Mathematical Sciences Letters
An International Journal

http://dx.doi.org/10.18576/msl/120201

Extended cosine Generalized Power Weibull Distribution: Properties, and Regression

Shei Baba Sayibu[*,1], *Albert Luguterah*[2]

[1]Department of Statistics, Faculty of Physical Sciences, University for Development Studies, West Africa, Ghana, Tamale.
[2]Department of Statistics and Actuarial Science, School of Mathematical Sciences, C. K. Tedam University of Technology and Applied Sciences, West Africa, Ghana, Navrongo.

Received: 25 Jul. 2022, Revised: 28 Nov. 2022, Accepted: 25 Feb. 2023
Published online: 1 May 2023

Abstract: This research introduces a new five-parameter model known as the extended cosine generalized power Weibull distribution. The novel model has many sub-models that are very useful in modeling real-life data such as extended cosine Weibull, extended cosine exponential, extended cosine Nadarajah Haghighi and the extended cosine Rayleigh. The model proposed demonstrates a high level of suitability in modeling both reliability and survival data. It is very flexible in accommodating various failure rates such as reversed J-shape, increasing, decreasing, and bathtub. Some of the statistical properties studied are the quantile function, moments, moment generating function, incomplete moment, and order statistics. The maximum likelihood and the ordinary least square estimation methods were employed to estimate the parameters. The log extended cosine generalize power Weibull location-scale regression model was also developed and its application demonstrated using the birth weight of babies. The flexibility of the proposed distribution was demonstrated with three data sets. The proposed distribution was compared with its sub-modes and other existing models. The results showed the proposed distribution provided a better fit to the data sets as compared to the other competing distributions.

Keywords: Extended cosine; failure rate; reliability; regression; location; cumulative hazard.

1 Introductions

In the recent past, the modeling of survival/reliability data was limited to the use of some classical probability distributions such as Weibull, exponential, Rayleigh, gamma, and beta distributions. Data churned out in recent times are so complex, rendering the classical models deficient in modeling such data sets. To improve upon the flexibility of these classical models, researchers in probability distribution theory modify the existing distributions. The modifications of the classical distributions are aimed at improving their goodness of fit. The Weibull distribution developed by [1] has seen a lot of usage for modeling lifetime data. However, this model does not give a better fit to data sets that exhibit non-monotonic failure rates such as bathtub, modified bathtub, or uni-modal shapes [2]. Given the shortfall, it has undergone a lot of modifications in literature. Some of these include the exponentiated generalized power generalized Weibull distribution by [3], exponentiated Weibull distribution by [4], modified Weibull distribution by [5], a new modified Weibull by [6], and generalized power Weibull (GPW) by [7]. The GPW distribution takes the form of the Weibull model when $\gamma = 1$. [7], enumerated the various failure rates of the GPW such as constant, decreasing, increasing, and bathtub shapes. The random variable X, follows GPW distribution if its cumulative distribution function (CDF) is

$$G(x) = 1 - e^{1-(1+\lambda(x)^{\theta})^{\gamma}}, \tag{1}$$

where

$$x > 0, \gamma > 0, \theta > 0, \lambda > 0$$

$$\lambda \text{ is a scale parameter}, \gamma \text{ and } \theta \text{ are shape parameters}.$$

The corresponding probability density function (PDF) is given as

$$g(x) = \gamma\lambda\theta(x)^{\theta-1}(1+\lambda(x)^{\theta})^{\gamma-1}e^{1-(1+\lambda(x)^{\theta})^{\gamma}}. \tag{2}$$

* Corresponding author e-mail: b.shei@yahoo.com

[8] introduced the CDF of the extended cosine-generated family as follows

$$F(x) = \left[1 - \cos\left(\frac{\pi}{2} G(x)^{\alpha}\right)\right]^{\beta}. \quad (3)$$

The corresponding PDF of the extended cosine-generated distribution is

$$f(x) = \frac{\pi}{2} \alpha\beta g(x) G(x)^{\alpha-1} \sin\left[\frac{\pi}{2} G(x)^{\alpha}\right] \times \left[1 - \cos\left(\frac{\pi}{2} G(x)^{\alpha}\right)\right]^{\beta-1}, \quad (4)$$

where $x \in R$, $\alpha > 0$, $\beta > 0$, $g(x) = \frac{dG(x)}{dx}$,$g(x)$ and $G(x)$ are the PDF and CDF of the baseline distribution, α and β are additional shape parameters

Cosine distributions are used for a wide range of applications including fields as diverse as psychological investigations, semiconductor manufacturing, atomic sputtering, and antenna design. Examples of some trigonometric generalized distributions include beta trigonometric distribution [9], trigonometric classes of probabilistic distributions [10], cosine sine distribution [11], a sine-G family of distributions [12], cosine geometric distribution [13], sine Topp-Leone-G family of distributions [14], and sine power Lomax model [15]. Motivated by the work of [8] and to improve upon the flexibility of the generalized power Weibull, this study wishes to present a modified GPW distribution with the ability of:

– Modeling non-monotonic failure rate datasets in reliability/survival analysis.
– modeling data sets that are directional and proportional in nature which therefore require trigonometric function-base distributions using the extended cosine generalized generated family of distributions.
– providing a high level of flexibility, accuracy, and, goodness of fit due to the injection of additional shape parameters.

In this research paper, the works of [8] and [7] are employed in the study of the new model known as the extended cosine generalized power Weibull distribution (ECGPWD). Organization of the other sections of the paper is as follows: section 2 covers the development of the ECGPWD and its sub-models, some of the statistical properties of the ECGPWD are considered under Section 3, parameter estimation is covered under Section 4, Monte Carlo simulation is covered under section 5, empirical application of the ECGPWD to real data is considered in section 6, the location-scale regression is considered in section 7, and The concluding remarks are under section 8.

2 Extended cosine generalized power Weibull distribution

This section introduces the PDF and the CDF of the ECGPWD using the GPW as the baseline function.

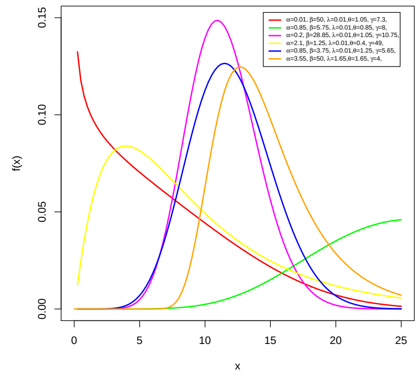

Figure 1: Some PDF shapes of ECGPWD

Substituting equation (1) into equation (3), the CDF of the ECGPWD is obtained as

$$F(x) = \left[1 - cos\left(\frac{\pi}{2}\left(1 - e^{1-(1+\lambda x^{\theta})^{\gamma}}\right)^{\alpha}\right)\right]^{\beta} \quad (5)$$

where

$$x > 0, \theta > 0,\ \gamma > 0, \alpha > 0,\ \beta > 0,\ \lambda > 0$$

. Differentiating equation (5) gives the PDF of the ECGPWD as,

$$f(x) = \frac{\alpha\beta\lambda\theta\gamma\pi x^{\theta}(1+\lambda x^{\theta})^{\gamma}}{2x(1+\lambda^{\theta})} (\Gamma_i)^{\alpha-1} \times \sin\left[\frac{\pi}{2}(\Gamma_i)^{\alpha}\right] \times \left[1 - cos\left(\frac{\pi}{2}(\Gamma_i)^{\alpha}\right)\right]^{\beta-1}. \quad (6)$$

where $\Gamma_i = 1 - e^{1-(1+\lambda x^{\theta})^{\gamma}}$ The desired PDF shapes of the ECGPWD are shown in Figure 1. The desired shapes are symmetric, increasing, decreasing, and right skewed. The survival function is defined as

$$S(x) = 1 - \left[1 - cos\left(\frac{\pi}{2}\left(1 - e^{1-(1+\lambda x^{\theta})^{\gamma}}\right)^{\alpha}\right)\right]^{\beta}. \quad (7)$$

This hazard function is defined as

$$H(x) = \frac{Nx^{\theta}VD(\Gamma_i)^{ac}\sin\left(\frac{\pi}{2}(\Gamma_i)^{\alpha}\right)\left(\tau_i\left(\frac{\pi}{2}(\Gamma_i)^{\alpha}\right)\right)^{\beta-1}}{2x(1+\lambda x^{\theta})(1-V)\left(1-\left(\tau_i\left(\frac{\pi}{2}(1-e^{1-V})^{\alpha}\right)\right)^{\beta}\right)}$$

where $N = \alpha\beta\lambda\theta\gamma\pi$ and $V = (1+\lambda x^{\theta})^{\gamma}$, $D = e^{1-V}$, $\tau_i = 1 - cos$.

Some of the desired shapes of the hazard function include reversed J-shape, increasing, decreasing, and bathtub as shown in Figure 2. The reversed hazard function is expressed as the density function divided by the cumulative density function and it is presented as

$$h(x) = \frac{Nx^{\theta}VD(\Gamma_i)^{ac}\sin\left(\frac{\pi}{2}(\Gamma_i)^{\alpha}\right)\left(\tau_i\left(\frac{\pi}{2}(\Gamma_i)^{\alpha}\right)\right)^{\beta-1}}{2x(1+\lambda x^{\theta})(1-V)\left(\tau_i\left(\frac{\pi}{2}(\Gamma_i)^{\alpha}\right)\right)^{\beta}}$$

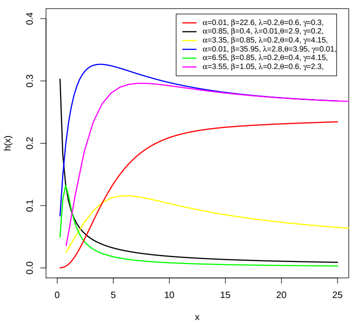

Figure 2: shapes of the ECGPWD Hazard Function

Mills ratio is defined as the ratio of the survival function to the density function. A common application of the Mills ratio arises in regression analysis which caters for biases in sample selection. The Mill's ratio is given by,

$$M(x) = \frac{2x(1+\lambda x^{\theta})(1-V)\left(1-\left(1-\cos\left(\frac{\pi}{2}\left(\Gamma_i\right)^{\alpha}\right)\right)^{\beta}\right)}{Nx^{\theta}VD\left(\Gamma_i\right)^{ac}\sin\left(\frac{\pi}{2}\left(\Gamma_i\right)^{\alpha}\right)\left(\tau_i\left(\frac{\pi}{2}\left(\Gamma_i\right)^{\alpha}\right)\right)^{\beta-1}}.$$

The cumulative hazard function is the measure of the overall accumulated risk that has been accumulated up to a certain point in time, t. The cumulative rate can either increase or remain the same. It is presented as,

$$m(x) = -\ln 1 - \left[1-cos\left(\frac{\pi}{2}\left(1-e^{1-(1+\lambda x^{\theta})^{\gamma}}\right)^{\alpha}\right)\right]^{\beta}. \quad (8)$$

2.1 *Sub-models of the ECGPWD*

The ECGPWD converges to different important distributions given certain values of its parameters. If $X\ ECGPWD(\alpha,\beta,\lambda,\theta\lambda)$, the following sub-models are obtained for some given parameter values.

- For $\gamma = 1$, the PDF reduces to the extended cosine Weibull distribution (ECWD).
- For $\gamma = 1$ and $\theta = 1$, the PDF becomes extended cosine exponential distribution (ECED) which is a new model.
- For $\theta = 1$, the PDF turns into extended cosine Nadarajah-Haghighi distribution (ECNHD) which is also a new model.
- For $\gamma = 1$ and $\theta = 2$, the PDF turns into extended cosine Rayleigh distribution (ECRD) which is also a new model.

3 Statistical Properties of the ECGPW Distribution

The statistical properties that are considered in this section are the quantile function, moments, moment generating function, characteristic function, incomplete moment, and order statistics.

3.1 *Quantile function*

The distribution of the random variable X, is well explained by the quantile function. The quantile function is significant with regard to the generation of random numbers as well as the computation of the variation and dispersion.
Lemma 1. The quantile function of the ECGPWD is given as

$$x_u = \rho\left[\left[1-log\left[1-\left[\frac{\pi}{2}\arccos\left(1-u^{\frac{1}{\beta}}\right)\right]^{\frac{1}{\alpha}}\right]\right]^{\frac{1}{\gamma}}-1\right]^{\frac{1}{\theta}},$$

where $\rho = \frac{1}{\lambda^{\theta}}$, $u \in [0,1]$.
Proof. For a strictly monotonic function, the quantile function Q, returns a threshold value x below which random draws from the given CDF could fall u percent of the time. $Q(u) \leq x$ if and only if $u \leq F(x)$. Therefore if $F(x)$ is continuous and strictly monotonically increasing, the inequalities are replaced by equality signs. Hence, the quantile function can be written in terms of u as; $Q(u) = F^{-1}(u)$.
Using the CDF given in equation (5),

$$\left[1-cos\left(\frac{\pi}{2}\left(1-e^{1-(1+\lambda x^{\theta})^{\gamma}}\right)^{\alpha}\right)\right]^{\beta} = u$$

Simplifying further gives,

$$\left[1-cos\left(\frac{\pi}{2}\left(1-e^{1-(1+\lambda x^{\theta})^{\gamma}}\right)^{\alpha}\right)\right] = u^{\frac{1}{\beta}}$$

After some algebraic manipulation gives, $e^{1-(1+\lambda x^{\theta})^{\gamma}} = 1-\left[\frac{2}{\pi}\cos^{-1}\left(1-u^{\frac{1}{\beta}}\right)\right]^{\frac{1}{\alpha}}$
Finally,

$$x_u = \frac{1}{\lambda^{\theta}}\left[\left[1-log\left[1-\left[\frac{\pi}{2}\arccos\left(1-u^{\frac{1}{\beta}}\right)\right]^{\frac{1}{\alpha}}\right]\right]^{\frac{1}{\gamma}}-1\right]^{\frac{1}{\theta}}$$

3.2 *Expansion of the PDF*

The PDF in a mixture form is as follows. Using the series expansions, the PDF can be presented in a linear sum of powers.
Lemma 2. The density function of the ECGPWD in a mixture form is,

$$f(x) = \alpha\theta\lambda\gamma x^{\theta-1}(1+\lambda x^{\theta})^{\gamma-1}w_{ijk}e^{(j+1)(1-(1+\lambda x^{\theta})^{\gamma})}, \quad (9)$$

where
$w_{ijk} = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty} b_k(i)(-1)^{i+j}\binom{\beta}{i}\binom{2\alpha k-1}{j}(\frac{\pi}{2})^{2k}$
Proof. Using binomial series expansion

$$(1-x)^n = \sum_{i=0}^{\infty}(-1)^{-1}\binom{n}{i}x^i$$

then

$$\left[1-cos\left(\frac{\pi}{2}\left(1-e^{1-(1+\lambda x^{\theta})^{\gamma}}\right)^{\alpha}\right)\right]^{\beta-1} = \sum_{i=0}^{\infty}(-1)^i \times \binom{\beta-1}{i}\left[\cos\left(\frac{\pi}{2}\left(1-e^{1-(1+\lambda x^{\theta})^{\gamma}}\right)\right)\right]^i$$

Also

$$\left[cos\left(\frac{\pi}{2}\left(1-e^{1-(1+\lambda x^{\theta})^{\gamma}}\right)^{\alpha}\right)\right]^i = \sum_{j=0}^{\infty} b_j(i) \times \left(\frac{\pi}{2}\left(1-e^{1-(1+\lambda x^{\theta})^{\gamma}}\right)^{\alpha}\right)^{2j}$$

where $b_0(i) = 1, b_1(i) = \frac{i(3i-2)}{24}$,...., [8] The linear expansion of the PDF of the extended cosine family according to [8] is,

$$f(x) = \sum_{i=1}^{\infty} B_{i,k}C_{2k\alpha}(x)$$

where $C_{2k\alpha}(x) = 2k\alpha g(x)G(x)^{2k\alpha-1}$ is the exponentiated density with $2k\alpha$ as the power parameter and $B_{i,k} = \sum_{i=0}^{\infty} b_k(i)(-1)^i\binom{\beta}{i}(\frac{\pi}{2})^{2k}$ Thus

$$f(x) = 2k\alpha\gamma\theta\lambda\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty} b_k(i)(-1)^{i+j}\binom{\beta}{i}\binom{2\alpha k-1}{j} (\frac{\pi}{2})^{2k}x^{\theta-1}(1+\lambda x^{\theta})^{\gamma-1}e^{(j+1)\left(1-(1+\lambda x^{\theta})^{\gamma}\right)}$$

It is finally presented as,

$$f(x) = \alpha\theta\lambda\gamma x^{\theta-1}(1+\lambda x^{\theta})^{\gamma-1}w_{ijk}e^{(j+1)(1-(1+\lambda x^{\theta})^{\gamma})}.$$

Equation (9) shows that the new PDF is a linear combination of exponentiated-G distribution. Utilizing the properties of the exponentiated-G, various statistical properties of the new distribution can be obtained.

3.3 Moments

The moments in statistical analysis are very useful in the calculation and determination of central tendencies and measures of dispersion.
Proposition 1. If X follows the ECGPWD, then the r^{th} moment of the random variable X is

$$\mu_r^{'} = 2k\alpha\omega_{ijkl}e^{j+1}\left(\frac{1}{\lambda}\right)^{\frac{r}{\theta}}\left(\frac{1}{j+1}\right)^{\frac{r+\theta(\gamma-l)}{\theta\gamma}} \times\Gamma\left(\frac{r+\theta(\gamma-l)}{\theta\gamma},J+1\right) \quad (10)$$

where $\omega_{ijkl} = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty} b_k(-1)^{i+j+l}\binom{\beta}{i} \times \binom{2\alpha k-1}{j}\binom{\frac{r}{\theta}}{l}\left(\frac{\pi}{2}\right)^{2k}, r = 1,2,\dots$ Proof. The r^{th} non-central moment is defined as, $\mu_i^{'} = E[X^r] = \int_0^{\infty} x^r f(x)dx$
Substituting the expanded form of the PDF in equation (13) into equation moment definition gives

$$\mu_i^{'} = 2k\alpha\theta\lambda\gamma\omega_{ijk}\int_0^{\infty} x^{r+\theta-1}(1+\lambda x^{\theta})^{\gamma-1} \times e^{(j+1)(1-(1+\lambda x^{\theta})^{\gamma})}dx$$

Let $u = (j+1)(1+\lambda x^{\theta})^{\gamma}$, then $x = \left(\frac{1}{\lambda}\left(\frac{u}{j+1}\right)^{\frac{1}{\gamma}}-1\right)^{\frac{1}{\theta}}$
and $dx = \frac{du}{\theta\lambda\gamma x^{\theta-1}}(1+\lambda x^{\theta})^{\gamma-1}$
Substituting dx and x back into the integral part, with the new limits as $(j+1,\infty)$, it gives
$\mu_i^{'} = 2k\alpha\omega_{ijk}e^{j+1}\int_{j+1}^{\infty}\left(\frac{1}{\lambda}\left(\frac{u}{j+1}\right)^{\frac{1}{\lambda}}-1\right)^{\frac{1}{\theta}} e^{-u}\frac{du}{j+1}$.
Using the binomial theorem with the power series identity and further simplifying yields,

$$\mu_r^{'} = 2k\alpha\omega_{ijkl}e^{j+1}\left(\frac{1}{\lambda}\right)^{\frac{r}{\theta}}\left(\frac{1}{j+1}\right)^{\frac{r+\theta(\gamma-l)}{\theta\gamma}} \times\Gamma\left(\frac{r+\theta(\gamma-l)}{\theta\gamma},J+1\right)$$

3.4 Moment Generating Function

Proposition 2. If X follows the ECGPWD, for any integer value, the moment generating function, $M_X(t)$ of the random variable X is

$$M_X(t) = 2k\alpha\omega_{ijkl}\sum_{r=0}^{\infty}\frac{t^r}{r!}e^{j+1}\left(\frac{1}{\lambda}\right)^{\frac{r}{\theta}}\left(\frac{1}{j+1}\right)^{\frac{r+\theta(\gamma-l)}{\theta\gamma}} \times\Gamma\left(\frac{r+\theta(\gamma-l)}{\theta\gamma},J+1\right) \quad (11)$$

Proof. The moment-generating function is defined as follows;
$M_X(t) = E\left[e^{tX}\right] = \int_0^{\infty} e^{tx}f(x)dx = E\left[\sum_{r=0}^{\infty}\frac{(tX)^r}{r!}\right]$
This implies $M_X(t) = \sum_{r=0}^{\infty}\frac{t^r}{r!}E[X^r]$
using $\mu_r^{'} = E[X^r]$ as given in equation (10) then, $M_X(t) = E\left[e^{tX}\right] = \int_0^{\infty} e^{tx}f(x)dx = E\left[\sum_{r=0}^{\infty}\frac{(tX)^r}{r!}\right]$
which is finally simplified as

$$M_X(t) = 2k\alpha\omega_{ijkl}\sum_{r=0}^{\infty}\frac{t^r}{r!}e^{j+1}\left(\frac{1}{\lambda}\right)^{\frac{r}{\theta}}\left(\frac{1}{j+1}\right)^{\frac{r+\theta(\gamma-l)}{\theta\gamma}} \times\Gamma\left(\frac{r+\theta(\gamma-l)}{\theta\gamma},J+1\right)$$

3.5 Characteristic Function

The characteristic function is just like the moment-generating function and is mostly used to demonstrate the uniqueness of two distributions.
Proposition 3. If X follows the ECGPWD, for any integer value, the characteristic function, $\phi_X(t)$ of the random variable X is

$$\phi_X(t) = 2k\alpha\omega_{ijkl}\sum_{m=0}^{\infty}\frac{(qt)^m}{m!}e^{j+1}\left(\frac{1}{\lambda}\right)^{\frac{r}{\theta}}\left(\frac{1}{j+1}\right)^{\frac{r+\theta(\gamma-l)}{\theta\gamma}} \times\Gamma\left(\frac{r+\theta(\gamma-l)}{\theta\gamma},J+1\right) \quad (12)$$

Proof. The characteristic function $\phi_X(t)$ is defined as, $\phi_X(t) = E\left[e^{-qtX}\right] = \sum_{r=0}^{\infty}\frac{(qt)^m}{m!}E\left[X^r\right]$ where $q=\sqrt{-1}$ using $\mu_r^{'} = E\left[X^r\right]$ as given in equation (10) then
$\phi_X(t) = 2k\alpha\omega_{ijkl}\sum_{m=0}^{\infty}\frac{(qt)^m}{m!}e^{j+1}\left(\frac{1}{\lambda}\right)^{\frac{r}{\theta}}\left(\frac{1}{j+1}\right)^{\frac{r+\theta(\gamma-l)}{\theta\gamma}}$
$\times\Gamma\left(\frac{r+\theta(\gamma-l)}{\theta\gamma},J+1\right)$

3.6 Incomplete Moment

The incomplete moment is useful in determining the shapes of the moments of a distribution.
Proposition 4. The r^{th} incomplete moment of a random variable X defined by $M_r(y)$ is given by;

$$M_r(y) = 2k\alpha\omega_{ijklm}e^{(j+1)}\times \Gamma\left(\frac{r+\theta(\gamma-m)}{\theta\gamma},(j+1)(1+\lambda x^\theta)^\gamma\right), \quad (13)$$

where
$w_{ijkl} = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\sum_{m=0}^{\infty}b_k(-1)^{i+j+l+m}\binom{\beta}{i}$
$\times\binom{2\alpha k-1}{j}\binom{\frac{r}{\theta}}{l}\left(\frac{\pi}{2}\right)^{2k}\left(\frac{1}{\lambda}\right)^{\frac{r}{\theta}}\left(\frac{1}{j+1}\right)^{\frac{r+\theta(\gamma-m)}{\theta\gamma}}, r=1,2,\ldots$
Proof. The r^{th} incomplete moment of X is defined as follows: $M_r(y) = E\left[X^r|X>y\right] = \int_y^\infty x^r f(x)dx$
Substituting the mixture form of the PDF in (9) into the $M_r(y)$ expression yields,
$M_r(y) = 2k\alpha\theta\lambda\gamma\omega_{ijk}\int_y^\infty x^{r+\theta-1}(1+\lambda x^\theta)^{\gamma-1}$
$\times e^{(j+1)(1-(1+\lambda x^\theta)^\gamma)}dx$
By setting $u = (j+1)((1+\lambda x^\theta)^\gamma)$, then $x=\left(\frac{1}{\lambda}\left(\frac{u}{j+1}\right)^{\frac{1}{\gamma}}-1\right)^{\frac{1}{\theta}}$ and $dx = \frac{du}{\theta\lambda\gamma x^{\theta-1}}(1+\lambda x^\theta)^{\gamma-1}$
Substituting dx and x back into the integral with the new limits as $((j+1)(1+\lambda x^\theta)^\gamma,\infty)$ gives
$M_X(y) = 2k\alpha\omega_{ijk}e^{j+1}$
$\times\int_{(j+1)(1+\lambda x^\theta)^\gamma}^\infty\left(\frac{1}{\lambda}\left(\frac{u}{j+1}\right)^{\frac{1}{\gamma}}-1\right)^{\frac{r}{\theta}}e^{-u}du.$
Using the binomial theorem with the power series identity and further simplifying yields,
$M_X(y) = 2k\alpha\omega_{ijkl}e^{(j+1)}\Gamma\left(\frac{r+\theta(\gamma-m)}{\theta\gamma},(j+1)(1+\lambda x^\theta)^\gamma\right)$

3.7 Order Statistics

Proposition 5. The r^{th} order statistic of the random sample $X_1,\ldots,X_n$ is is presented as

$$f_{X(r)(x)} = KC_m\frac{\sum_{i,j,k,l=0}^{\infty}(-1)^{i+j+l}2^{-2(j+1)}\pi^{2(j+1)}}{2\alpha(j+l+1)}x^\theta \times\left(1+\lambda x^\theta\right)^{\gamma-1}e^{1-(1+\lambda x^\theta)^\gamma}\left(1-e^{1-(1+\lambda x^\theta)^\gamma}\right)^{2\alpha(j+l+1)} \quad (14)$$

where
$C_m = 2K\alpha^2\beta(j+l+1)\binom{n-r}{i}\binom{\beta(r+i-1)}{l}\binom{2\alpha(j+l+1)}{m}$
Proof: The PDF of X_r is defined as
$f_{X(r)(x)} = \frac{n!}{(n-r)!(r-1)!}\left[F_X(x)\right]^{r-1}\left[1-F_X(x)\right]^{n-r}]f_X(x)$
Using the Binomial expansion, this is simplified and presented as
$f_{X(r)(x)} = \frac{n!}{(n-r)!(r-1)!}\sum_{i=1}^{n-r}(-1)^i\binom{n-r}{i}\left[F_X(x)\right]^{r+i-1}f_X(x)$
the order statistics as defined by [8],
$f_{X(r)(x)} = 2K\alpha^2\beta(j+l+1)\binom{n-r}{i}\binom{\beta(r+i-1)}{l}$
$\times\frac{\sum_{i,j,k,l=0}^{\infty}(-1)^{i+j+l}2^{-2(j+1)}\pi^{2(j+1)}}{2\alpha(j+l+1)}g(x)G(x)^{2\alpha(j+l+1)}$
where $k = \frac{n!}{(n-r)!(r-1)!}$ substituting the expressions for $g(x)$ and $G(x)$ then
$f_{X(r)(x)} = 2K\alpha^2\beta(j+l+1)\binom{n-r}{i}\binom{\beta(r+i-1)}{l}$
$\times\frac{\sum_{i,j,k,l=0}^{\infty}(-1)^{i+j+l}2^{-2(j+1)}\pi^{2(j+1)}}{2\alpha(j+l+1)}x^{\theta-1}\left(1+\lambda x^\theta\right)^{\gamma-1}$
$\times e^{1-(1+\lambda x^\theta)^\gamma}\left(1-e^{1-(1+\lambda x^\theta)^\gamma}\right)^{2\alpha(j+l+1)}$
It is further simplified as
$f_{X(r)(x)} = KC_m\frac{\sum_{i,j,k,l=0}^{\infty}(-1)^{i+j+l}2^{-2(j+1)}\pi^{2(j+1)}}{2\alpha(j+l+1)}x^{\theta-1}$
$\times\left(1+\lambda x^\theta\right)^{\gamma-1}e^{1-(1+\lambda x^\theta)^\gamma}\left(1-e^{1-(1+\lambda x^\theta)^\gamma}\right)^{2\alpha(j+l+1)}$

4 Estimation and Inference of ECGPWD

This section presents two different techniques employed to estimate the unknown parameters of the ECGPWD. The methods are maximum likelihood (MLE) and the ordinary least square (OLS) estimation methods.

4.1 Maximum Likelihood method of Estimation

Let $x_1,x_2,\ldots,x_n$ be a random sample of size n, that follows the ECGPWD model. The log-likelihood function with respect to equation (6) is,

$$\log\ell = n\log\left(\frac{\pi}{2}\right)n\log\alpha+n\log\beta+n\log\lambda+n\log\theta +n(\theta-1)\sum_{i=1}^n\log x_i+n(\gamma-1)\sum_{i=1}^n\log\left(1+\lambda x_i^\theta\right)+ n-\sum_{i=1}^n V+n(\gamma-1)\sum_{i=1}^n\log(\Gamma_i)+n\sum_{i=1}^n\log\times \sin\left[\frac{\pi}{2}(\Gamma_i)^\alpha\right]+n(\beta-1)\sum_{i=1}^n log\left[\tau_i\left[\frac{\pi}{2}(\Gamma_i)^\alpha\right]\right].$$

The maximum likelihood estimators (MLEs) are the solutions of the equations:
$\frac{\partial}{\partial\alpha}\ell(x:,\alpha,\beta,\lambda,\theta,\gamma)=0$, $\frac{\partial}{\partial\beta}\ell(x:,\alpha,\beta,\lambda,\theta,\gamma)=0$, $\frac{\partial}{\partial\lambda}\ell(x:,\alpha,\beta,\lambda,\theta,\gamma)=0$, $\frac{\partial}{\partial\theta}\ell(x:,\alpha,\beta,\lambda,\theta,\gamma)=0$, $\frac{\partial}{\partial\gamma}\ell(x:,\alpha,\beta,\lambda,\theta,\gamma)=0$,
Thus:

$$\frac{\partial\ell}{\partial\alpha}=\frac{n}{\alpha}+n\sum_{i=1}^{n}\frac{\pi}{2}\Gamma_i^{\alpha}\log\Gamma_i+\sum_{i=1}^{n}\log\Gamma_i+n(\beta-1)\times\sum_{i=1}^{n}\frac{\pi\Gamma_i^{\alpha}\log\Gamma_i\sin\left[\frac{\pi}{2}\Gamma_i^{\alpha}\right]}{2\left[\tau_i\left[\frac{\pi}{2}\Gamma_i^{\alpha}\right]\right]}$$

$$\frac{\partial\ell}{\partial\beta}=\frac{n}{\beta}+n\sum_{i=1}^{n}log\left[\tau_i\left[\frac{\pi}{2}\Gamma_i^{\alpha}\right]\right]$$

$$\frac{\partial\ell}{\partial\lambda}=\frac{n}{\lambda}-\gamma x_i^{\theta}V_i^{\gamma-1}+(\gamma-1)\sum_{i=1}^{n}\frac{\gamma x_i^{\theta}}{V_{\;i}}+\frac{n}{2}\sum_{i=1}^{n}e^{1-V_i}\times\Gamma_i\alpha\lambda\pi\cot\left[\frac{\pi}{2}\Gamma_i^{\alpha}\right]x_i^{\theta}V_i^{\gamma-1}+n(\beta-1)(\alpha-1)\times\sum_{i=1}^{n}\frac{\gamma x_i^{\theta}V_i^{\gamma-1}}{1-V_i^{\gamma}}\times\sum_{i=1}^{n}\frac{e^{1-V_i^{\gamma}\Gamma_i}\alpha\pi\gamma\sin\left[\frac{\pi}{2}(\Gamma_i)^{\alpha}\right]x_i^{\theta}V_i^{\gamma-1}}{\tau_i\left[\frac{\pi}{2}\Gamma_i^{\alpha}\right]}$$

$$\frac{\partial\ell}{\partial\theta}=\frac{n}{\theta}+\gamma\lambda\log x_ix_i^{\theta}V_i^{\gamma-1}+\sum_{i=1}^{n}\log x_i+(\gamma-1)\times\sum_{i=1}^{n}\frac{\gamma\lambda\log x_ix_i^{\theta}}{V_{\;i}}+\frac{n}{2}\sum_{i=1}^{n}e^{1-V_i^{\gamma}}\times\Gamma_i^{\alpha-1}\pi\alpha\gamma\lambda\cot\left[\frac{\pi}{2}\Gamma_i^{\alpha}\right]logx_ix_i^{\theta}(\beta+1)\times\sum_{i=1}^{n}\frac{e^{1-V_i^{\gamma}}\Gamma_i^{\alpha-1}\pi\alpha\gamma\lambda\log x_i\sin\left[\frac{\pi}{2}[\Gamma_i^{\alpha}]\right]x_i^{\theta}V_i^{\gamma-1}}{2\left[\tau_i\left[\frac{\pi}{2}\Gamma_i^{\alpha}\right]\right]}+(\alpha-1)\sum_{i=1}^{n}-\frac{logV_iV_i^{\gamma}}{1-V_i^{\gamma}}$$

$$\frac{\partial\ell}{\partial\gamma}=\frac{n}{\gamma}-logV_iV_i^{\gamma}+(\gamma-1)\sum_{i=1}^{n}logV_i+\sum_{i=1}^{n}logV_i^{\gamma}+\frac{n}{2}\sum_{i=1}^{n}e^{1-V_i^{\gamma}}\Gamma_i^{\alpha-1}\pi x_i^{\theta}\cot\left[\frac{\pi}{2}\Gamma_i^{\alpha}\right]\log(\beta+1)\times\sum_{i=1}^{n}\frac{e^{1-V_i^{\gamma}}\Gamma_i^{\alpha-1}\pi\alpha\log V_i\sin\left[\frac{\pi}{2}[\Gamma_i^{\alpha}]\right]V_i^{\gamma}}{2\left[\tau_i\left[\frac{\pi}{2}\Gamma_i^{\alpha}\right]\right]}+(\alpha-1)\sum_{i=1}^{n}-\frac{logV_iV_i^{\gamma}}{1-V_i^{\gamma}}$$

Setting each of these equations to zero and then solving them simultaneously produces the maximum likelihood estimates of the parameters. With regard to the interval estimation of the model parameters, we need to obtain the observed information matrix, J^{-1} as follows:

$$J^{-1}=\begin{pmatrix}J_{\alpha\alpha} & J_{\alpha\beta} & J_{\alpha\lambda} & J_{\alpha\theta} & J_{\alpha\gamma}\\ . & J_{\beta\beta} & J_{\beta\lambda} & J_{\beta\theta} & J_{\beta\gamma}\\ . & . & J_{\lambda\lambda} & J_{\lambda\theta} & J_{\lambda\gamma}\\ . & . & . & J_{\theta\theta} & J_{\theta\gamma}\\ . & . & . & . & J_{\gamma\gamma}\end{pmatrix}$$

4.2 Ordinary Least Square Estimation Method

The ordinary least square (OLS) estimation method is one of the estimation procedures which estimates the parameters by minimizing the objective function. The OLS method was introduced by [16]. If $x_1,x_2,\ldots,x_n$ are the order statistics of a random sample of size n obtained from the ECGPWD. The OLS estimates $\hat{\alpha}_{OLS}$,$\hat{\beta}_{OLS}$,$\hat{\lambda}_{OLS}$,$\hat{\theta}_{OLS}$,$\hat{\gamma}_{OLS}$ for the ECGPWD parameters can be obtained by minimizing the objective function $\tau(\alpha,\beta,\lambda,\theta,\gamma)=\sum_{i=1}^{n}\left[F(x_i|\alpha,\beta,\lambda,\theta,\gamma)-\frac{1}{n+1}\right]^2$ with respect to $\alpha,\beta,\lambda,\theta,\gamma$

5 Monte Carlo simulation

The section discusses the simulation studies carried out to establish the values of the Average Bias (AB) and the root mean square error (RMSE) for the parameters of the ECGPWD employing the MLE and OLS methods. The simulation was done using different parameter values as well as varying the sample size. The random samples of the ECGPWD were obtained using the quantile function. The simulation process was repeated for $N=1000$ times each containing sample sizes $30,50,80,120,150$ and 250. The values of the first and second sets of the parameters are $\alpha=1.0,\beta=1.3,\lambda=1.1,\theta=0.8,\gamma=1.4$ and $\alpha=0.8,\beta=1.1,\lambda=0.7,\theta=0.4,\gamma=1.0$. The AB and the RMSE of the parameters are computed using the relations:
$AB=\frac{1}{N}\sum_{i=1}^{N}(\hat{\tau}_i-\tau)$ and $RMSE=\sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{\tau}_i-\tau)^2}$
where $\tau=\alpha,\beta,\lambda,\theta,\gamma$ Table 1 shows the simulation results for two estimators (MLE and OLS) for $\alpha=1.0,\beta=1.3,\lambda=1.1,\theta=0.8,\gamma=1.4$. It can be deduced from Table 1 that both the AB and MSE estimates are positive and diminishes towards zero as the sample size increases which shows the consistency of the parameters. The AB and the MSE of the MLE are smaller than those produced by the OLS method. This makes MLE a better estimator than the OLS estimation method.

Table 2 shows the simulation results for the two estimators (MLE and OLS) for $\alpha=0.8,\beta=1.1,\lambda=0.7,\theta=0.4,\gamma=1.0$. It can be deduced that both the AB and RMSE estimates again are positive and diminishes to zero as the sample size increases showing the consistency of the parameters. The AB and the RMSE of the MLE are smaller than those produced by the OLS method. This makes MLE a better estimator than the OLS estimation method.

Table 1: $\alpha = 1.0, \beta = 1.3, \lambda = 1.1, \theta = 0.8, \gamma = 1.4$

		MLE		OLS	
Parameter	n	AB	RMSE	AB	RSME
$\hat{\alpha}$	30	2.6589	4.2732	5.8892	36.3279
	50	2.3425	4.0717	4.6120	25.8264
	80	1.9070	3.2142	4.4301	24.0044
	120	1.7924	3.1102	3.7140	17.3512
	150	1.0743	1.5085	2.9217	11.7555
	200	0.9850	1.3191	2.3821	9.0223
	250	0.6790	0.6826	1.8989	6.3861
$\hat{\beta}$	30	4.5865	5.6853	5.1784	28.7970
	50	3.6683	5.0306	4.6263	24.6797
	80	3.4238	4.8763	4.4551	24.1557
	120	3.3799	4.8846	4.4122	23.3421
	150	2.5334	4.0580	4.0751	23.2049
	200	2.4251	3.9713	3.09568	15.7971
	250	2.1075	3.8306	2.5157	11.7415
$\hat{\lambda}$	30	2.7422	3.7731	8.9000	79.2100
	50	2.3009	3.1538	8.0258	71.2915
	80	1.8699	2.6886	8.0194	71.2899
	120	1.5693	2.0516	7.2731	63.5075
	150	1.4339	2.0476	6.2785	55.4551
	200	1.1078	1.3491	5.3968	47.5341
	250	1.0171	1.1827	4.5973	39.7005
$\hat{\theta}$	30	1.0917	1.6563	0.7972	0.6541
	50	1.0248	1.4732	0.7831	0.6135
	80	1.0162	1.4638	0.7729	0.5989
	120	0.9579	1.4194	0.7728	0.5977
	150	0.9554	1.1942	0.7648	0.5853
	200	0.8246	1.1102	0.7645	0.5848
	250	0.2868	0.4730	0.7369	0.5449
$\hat{\gamma}$	30	2.9878	4.2532	0.6807	0.4736
	50	2.5353	4.0910	0.6755	0.4571
	80	2.5124	3.9245	0.6540	0.4477
	120	2.3266	3.5245	0.6396	0.4400
	150	1.4196	1.8075	0.6089	0.3872
	200	1.2549	1.8022	0.5947	0.3731
	250	1.2456	1.6452	0.5536	0.3347

Table 2: $\alpha = 0.8, \beta = 1.1, \lambda = 0.7, \theta = 0.4, \gamma = 1.0$

		MLE		OLS	
Parameter	n	AB	RMSE	AB	RSME
$\hat{\alpha}$	30	2.6589	4.2732	5.8892	36.3279
	50	2.3425	4.0717	4.6120	25.8264
	80	1.9070	3.2142	4.4301	24.0044
	120	1.7924	3.1102	3.7140	17.3512
	150	1.0743	1.5085	2.9217	11.7555
	200	0.9850	1.3191	2.3821	9.0223
	250	0.6790	0.6826	1.8989	6.3861
$\hat{\beta}$	30	4.5865	5.6853	5.1784	28.7970
	50	3.6683	5.0306	4.6263	24.6797
	80	3.4238	4.8763	4.4551	24.1557
	120	3.3799	4.8846	4.4122	23.3421
	150	2.5334	4.0580	4.0751	23.2049
	200	2.4251	3.9713	3.09568	15.7971
	250	2.1075	3.8306	2.5157	11.7415
$\hat{\lambda}$	30	2.7422	3.7731	8.9000	79.2100
	50	2.3009	3.1538	8.0258	71.2915
	80	1.8699	2.6886	8.0194	71.2899
	120	1.5693	2.0516	7.2731	63.5075
	150	1.4339	2.0476	6.2785	55.4551
	200	1.1078	1.3491	5.3968	47.5341
	250	1.0171	1.1827	4.5973	39.7005
$\hat{\theta}$	30	1.0917	1.6563	0.7972	0.6541
	50	1.0248	1.4732	0.7831	0.6135
	80	1.0162	1.4638	0.7729	0.5989
	120	0.9579	1.4194	0.7728	0.5977
	150	0.9554	1.1942	0.7648	0.5853
	200	0.8246	1.1102	0.7645	0.5848
	250	0.2868	0.4730	0.7369	0.5449
$\hat{\gamma}$	30	2.9878	4.2532	0.6807	0.4736
	50	2.5353	4.0910	0.6755	0.4571
	80	2.5124	3.9245	0.6540	0.4477
	120	2.3266	3.5245	0.6396	0.4400
	150	1.4196	1.8075	0.6089	0.3872
	200	1.2549	1.8022	0.5947	0.3731
	250	1.2456	1.6452	0.5536	0.3347

6 Empirical Applications

The applications of the ECGPWD are demonstrated in this section using real data sets. The performance of the ECGPWD in fitting the data sets is compared to its sub-models and other existing distributions. The comparison is done using the log-likelihood ℓ, Akaike Information Criterion (AIC), corrected Akaike Information Criterion (AICc), Bayesian Information Criterion (BIC), *Crame*′r-von Mises test (W*), and the Kolmogorov-Smirnov test (K-S).

6.1 First Application

The data set represents 101 observations that show the failure times (in hours) of Kevlar 49/epoxy strands subjected to constant sustained pressure at a 90 percent stress level. Table 11 in the Appendix shows the data set and the data can be found in [17] and was recently used by [18]. The descriptive statistics for the Kevlar 49/epoxy data are shown in Table 3. It has a mean of 1.025, a median value of 0.800, and a standard deviation of 1.119. The results also show the data is skewed to the right and it is highly peaked as compared to the normal distribution since the kurtosis value is more than three.

Table 3: Discriptive Statistics for the Kevlar dataset

Mean	Median	standard Dev.	Skewness	Kurtosis
1.0250	0.8000	1.1190	3.0800	14.470

The TTT transform graph is used to obtain the hazard failure rate shape for the data. The TTT transform plot of the Kevlar 49/epoxy data is given in Figure 3. The failure

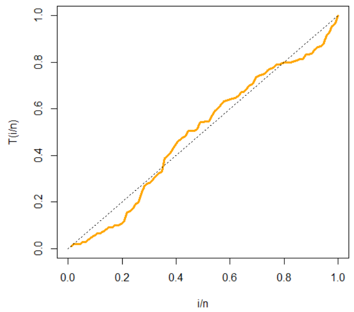

Figure 3: TTT plots of the Kevlar data set

rate is first a convex shape, followed by a concave shape, and then a convex shape again. This indicates that the data set exhibits a modified bathtub shape.

The maximum likelihood parameter estimates as well as their corresponding standard errors for the Kevlar data are illustrated in Table 4 for the ECGPWD, its sub-models, and other existing distributions. The other models used in the comparison are: Nadarajah Haghighi generalized power Weibull distribution (NHGPWD), Weibull Nadarajah Haghighi distribution (WNHD), generalized power Weibull distribution (GPWD), and exponentiated generalized Poisson inverse exponential distribution (EGPIED) models [17]. To test for the significance of the parameters, the standard error test was used. The results in Table 4 indicate that all the parameters of the ECGPWD are significant except at the 5 percent level of significance. The log-likelihood, goodness of fit statistics and the information criteria values of the ECGPWD and the other competing models for the Kevlar data are presented in Table 5. It can be observed that the ECGPWD gives a better fit to the data than the other models since it has the highest log-likelihood value and smallest value of the AIC, AICc, BIC, and the K-S.

The PDFs and the CDFs of the fitted models as well as the empirical densities fitted on the histogram for the Kevlar data are shown in Figure 4. The PDF and CDF of the ECGPWD followed the empirical density and the empirical CDF better than the other competing models.

The variance-covariance matrix, V_{cov} of the parameters of the ECGPWD for the Kevlar data was estimated and presented below as

$$V_{cov} = \begin{pmatrix} 3.7855 & -0.1454 & 0.0924 & -0.1690 & 0.1214 \\ -0.1454 & 0.0123 & -0.0034 & -0.0274 & 0.0268 \\ 0.0924 & -0.0034 & 0.0841 & 0.0121 & -0.0692 \\ -0.1690 & -0.0274 & 0.0121 & 0.1908 & -0.1844 \\ 0.1214 & 0.0268 & -0.0692 & -0.1844 & 0.2244 \end{pmatrix}$$

The variances of the MLE of the parameters are: $\text{var}(\hat{\alpha}) = 3.7855$, $\text{var}(\hat{\beta}) = 0.0123$, $\text{var}(\hat{\lambda}) = 0.0841$, $\text{var}(\hat{\theta}) = 0.1908$, and $\text{var}(\hat{\gamma}) = 0.2244$. The 95 percent confidence intervals of the estimated parameters are: (0, 6.806), (0,0.343), (0.399, 1.395), (0.152, 2.276), and (0.120, 1.512).

The P-P plots of the ECGPWD and the other competing models are shown in Figure 5. The plots indicate that the ECGPWD gives the Kevlar data set a better fit than the competing models since it has almost all of its points on the diagonal line.

Table 4: Parameters estimates of the kevlar dataset

Model	Parameter	Estimates	Standard Error
ECGPWD	α	2.4820	2.2060
	β	0.1410	0.1030
	λ	0.8970	0.2540
	θ	1.2140	0.5420
	γ	0.8160	0.3550
ECWD	α	2.7990	1.0270
	β	0.1510	0.1160
	λ	0.8160	0.2860
	γ	1.0270	0.2160
ECGNHD	α	2.8810	2.3830
	β	0.1510	0.1250
	λ	0.8160	0.4130
	γ	1.0230	0.2850
ECRD	α	0.3900	0.5280
	β	0.4190	0.5920
	λ	0.1570	0.0330
NHGPWD	α	3.4350	1.6180
	β	0.4140	0.1550
	λ	5.6380	0.0220
	γ	1.0070	0.0020
	θ	114.5700	0.0000
WNHD	α	1.1450	0.4310
	β	0.6930	0.1890
GPWD	α	0.5930	0.4650
	β	0.7620	0.1250
	λ	1.3080	0.6450
EGPIED	λ	26.0690	0.0090
	c	7.3200	1.7700
	d	0.1750	0.0190
	γ	0.0020	0.0020

Table 5: Model selection criteria

Model	ℓ	AIC	AICc	BIC	$K-S$
ECGPWD	-93.18	196.361	197.058	208.970	0.006
ECWD	-93.19	196.378	197.076	208.987	0.000
ECNHDD	-93.19	196.378	197.075	208.987	0.000
ECRD	-97.63	205.264	205.961	217.873	0.000
NHGPWD	-95.01	199.89	200.580	213.253	0.721
NH	-103.34	210.683	211.373	215.913	0.005
GPWD	-102.8	245.124	245.814	252.969	0.000
EGPIED	-116.66	241.314	241.947	251.774	0.000

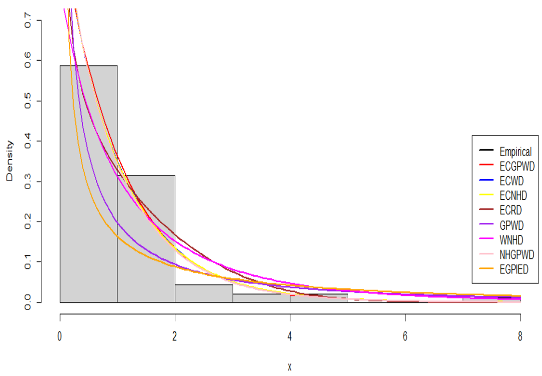

Figure 4: PDF plots of the Kevlar data set

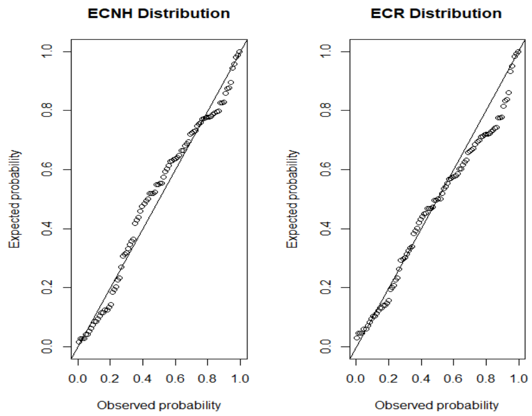

Figure 7: P-P plot2 of the Kevlar data set

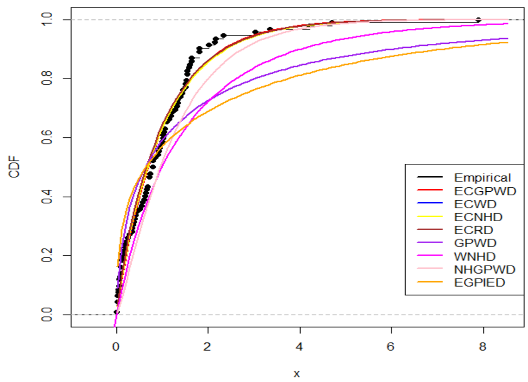

Figure 5: CDF plots of the Kevlar data set

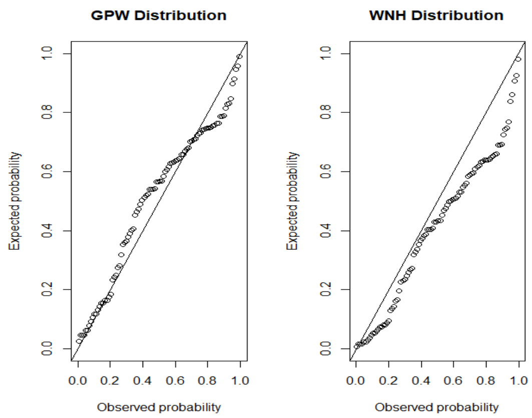

Figure 8: P-P plot3 of the Kevlar data set

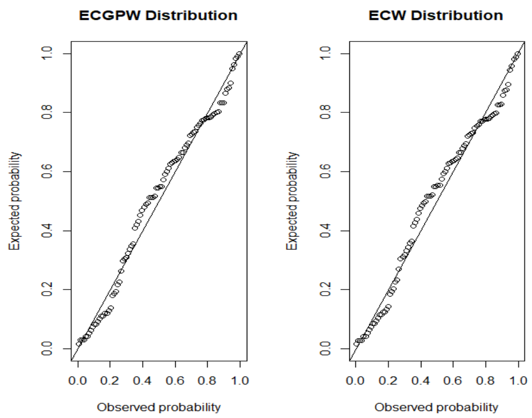

Figure 6: P-P plot1 of the Kevlar data set

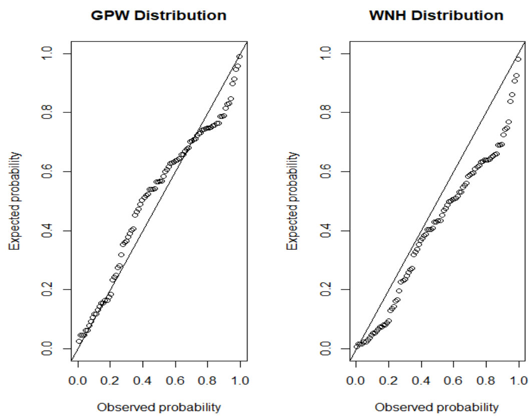

Figure 9: P-P plot4 of the Kevlar data set

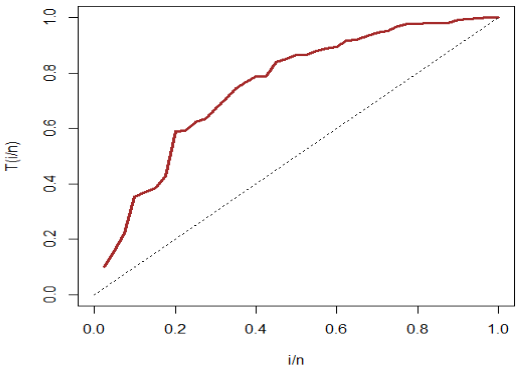

Figure 10: TTT plots of leukemia patients' survival data

6.2 Second Application

The second data set is in Table 12 under the appendix which represents the survival times of leukemia patients and can be found in [19]. The descriptive statistics of the data set on leukemia are shown in Table 6. The mean, median, the standard deviation of the observed data set are respectively found to be 1137, 1222, and 481.6. The results also indicate that the observed data is skewed to the left and less peaked in comparison to the normal distribution. The TTT plot of the failure rate of the leukemia data shows an increasing shape as the curve falls above the 45^o line as shown in Figure 6.

Table 6: Descriptive Statistics for the leukemia data set

Mean	Median	standard Dev.	Skewness	Kurtosis
1137	1222	481.6	-0.510	-0.66

The maximum likelihood parameter estimates and their corresponding standard errors are shown in Table 7 for the ECGPWD, its sub-models, and other existing models. The existing models used in the comparison are: generalized inverse Weibull (GIW), exponentiated Lomax (ELx), generalized odd inverse exponential Weibull (GOIEW), and generalized odd inverse exponential Lomax (GOIEL) distributions [19]. The evidence from Table 7 indicates that all the parameters of the ECGPWD are significant at a 5 percent level of significance except λ. The log-likelihood, information criteria, and goodness of fit statistics of the models fitted are investigated and presented in Table 8. The results shows that ECGPWD performs better in fitting the given data set since it recorded the highest log-likelihood value, and the smallest values of the goodness of fit statistics of AIC, BIC, AICc, and the K-S. The histogram plots as well as the densities for the models fitted together with their CDFs of the leukemia data are shown in Figure 7. The fitted models mimic the empirical PDF and CDF of the leukemia data set.

Table 7: Parameters estimates of the leukemia data set

Model	Parameter	Estimates	Standard Error
ECGPWD	$\hat{\alpha}$	12.8430	0.0061
	$\hat{\beta}$	0.2360	0.1194
	$\hat{\lambda}$	0.0084	0.0102
	$\hat{\theta}$	0.5849	0.1580
	$\hat{\gamma}$	2.9128	0.3288
ECWD	$\hat{\alpha}$	0.8841	0.8720
	$\hat{\beta}$	3.5102	0.8146
	$\hat{\lambda}$	0.0074	0.0531
	$\hat{\gamma}$	0.8146	0.8831
ECED	$\hat{\alpha}$	0.0478	0.0119
	$\hat{\beta}$	55.0940	1.0484×10^{-6}
	λ	0.0018	0.0003
ECNHD	$\hat{\alpha}$	2.6128	0.1464
	$\hat{\beta}$	0.6839	0.4262
	$\hat{\lambda}$	0.0015	0.0031
	$\hat{\gamma}$	1.0310	1.0783
GOIEWD	$\hat{\alpha}$	1.7676	0.5999
	$\hat{\beta}$	42.2469	0.0143
	$\hat{\gamma}$	0.0079	0.0078
	$\hat{\theta}$	0.8782	0.1352
GIWD	$\hat{\gamma}$	20.4052	6.3882
	$\hat{\theta}$	1.1893	0.1184
	$\hat{b}$	69.0472	1.5874
GOIELD	$\hat{\alpha}$	6.7875	0.0135
	$\hat{\beta}$	55.0620	6.6×10^{-4}
	$\hat{\gamma}$	$5.2442\times10-3$	1.6279×10^{-3}
	$\hat{\theta}$	3.3683	0.0468
ELxD	$\hat{\alpha}$	207.5900	5.0510×10^{-5}
	$\hat{\lambda}$	0.0106	0.0047
	$\hat{\theta}$	1.2399	0.1393

Table 8: Model selection criteria

Model	ℓ	AIC	AICc	BIC	$K-S$
ECGPWD	-309.49	628.989	630.754	637.433	0.260
ECWD	-312.29	634.571	636.336	643.015	0.190
ECED	-310.48	630.953	632.718	639.398	0.117
ECNHD	-310.39	630.788	632.553	639.233	0.198
GOIEWD	-330.48	668.927	669.980	675.972	0.165
GIWD	-352.32	710.645	711.261	715.929	0.006
GOIELD	-338.75	685.507	686.559	692.552	0.010
ELxD	-351.94	709.870	710.486	715.154	0.009

The variance-covariance matrix, V_{cov} for the estimated parameters of the ECGPWD for the leukemia data is shown below as

$$V_{cov}=\begin{pmatrix} 3.696D & 1.464C & 2.427D & 7.889D & -7.824B \\ 1.464C & 1.426A & 1.062B & -1.614A & 7.323B \\ 2.427D & 1.062B & 1.037C & -1.421B & -2.649B \\ 7.889D & -1.614A & -1.421B & 2.497A & -5.927A \\ -7.824B & 7.323B & -2.649B & -5.927A & 1.766 \end{pmatrix}$$

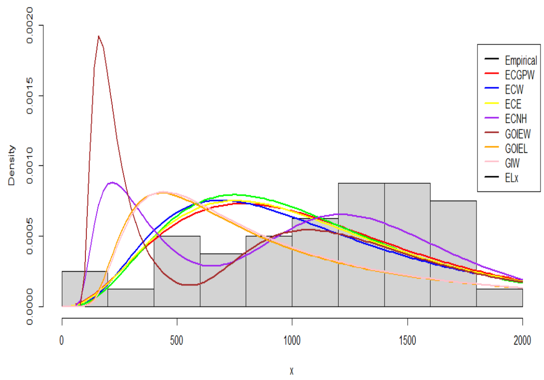

Figure 11: PDF plots of leukemia patients' survival data

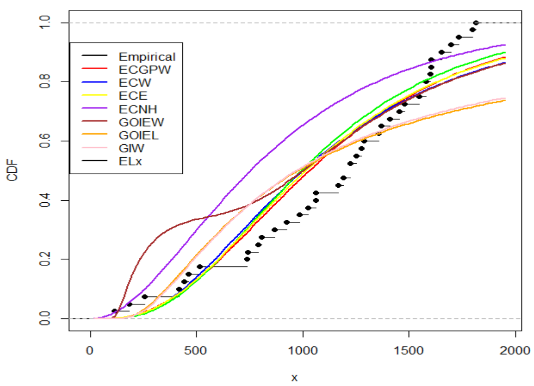

Figure 12: CDF plots of leukemia patients' survival data

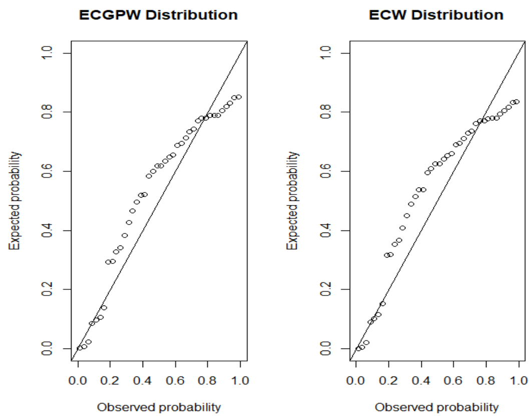

Figure 13: P-P plots of leukemia patients' survival data

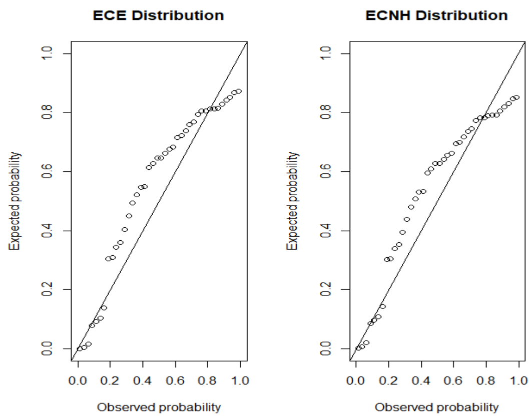

Figure 14: P-P plot2 of leukemia patients' survival data

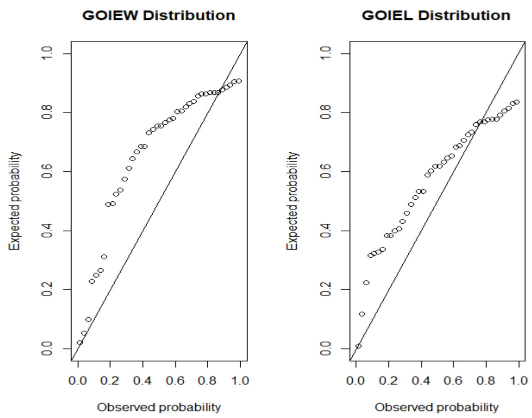

Figure 15: P-P plot3 of leukemia patients' survival data

where $A = 10^{-2}, B = 10^{-3}, C = 10^{-4}, D = 10^{-5}$ The variances of the MLE of the parameters are: $var(\hat{\alpha}) = 3.696 \times 10^{-5}, var(\hat{\beta}) = 1.4264 \times 10^{-2}, var(\hat{\lambda}) = -1.0366 \times 10^{-4}, var(\hat{\theta}) = 2.4969 \times^{-2}$, and $var(\hat{\gamma}) = 1.7658$. The 95 percent confidence intervals of the estimated parameters are: (12.8310, 12.8550), (0.0020, 0.4700), (0, 0.0284), (0.2752, 0.8946), and (0.3083, 5.5173).

Figure 8 shows the P-P plots of the ECGPWD and that of the other fitted models. The ECGPWD gives a better fit of the data set since almost all of its plotted points lie on the 45^o line.

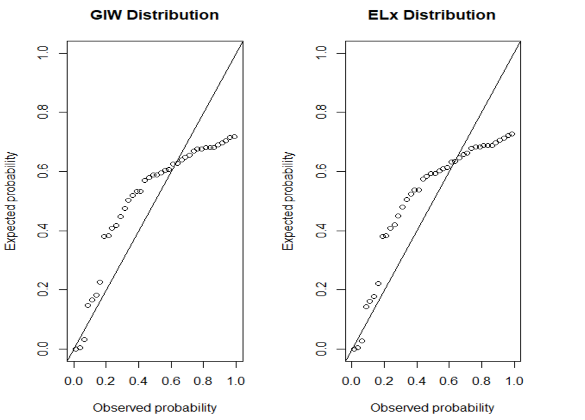

Figure 16: P-P plot4 of leukemia patients' survival data

7 ECGPWD Location-Scale Regression Model

The logarithm of the ECGPW location-scale regression model is presented in this section. If X is a random variable that follows the ECGPWD, then another random variable Y also follows the log ECGPWD (LECGPW). Assuming that $\lambda = e^{\frac{-\mu}{\sigma}}$ and $\theta = \frac{1}{\sigma}$, then the density function of the LECGPW regression model is given by

$$f_r(y;\zeta) = \frac{\rho \Gamma_i^{\alpha} \sin\left[\frac{\pi}{2}\Gamma_i^{\alpha}\right]\left[1-\cos\left[\frac{\pi}{2}\Gamma_i^{\alpha}\right]\right]^{\beta-1}}{2\sigma\tau\left[\Gamma_i\right]} \quad (15)$$

where $\zeta = \alpha, \beta, \gamma, \sigma, \mu$ $\rho = \pi\alpha\beta\gamma e^{z}$, $\tau = 1+e^{z}$, $z = \frac{y-\mu}{\sigma}$, $\alpha > 0, \beta > 0$, $\gamma > 0$ $\Gamma_i = 1-e^{1-\tau^{\gamma}}$, and α, β, γ are shape parameters, $\sigma > 0$ is a scale parameter, and $\mu \in R$ is a location parameter. The density of the LECGPW regression model is defined for $y \in R$. The LECGPW regression model converges to the log extended cosine generalized Weibull (LECW) regression model when $\gamma = 1$.
The CDF corresponding to equation (15) is given by

$$F_r(y:\zeta) = \left[1-cos\left(\frac{\pi}{2}\left(1-e^{1-(1+e^{\frac{y-\mu}{\sigma}})^{\gamma}}\right)^{\alpha}\right)\right]^{\beta} \quad (16)$$

The survival function of the LECGPWD is given by

$$S_r(y:\zeta) = 1-\left[1-cos\left(\frac{\pi}{2}\left(1-e^{1-(1+e^{\frac{y-\mu}{\sigma}})^{\gamma}}\right)^{\alpha}\right)\right]^{\beta} \quad (17)$$

Using the PDF in equation (15), the LECGPW location-scale regression model is defined in the following form.

$$y_i = q_T^i\phi + \sigma z_i, i = 1,2,3,...,n$$

where $\mu = q_T^i\phi$ represents the location parameter depending on several covariates, $\phi = (\phi_0, \phi_1, ...\phi k)'$ are the regression parameters, k represents number of covariates, $q_i = (q_{i1}, q_{i2}, ...q_{ik})'$ are the covariates and z_i represents the error term that follows the PDF as given in equation (15). The estimates of the parameters are obtained using the MLE technique. The log-likelihood function of the LECGPW model is presented as

$$\ell = n\log(\pi\alpha\beta\gamma) - \frac{n}{2}\log\sigma + \sum_{i=1}^{n} z_i + (y-1)\times \sum_{i=1}^{n}\log(1+e^{z_i}) + (1-V_i) + (\alpha-1)\sum_{i=1}^{n}\log\Gamma_i + \sum_{i=1}^{n}\log\sin\left[\frac{\pi}{2}\Gamma_i^{\alpha}\right] + (\beta-1)\sum_{i=1}^{n}\log\tau_i\Gamma_i^{\alpha}$$

where $z_i = \frac{(y_i - q_i^T)}{\sigma}$ and n represents the sample size. Estimates for the parameters are found through the maximization of the log-likelihood function. Cox-Snell residuals (Cox and Snell, 1968) are used to check the adequacy of the LECGPW model. The Cox-Snell residuals of the LECGPW model are given as

$$\hat{r}_i = -\log\left(S_r\left(y_i|\hat{\alpha}, \hat{\beta}, \hat{\gamma}, \hat{\mu}, \hat{\sigma}\right)\right), i = 1,2,3,...,n,$$

where $\left(S_r\left(y_i|\hat{\alpha}, \hat{\beta}, \hat{\gamma}, \hat{\mu}, \hat{\sigma}\right)\right)$. If the LECGPW model fits the data, its Cox-Snell residuals are expected to follow the standard exponential distribution.
Application of the LECGPW model illustrated by modeling the birth weight of babies (x_i), mothers' age at the time of delivery (ϕ_1) and the gestation period of pregnancy (ϕ_2). The data was obtained from the Bole hospital for 42 babies delivered in the month of January 2022. The observations are shown in Tables 13, 14, and 15 in the appendix. The data is modeled by the regression equation;

$$y_i = \phi_0 + q_1\phi_{1i} + q_2\phi_{2i} + \sigma z_i, i = 1,2,3,...,42$$

where $y_i = \log(x_i)$ follows the LECGPW distribution. The performance of the LEGPW regression model was compared with LECW location-scale regression model. The parameter estimates of the LECGPW and LECW location-scale regression models are shown in Table 9. From the goodness of fit statistics shown in Table 10, it can be established that the LECGPW regression model has performed better in fitting the data set than the log location scale Weibull regression model since it has the highest log-likelihood value and has the smallest values of AIC, AICc, BIC and K-S statistics. Mothers' age negatively affects birth weight but it however not significant in both models at 5 percent a level of significance and the gestation period positively affects birth weight at the 5 percent level of significance for the LECGPW location-scale regression model.

From the P-P plots in Figure 9, the LECGPW regression model performs better in fitting the data set on the birth weight of babies.

Table 9: Parameters estimates of the leukemia data set

Model	Parameter	Estimates	Standard Error
LECGPW	$\hat{\alpha}$	10.0750	0.0061
	$\hat{\beta}$	9.4985	0.0017
	$\hat{\gamma}$	6.5283	0.0221
	$\hat{\sigma}$	0.5426	0.0622
	$\hat{q}_0$	0.7451	0.1815
	$\hat{q}_1$	-0.0010	0.0015
	$\hat{q}$	0.0187	0.0036
LECW	$\hat{\alpha}$	2.9943	0.0068
	$\hat{\beta}$	22.8844	0.0008
	$\hat{\sigma}$	0.2519	0.0274
	$\hat{q}_0$	-0.2401	0.1283
	$\hat{q}_1$	-0.0007	0.0014
	$\hat{q}_2$	0.0179	0.0035

Table 10: Model selection criteria

Model	ℓ	AIC	AICc	BIC	$K-S$
LECGPW	59.09	-104.18	-100.89	-92.02	0.07
LECW	58.24	-104.47	-102.07	-94.05	0.11

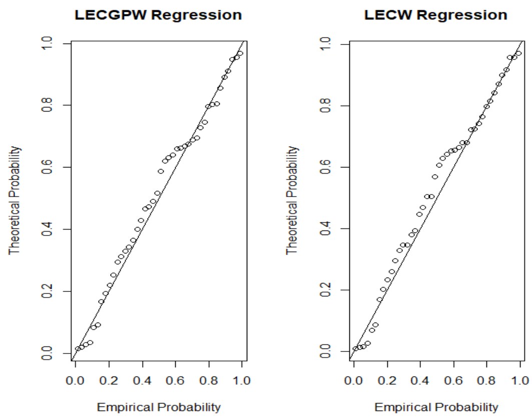

Figure 17: P-P plots of Cox-Snell residuals for the models

8 Concluding Remarks

This study introduces a novel model known as the extended cosine generalized power Weibull distribution. Various statistical properties, such as moments, moment-generating function, characteristic function, incomplete moment, and order statistics are examined in the study. Additionally, four special sub-models were derived from the proposed model. The study investigates the shapes of the density and failure rates for specific parameter values. The failure rates of the ECGPWD display various non-monotonic patterns. The parameters of the ECGPWD are estimated using maximum likelihood and ordinary least square methods. Monte Carlo simulations are conducted to assess the performance of the parameters under these estimation methods. The estimates obtained from the maximum likelihood for the average biases and the mean root square were however the least, making it the best estimator. The study also defined a log-location scale extended cosine generalized power Weibull regression model as a competitive regression model to other existing regression models. Three data sets were used to demonstrate the application of the proposed distribution. The proposed model proves to be more flexible in fitting the data set. Further studies can be done to estimate the parameters of the distribution using other estimation methods other than the MLE and OLE methods. Not all but also, the data sets used in this study are complete data sets. However, incomplete samples may arise in different fields of studies; such as in a follow-up study in medical investigation, cancer patients may die before the study ends or survived beyond the duration of the study. Hence further studies can consider the use of incomplete data to demonstrate the applications of the proposed distributions

Availability of data and material

All the data used by the authors are presented in the Appendix

Conflict of interest

The authors declare that they have no conflict of interest.

Funding

The authors did not receive any financial support for the research.

Authors' contributions

All authors contributed equally work.

Acknowledgement

The authors would like to thank the anonymous reviewers for carefully reading the article and also for their constructive and valuable comments, which have improved the paper in its present form.

References

[1] W. Weibull, A statistical distribution function of wide applicability, *Journal of Applied Mechanics*, **18**, 293-297 (1951).

[2] R. Sujata, M. B. Rajash, Bathtub distribution, *Communication in Statistics-Theory and methods*, **17**, 2597-2621, (1988).
[3] M. A. Selim, H. M. Salem, The exponentiated generalized power Weibull distribution: properties and applications, *Journal of Statistics Applications and Probability*, **12**(1), 19-38, (2023).
[4] G. S. Mudholkar, D. K. Srivastava, M. Freimer, The exponentiated Weibull family: a re-analysis of the bus-motor-failure data, *Tchnometrics*, **37**, 436-445 (1995).
[5] A. M. Sarhan, M. Zaindin, Modified Weibull distribution, *Applied Science*, **11**m 123-136 (2009).
[6] S. J. Almalki, J. Yuan, A new modified Weibull distribution, *Reliability Engineering and System Safety*, **111**(C), 164-170 (2013).
[7] M. Nikulin, F. Haghighi, A Chi-squared test for the generalized power Weibull family for the head-and-neck cancer censored data, *Journal of Mathematical Science*, **133**(3), 1333-1341 (2006).
[8] Z. Mahmood, T. M. Jawa, N. Sayed-Ahmed, E. M. Khalid, A. H. Muse, A. H. Tolba, An extended cosine generalized family of distributions for reliability modelling: characteristics and applications with simulation study, *Math. Probl. Eng.*, **2022**, 3634698 (2022).
[9] S. Nadarajah, S. Kotz, Beta trigonometric distribution, *Portuguese Economic Journal*, **5**(3), 207-224 (2006).
[10] L. Souza, New trigonometric classes of probabilistic distributions, *thesis, University Federal Rural de Pernambuco,* Brazil, (2015).
[11] C, Chesneau, H. S. Bakouch, T. Hussain, A new class of probability distribution via cosine and sine functions with applications. *Communication in Statistics-Simulation and Computation*, **48**(80), 2287-3000 (2019).
[12] Z. Mahmood, C. Chesneau, M. H. Tahir, A new Sine-G family of distributions: properties and applications, *Bull. Comput. Appl. MAth.*, **7**(1), 53-82 (2019).
[13] C. Chesneau, H. S. Bakouch, T. Hussain, B. Para, The cosine geometric distribution with count data modeling. *Journal of Applied Statistics*, **48**(1), 124-137 (2020).
[14] A. A. Al-Babtain, I. Elbatal, C. Chesneau, M. Elgarhy, Sine Topp-Leone-G family of distributions: theory and applications, *Open physics*, **18**(1), 574-593 (2020).
[15] A. Ali, Sine power Lomax model with application to bladder cancer data. *Nanoscience and NanoTechnology Letters*, **12**(5), 677-684 (2020).
[16] J. J. Swain, S. Venkatraman, J. R. Wilson, Least Squares estimation of distribution functions in Johnson's translation system. *Journal of Statistical Computation and Simulation*, **29**, 271-297 (1988).
[17] R. Barlow, R. Toland, T. Freeman, A Bayesian analysis of stress-rupture life of Kevlar 49/epoxy spherical pressure vessels, *In proceedings of a conference on applications statistics.* Marcel Dekker, New York, 1984.
[18] A. I. Abonongo, Nadarajah Haghighi Generalized power Weibull: properties and applications. *PhD thesis. www.udsspace.uds.edu.gh* (2022).
[19] A. Yakubu, A. Luguterah, S. Nasiru, On the odd inverse exponential class of distributions: properties, applications and cure fraction regression. *Journal of Statistics and Management systems*, **25**(4), 805-836 (2022).

9 Appendix

Table 11: failure time (in hours) of Kelar 49/epoxy strands data

0.01	0.02	0.02	0.02	0.03	0.03	0.04	0.05
0.06	0.07	0.07	0.08	0.09	0.10	0.10	0.11
0.11	0.12	0.13	0.18	0.19	0.20	0.23	0.24
0.29	0.34	0.35	0.36	0.38	0.40	0.42	0.43
0.52	0.54	0.56	0.60	0.63	0.65	0.67	0.68
0.72	0.72	0.72	0.73	0.79	0.79	0.80	0.80
0.85	0.90	0.92	0.95	0.99	1.00	1.01	1.02
1.03	1.05	1.10	1.10	1.15	1.18	1.20	1.29
1.31	1.33	1.34	1.40	1.43	1.45	1.50	1.51
1.53	1.54	1.54	1.55	1.58	1.60	1.63	1.64
1.80	1.80	1.81	2.02	2.14	2.17	2.33	3.03
3.34	4.20	4.69	7.89				

Table 12: Data on survival times of leukemia patients

115	461	807	1062	1251	1408	1578	1696
181	516	865	1063	1277	1455	1578	1735
255	739	924	1165	1290	1478	1599	1799
418	743	983	1191	1357	1222	1603	1815
441	789	1024	1222	1369	1549	1605	1852

Table 13: Birth weight of babies' data

5.8	4.2	6.4	4.5	5.8	6.8	5.2
6.1	7.5	8.0	8.6	7.1	6.6	7.0
6.6	6.3	7.3	8.5	5.5	7.5	6.0
8.3	9.5	6.9	7.2	7.7	7.9	7.3
8.0	6.0	9.0	7.0	7.3	7.8	7.0
6.9	10.0	8.6	8.0	10.0	8.5	8.9

Table 14: Mothers' age in years

24	20	26	41	20	28	20
19	20	18	29	31	30	31
27	19	23	30	22	23	21
24	19	29	24	26	31	27
20	37	35	21	27	37	27
22	32	24	21	20	28	20

Table 15: Gestation period in weeks

33	33	34	45	37	37	37
38	38	38	38	38	38	39
39	39	39	39	39	39	40
40	40	40	40	40	40	40
40	40	41	41	41	41	41
41	41	42	42	44	45	44