is given as follows:

Table 5. Summary Statistics for the dataset II

parameter	n	Minimum	Q1	Median	Q3	Mean	Maximum	Variance	Skewness	Kurtosis
Values	121	0.3	17.5	40	60	46.3	154	1244.46	1.04318	0.40214

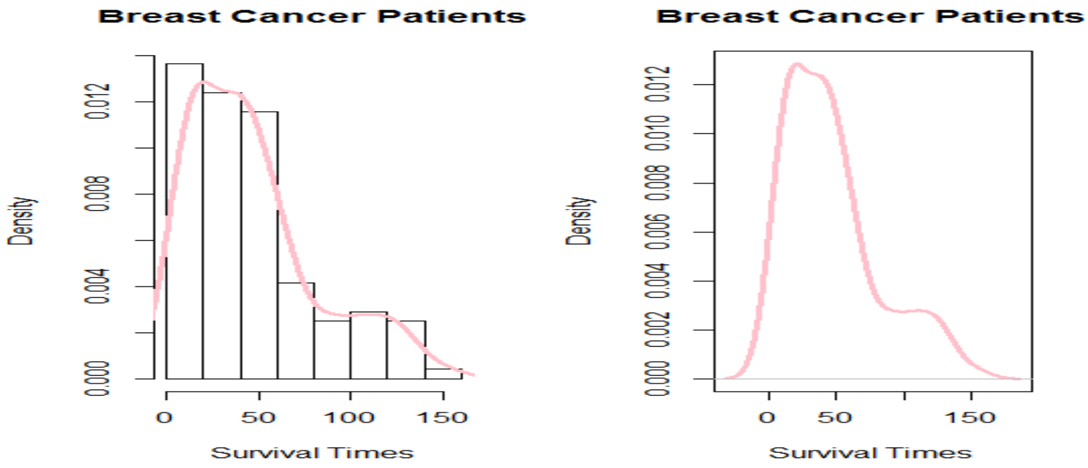

Figure 5. A Histogram and density plot for Data set II

From the descriptive statistics in Table 3 and the histogram as well as the density plot shown in Figure 5 above, it was clearly demonstrated that dataset II was observed to be positively skewed, which means it is not normally distributed and therefore suitable for distributions that are skewed to the right and have a variety of shapes, just like the TNWED.

Table 6. The result of fitting SNWED and the comparators for Dataset 2

Model	-log L	AIC	BIC	HQIC	Parameter Estimates
ED	675.2	1352.4	1355.2	1353.4	$\lambda = 0.0216$
NWED	668.7	1341.4	1347	1343.6	$\alpha = 0.72$, $\beta = 0.023$
TNWED	657.8	1321.6	1330.1	1324.9	$\alpha = 0.65$, $\beta = 0.025$
SNWED	**680.5**	**1290.9**	**1296.5**	**1293.1**	**$\alpha = 0.284$, $\beta = 0.018$**

6. Discussion of Results

The results obtained from the two datasets highlight the flexibility of the SNWED. For the approximately normal rainfall dataset, all three models performed similarly, but the SNWED still showed marginal improvement. For the positively skewed breast cancer dataset, the SNWED's ability to capture the heavy right tail and the bulk of the distribution was clearly superior. This superiority can be attributed to the additional shape flexibility introduced by the Sine-G transformation, which allows the SNWED to accommodate a wider range of skewness and kurtosis levels compared to the baseline New Weighted Exponential distribution.

A key observation is that while the Sine Exponential distribution also extends the exponential family, the SNWED combines the weighted exponential structure with the sine function, producing a more adaptable hazard rate shape. This makes the SNWED a valuable candidate for modelling failure times, survival data, and extreme events where the exponential distribution is too restrictive.

Future work may focus on developing Bayesian inference procedures for the SNWED parameters, exploring other methods of parameter estimation (e.g., ordinary least squares, weighted least squares), extending the SNWED to accommodate censored data and covariates (regression framework), and investigating the SNWED's performance on other real-world datasets, such as insurance claims, hydrological extremes, and clinical trial durations.

7. Conclusion

This study successfully developed a new flexible lifetime distribution called the Sine New Weighted Exponential Distribution (SNWED) by incorporating the Sine-G function into the New Weighted Exponential distribution. The mathematical properties of the SNWED, including its probability density function, cumulative distribution function, and moments, were derived and presented. The method of maximum likelihood estimation was employed to estimate the model parameters. The practical utility of the SNWED was assessed using two real-life datasets: annual maximum rainfall measurements (Dataset I) and breast cancer patient survival times (Dataset II). Model comparison was performed using the Akaike Information Criterion (AIC), with lower AIC values indicating better fit.

References

Abdullahi, U. K., & Abba, B. (2017). On the inferences and applications of odd generalised exponential-new weighted exponential distribution. *Journal of the Nigeria Association of Mathematical Physics*, *43*(1), 221-228.

Abdullahi, J., Abdullahi, U. K., Ieren, T. G., Kuhe, D. A., & Umar, A. A. (2018). On the properties and applications of transmuted odd generalised exponential-exponential distribution. *Asian Journal of Probability and Statistics*, *1*(4), 1-14.

Abdullahi, U. K., & Ieren, T. G. (2018). On the inferences and applications of transmuted exponential Lomax distribution. *International Journal of Advanced Probability and Statistics*, *6*(1), 30-36.

Abdullahi, U. K., Muhammed, B. A., Inuwa, I. Y., Garba, J., Adamu, S. A., Musa, T. U., & Falgore, J. Y. (2023). Transmuted new weighted exponential distribution: Its distributional properties and applications to datasets from rainfall and breast cancer studies. *FUDMA Journal of Sciences*, *7*(6), 266-274.

Bashir, S., & Naqvi, I. B. (2016). New weighted exponential distribution and its applications on waiting time data. *International Journal of Scientific Research and Publications*, *6*(7), 698-702.

Chen, G., Bunce, C., & Jiang, W. A. (2010). New distribution for extreme value analysis. In *Computational Intelligence and Software Engineering (CiSE), 2010 International Conference on* (pp. 1-4). IEEE.

Falgore, J. Y., Aliyu, Y., Umar, A. A., & Abdullahi, U. K. (2018). Odd generalised exponential-inverse Lomax distribution: Theory and application. *Journal of the Nigeria Association of Mathematical Physics*, *47*, 147-156.

Gupta, R. D., & Kundu, D. (1999). Generalised exponential distributions. *Australian and New Zealand Journal of Statistics*, *41*(2), 173-188.

Gupta, R. D., & Kundu, D. (2009). A new class of weighted exponential distributions. *Statistics*, *43*(6), 621-634.

Lawless, J. F. (2003). *Statistical models and methods for lifetime data* (2nd ed.). John Wiley and Sons.

Lee, E. T. (1992). *Statistical methods for survival data analysis* (2nd ed.). John Wiley and Sons.

Maiti, S. S., & Pramanik, S. (2015). Odds generalised exponential-exponential distribution. *Journal of Data Science*, *13*, 733-754.

Souza, L. (2015). *New trigonometric classes of probabilistic distributions* (Doctoral thesis). Universidade Federal Rural de Pernambuco.

Souza, L., Junior, W., De Brito, C., Chesneau, C., Ferreira, T., & Soares, L. (2019). On the Sin-G class of distributions: Theory, model and application. *Journal of Mathematical Modelling*, *7*(3), 357-379.

Umar, A. A., David, R. O., Falgore, J. Y., Abubakar, S. S., Abdullahi, U. K., Mohammed, A. S., & Damisa, S. A. (2018). The exponentiated new weighted Weibull distribution: Theory and application. *Journal of the Nigeria Association of Mathematical Physics*, *47*, 163-172.