Asian Journal of Probability and Statistics

Volume 25, Issue 1, Page 13-34, 2023; Article no.AJPAS.106376
ISSN: 2582-0230

Odd Chen Exponential Distribution: Properties and Applications

Henry Otoo [a*], Justice Inkoom [a] and Eric N. Wiah [a]

[a] Department of Mathematical Sciences, UMaT, Tarkwa, Ghana.

Authors' contributions

This work was carried out in collaboration among all authors. All authors read and approved the final manuscript.

Article Information

DOI: 10.9734/AJPAS/2023/v25i1535

Open Peer Review History:
This journal follows the Advanced Open Peer Review policy. Identity of the Reviewers, Editor(s) and additional Reviewers, peer review comments, different versions of the manuscript, comments of the editors, etc are available here: https://www.sdiarticle5.com/review-history/106376

Original Research Article

Received: 09/07/2023
Accepted: 13/09/2023
Published: 19/09/2023

Abstract

In this study, a new statistical distribution with three parameters called the Odd-Chen Exponential has been proposed. The statistical properties of the proposed distribution, such as the quantile, moments, incomplete moments, moment-generating function, and mean residual life, were developed. The density shows different shapes, making it more flexible for analyzing different forms of data. The hazard function also exhibits different shapes, including the well-known bathtub shape, which means that the distribution is flexible with real-life data. To estimate the distribution parameters, ordinary least squares estimators, Cramér-von Mises estimators, and maximum likelihood estimators were derived. The results were compared using a Monte Carlo simulation. Two-time datasets; one from the mining field and the other from survival analysis, were used to check the applicability of the proposed distribution. The results revealed that the OCE distribution performed better than the Odd Chen Weibull, Odd Chen Rayleigh, Rayleigh, Cauchy, Generalised Inverse Weibull and the Modified Extended Chen distributions.

Keywords: Odd chen exponential; Odd chen generator; exponential; hazard function.

**Corresponding author: Email: hotoo@umat.edu.gh;*

1 Introduction

The challenge in empirical statistical modeling is to find parameter estimates for a distribution function that are as close to the true values of the theoretical model parameters as possible. The accuracy of parametric statistical inference and dataset modeling is largely determined by how well the probability distribution fits the given dataset once all distributional assumptions are met. The concept of statistical distribution is widely used in many disciplines. It is used in the area of computer science, engineering, actuarial science, social science and health sciences due to its importance when it comes to modelling datasets generated from some random experiment [1]. However, at times researchers face several challenges since most of the existing distributions are not best suited for modelling some existing and new datasets [2]. Therefore, there is the need to continuously research old distributions and even create new ones to improve upon the goodness-of-fit of the existing distribution and as well as deducing more powerful distributions to help model datasets. As a result, several studies have developed a more flexible distribution. See [3- 9].

For modeling lifetime datasets, the Weibull distribution is one of the most commonly used parametric distributions. However, because most lifetime data exhibit non-monotonic failure rates, Weibull's inability to exhibit bathtub-shaped failure rate functions is one of its major drawbacks. As a result, several modifications to the Weibull distribution have been made over time to make it more flexible for modeling present-day datasets with varying shapes of hazard rate functions [10]. The Frechet, Reverse Weibull, Log-Weibull, Generalized Weibull, etc. are a few of such modifications of the Weibull distribution. As new distributions emerge now and then the chances that there are now some distributions that can compete with Weibull distribution and other known distributions are very high.

The Chen family of distributions is also known for some reliable modelling in several areas of research. The Exponential, Gamma, Extreme Value, Burr Distributions, etc. are also some examples of good-performing statistical distributions. The Odd Chen G distribution though a new distribution appears to have some promising features that put it in the same league as some of the top distributions known in survival analysis and engineering.

The Odd Chen G distribution (OCG) as defined by [10] is a generator distribution with only two parameters which happen to be shape parameters. The absence of a scale parameter in the OCG distribution makes it difficult to control the spread and the variability of the distribution. To prevent this deficiency, a modification of the OCG with a distribution that has a scale parameter will do.

There have been several modifications of the Odd Chen G distribution [10, 11, 13]. For instance [12] proposed a new family of distribution called the Odd Chen-G family of distribution using the T-X approach. [11], also proposed an exponentiated Odd Chen distribution capable of modeling characteristics of the data sets such as skewness, symmetric, kurtosis and various shapes of failure rate (increasing – decreasing – J – inverse). However, some of these well-known models are insufficient to model these kinds of data sets.

The exponential distribution happens to be one of the simple distributions that have wide usage and also has the scale parameter present once the rate parameter is further parameterized. Modifying the OCG with such a distribution will provide control over the variability and also generate a new member of the Odd Chen family of distributions. Therefore, this study seeks to propose a new distribution namely, the Odd Chen Exponential, (OCE). This study will comprehensively explore the properties, estimation methods, and practical applications associated with the OCE.

2 Formulation of the Odd Chen Exponential Distribution

2.1 Odd Chen G Distribution

The Odd Chen G distribution can be obtained by performing integration and differentiation on the Chen distribution. The cumulative density function (cdf) (denoted by F(t)) for Chen distribution is given by; $F(t) = 1 - \exp\left(\lambda\left(1 - e^{t^{\beta}}\right)\right), t > 0$ [7]. Suppose $G(x;\psi)$ is the baseline cdf of an arbitrary continuous random variable X on any continuous support say $(-\infty,\infty)$ and ψ is a $(p \times 1)$ vector of associated parameters, the cdf of the OC family of distributions is defined as:

$$F(x) = \int_0^{\frac{G(x;\psi)}{1-G(x;\psi)}} f(t)dt = 1 - \exp\left(\lambda\left(1 - e^{\frac{G(x;\psi)}{1-G(x;\psi)}}\right)\right), x > 0, \lambda > 0, \beta > 0 \quad (2.1)$$

where λ and β are extra shape parameters [10].

By differentiating the cdf in equation 2.1, the probability density function (pdf) of equation 2.1 is given by;

$$f(x) = \lambda\beta g(x;\psi)G(x;\psi)^{\beta-1}[1 - G(x;\psi)]^{-(\beta-1)} \exp\left(\frac{G(x;\psi)}{1-G(x;\psi)}\right)^{\beta} \exp\left(\lambda\left(1 - \exp\left(\frac{G(x;\psi)}{1-G(x;\psi)}\right)^{\beta}\right)\right), x > 0 \quad (2.2)$$

The exponential distribution with a location parameter, α has its cdf defined as;

$$M(x) = 1 - e^{-x/\alpha} \quad (2.3)$$

The pdf of the exponential distribution is also defined as;

$$m(x) = \frac{1}{\alpha} e^{-x/\alpha} \quad (2.4)$$

Now, substituting equation 2.3 into equation 2.1 gives the cdf of the Odd Chen-Exponential distribution. Therefore,

$$F(x) = 1 - \exp\left(\lambda - \lambda \exp\left(e^{x/\alpha} - 1\right)^{\beta}\right) \quad (2.5)$$

Substituting equations 2.4 and 2.3 into equation 2.2, the pdf for the Odd Chen-Exponential (OCE) distribution is given by;

$$f(x) = \frac{\lambda\beta}{\alpha} e^{-x/\alpha}\left(e^{-x/\alpha}\right)^{-(\beta+1)}\left(1 - e^{-x/\alpha}\right)^{(\beta-1)} \phi_{(x,\alpha)} \quad (2.6)$$

where,

$$\phi_{(x,\alpha)} = \exp\left(\left(e^{x/\alpha} - 1\right)^{\beta}\right) \exp\left(\lambda\left(1 - \exp\left(\left(e^{x/\alpha} - 1\right)^{\beta}\right)\right)\right) \quad (2.7)$$

Survival and failure rate functions are critical in reliability analysis and other fields. The survival (or reliability) function expresses the likelihood of successfully completing a specified task under specified conditions over a specified period. Thus, reliability can be used to assess a system's ability to perform its function correctly. S(x) is the mathematical expression for the survival function and is given by;

$$S(x) = 1 - F(x) \quad (2.8)$$

Substituting the CDF of the OCG distribution into equation 2.8 yields the survival function of the distribution. Thus,

$$S(x) = \exp\left(\lambda - \lambda \exp\left(e^{x/\alpha} - 1\right)^{\beta}\right) \quad (2.9)$$

The hazard function is a primary tool in the field of data analysis. The hazard function can determine the type of distribution that can be used in modelling data sets. The failure rate or the hazard function, $h(x)$, is the instantaneous failure rate and is mathematically expressed as;

$$h(x) = \frac{f(x)}{1-F(x)} = \frac{f(x)}{S(x)} \quad (2.10)$$

Substituting the survival function, S(x) and the pdf of the OCE distribution into equation 2.10 results in the hazard function. Therefore,

$$h(x)=\frac{\frac{\lambda\beta}{\alpha}e^{-x/\alpha}\left(e^{-x/\alpha}\right)^{-(\beta+1)}\left(1-e^{-x/\alpha}\right)^{(\beta-1)}\exp\left(\left(e^{x/\alpha}-1\right)^{\beta}\right)\exp\left(\lambda\left(1-\exp\left(\left(e^{x/\alpha}-1\right)^{\beta}\right)\right)\right)}{\exp\left(\lambda\left(1-\exp\left(\left(e^{x/\alpha}-1\right)^{\beta}\right)\right)\right)}$$

Simplifying further, the hazard function is given by;

$$h(x)=\frac{\lambda\beta}{\alpha}e^{-x/\alpha}\left(e^{-x/\alpha}\right)^{-(\beta+1)}\left(1-e^{-x/\alpha}\right)^{(\beta-1)}\exp\left(\left(e^{x/\alpha}-1\right)^{\beta}\right) \quad (2.11)$$

2.2 Mixture Representation

The mixture representation of the pdf is essential in the derivation of the statistical properties of the OCE family of distributions. An alternative form for the OCE distribution's pdf is provided in this section. Through the concept of power series, the density function of the OCE distribution is expressed in a mixture form to obtain certain statistical features of the distribution.

Lemma 2.1 The density function of the OCE distribution can be expressed in a series representation as;

$$f(x)=\frac{\lambda\beta}{\alpha}e^{\lambda}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\sum_{m=0}^{\infty}\omega_{ijkm}x^{m}e^{-kx/\alpha} \quad (2.12)$$

Where,

$$\omega_{ijkm}=\binom{\beta(j+1)-1}{k}\frac{(-1)^{i+k}\lambda^{i}}{i!}\cdot\frac{(i+1)^{j}}{r!}\cdot\frac{\beta^{m}(j+1)^{m}}{m!\alpha^{m}} \quad (2.13)$$

Proof. Given the density function,

$$f(x)=\frac{\lambda\beta}{\alpha}e^{-x/\alpha}\left(e^{-x/\alpha}\right)^{-(\beta+1)}\left(1-e^{-x/\alpha}\right)^{(\beta-1)}\exp\left(\left(e^{x/\alpha}-1\right)^{\beta}\right)\exp\left(\lambda\left(1-\exp\left(\left(e^{x/\alpha}-1\right)^{\beta}\right)\right)\right)\ \alpha>0,\beta>0,\lambda>0 \quad (2.14)$$

Applying the Taylor series expansion yields;

$$f(x)=\frac{\lambda\beta}{\alpha}e^{\lambda}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\frac{(-1)^{i}\lambda^{i}}{i!}\cdot\frac{(i+1)^{j}}{j!}\left(1-e^{-x/\alpha}\right)^{\beta(j+1)-1}e^{\beta(j+1)x/\alpha} \quad (2.15)$$

Further applying the Taylor's series and binomial expansion yields;

$$f(x)=\frac{\lambda\beta}{\alpha}e^{\lambda}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\sum_{m=0}^{\infty}\binom{\beta(j+1)-1}{k}\frac{(-1)^{i+k}\lambda^{i}}{i!}\cdot\frac{(i+1)^{j}}{r!}\cdot\frac{\beta^{m}(j+1)^{m}}{m!\,\alpha^{m}}x^{m}e^{-kx/\alpha}$$

This completes the proof.

2.3 Statistical Properties of the Odd Chen Exponential Distribution

In this section, the structural properties of the OCE distribution such as the quantile function, moments, moment-generating function, incomplete moment, entropy, order statistic, and inequality measures are derived and discussed.

2.3.1 Quantile Function

In probability and statistics, the quantile function, associated with a probability distribution of a random variable, specifies the value of the random variable such that the probability of the variable being less than or equal to that value equals the given probability. The quantile function is the distribution function's inverse. The quantile function can be used to calculate random variable characteristics such as skewness, kurtosis, and median [12].

Proposition 2.1 The quantile function of the OCE distribution is given by;

$$Q(p) = \alpha \log\left\{1 + \left[\log\left(1 - \tfrac{1}{\lambda}\log(1-p)\right)\right]^{\frac{1}{\beta}}\right\}, 0 < p < 1 \tag{2.16}$$

Proof. Suppose the random variable p follows the standard uniform distribution then, $0 < p < 1$.

$$Let\ p = 1 - \exp\left(\lambda - \lambda \exp\left(e^{x/\alpha} - 1\right)^{\beta}\right) \tag{2.17}$$

$$\exp\left(\lambda - \lambda \exp\left(e^{x/\alpha} - 1\right)^{\beta}\right) = 1 - p$$

$$\lambda\left[1 - \exp\left(e^{x/\alpha} - 1\right)^{\beta}\right] = \log(1-p)$$

$$\exp\left(e^{x/\alpha} - 1\right)^{\beta} = 1 - \frac{1}{\lambda}\log(1-p)$$

$$\left(e^{x/\alpha} - 1\right)^{\beta} = \log\left(1 - \frac{1}{\lambda}\log(1-p)\right)$$

$$x = \alpha \log\left\{1 + \left[\log\left(1 - \tfrac{1}{\lambda}\log(1-p)\right)\right]^{\frac{1}{\beta}}\right\} \tag{2.18}$$

By observations, x is a function of p, hence the quantile function is a function of p. Therefore,

$$Q(p) = \alpha \log\left\{1 + \left[\log\left(1 - \tfrac{1}{\lambda}\log(1-p)\right)\right]^{\frac{1}{\beta}}\right\} \tag{2.19}$$

This completes the proof.

Also, the measures of skewness and kurtosis can be computed based on the quantile measures. The Bowley measure of skewness and the Moors measure of kurtosis are respectively defined as;

$$Skewness = \frac{Q(1/4) + Q(3/4) - 2Q(1/2)}{Q(3/4) - Q(1/4)} \tag{2.20}$$

and

$$Kurtosis = \frac{Q(7/8) - Q(5/8) + Q(3/8) - Q(1/8)}{Q(3/4) - Q(1/4)} \tag{2.21}$$

2.3.2 Moment

Proposition 2.2 The r^{th} moment of the OCE distributed random variable X is given as;

$$u_r' = \frac{\lambda\beta}{\alpha} e^{\lambda} \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\sum_{m=0}^{\infty} \omega_{ijkm} \int_0^{\infty} x^{m+r} e^{-kx/\alpha} dx \tag{2.22}$$

Proof. By definition, the r^{th} non-central moment is given by;

$$u_r' = \int_0^{\infty} x^r f(x) dx \tag{2.23}$$

Applying the series representation of the pdf results in equation 2.24;

$$u_r' = \int_0^{\infty} \frac{\lambda\beta}{\alpha} e^{\lambda} \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\sum_{m=0}^{\infty} \omega_{ijkm} x^{m+r} e^{-kx/\alpha}\, dx \tag{2.24}$$

$$u_r' = \frac{\lambda\beta}{\alpha} e^{\lambda} \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\sum_{m=0}^{\infty} \omega_{ijkm} \int_0^{\infty} x^{m+r} e^{-kx/\alpha} dx \tag{2.25}$$

This therefore completes the proof.

2.3.3 Incomplete Moment

Proposition 2.3 The incomplete moment of the OCE distribution is given as;

$$M_r(x) = \int_0^x \frac{\lambda\beta}{\alpha} e^{\lambda} \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\sum_{m=0}^{\infty} \omega_{ijkm} y^{m+r} e^{-ky/\alpha}\, dy \quad (2.26)$$

Where, ω_{ijkm} is the same as in equation 2.13.

Proof. The incomplete moment of a random variable is given as;

$$M_r(x) = \int_0^x y^r f(y) dy \quad (2.27)$$

Thus, by replacing the series representation of the density function of the OCE distribution into the definition of the incomplete moment, the incomplete moment is given by;

$$M_r(x) = \int_0^x \frac{\lambda\beta}{\alpha} e^{\lambda} \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\sum_{m=0}^{\infty} \omega_{ijkm} y^{m+r} e^{-ky/\alpha}\, dy \quad (2.28)$$

This therefore completes the proof.

2.3.4 Inequality Measures

Several fields like insurance, econometrics and reliability studies employ Lorenz and Bonferroni curves in the study of inequality measures like income and poverty.

2.3.4.1 Lorenz curve

The Lorenz Curve is defined as;

$$L_f(x) = \frac{1}{\mu}\int_0^x y f(y) dy \quad (2.29)$$

The Lorenz curve for the OCE distribution is obtained by substituting the mixture representation of the density in equation 2.13 into the definition of the Lorenz curve. Therefore, the Lorenz Curve for the OCE distribution is given by;

$$L_f(x) = \frac{\lambda\beta}{\alpha\mu} e^{\lambda} \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\sum_{m=0}^{\infty} \omega_{ijkm} \int_0^x y^{m+1} e^{-ky/\alpha} dy \quad (2.30)$$

2.3.4.2 Bonferroni curve

The Bonferroni Curve is defined as;

$$B_F(x) = \frac{L_f(x)}{F(x)} \quad (2.31)$$

The Bonferroni curve for the OCE distribution is obtained by substituting the Lorenz curve in equation 2.31 into the definition of the Bonferroni curve. Therefore, the Bonferroni curve for the OCE distribution is given by;

$$B_F(x) = \frac{\lambda\beta}{\alpha\mu F(x)} e^{\lambda} \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\sum_{m=0}^{\infty} \omega_{ijkm} \int_0^x y^{m+1} e^{-ky/\alpha} dy \quad (2.32)$$

2.3.5 Entropy

Proposition 2.4 The Renyi entropy of the OCE distribution is given by;

$$I_R(\psi) = \frac{1}{(1-\psi)} \log\left\{\int_0^{\infty} \left(\frac{\lambda\beta e^{\lambda}}{\alpha}\right)^{\psi} \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\sum_{m=0}^{\infty} \phi_{ijkm} x^m e^{-kx/\alpha}\, dx\right\}, \psi \neq 1, \psi > 0 \quad (2.33)$$

Proof: The Renyi entropy, which is denoted by $I_R(\psi)$ for the random variable X is defined as;

$$I_R(\psi) = \frac{1}{(1-\psi)}\log\left\{\int_0^\infty f(x)^\psi dx\right\}, \psi > 0 \tag{2.34}$$

Using the density function, f(x) as it is in equation 2.6,

$$f(x)^\psi = \left[\frac{\lambda\beta}{\alpha}e^{-x/\alpha}\left(e^{-x/\alpha}\right)^{-(\beta+1)}\left(1-e^{-x/\alpha}\right)^{(\beta-1)}\exp\left(\left(e^{x/\alpha}-1\right)^\beta\right)\exp\left(\lambda\left(1-\exp\left(\left(e^{x/\alpha}-1\right)^\beta\right)\right)\right)\right]^\psi$$

By some algebraic simplification;

$$f(x)^\psi = \left(\frac{\lambda\beta}{\alpha}e^\lambda\right)^\psi \sum_{i=0}^\infty\sum_{j=0}^\infty\sum_{k=0}^\infty\sum_{m=0}^\infty \phi_{ijkm}x^m e^{-kx/\alpha}$$

where,

$$\phi_{ijkm} = \binom{\beta(j+1)-\psi}{k}\frac{(-1)^{i+k}\lambda^i\psi^i}{i!}.\frac{(i+\psi)^j}{r!}.\frac{\beta^m(j+\psi)^m}{m!\,\alpha^m}$$

Hence, the Renyi entropy is becomes;

$$I_R(\psi) = \frac{1}{(1-\psi)}\log\left\{\int_0^\infty\left(\frac{\lambda\beta e^\lambda}{\alpha}\right)^\psi \sum_{i=0}^\infty\sum_{j=0}^\infty\sum_{k=0}^\infty\sum_{m=0}^\infty \phi_{ijkm}x^m e^{-kx/\alpha}\,dx\right\}, \psi \neq 1, \psi > 0 \tag{2.35}$$

This therefore completes the proof.

2.4 The Moment Generating Function (MGF)

The Moment Generating Function (MGF) is a special function used to compute statistical measures such as the mean and variance of a given random variable.

Proposition 2.5

The MGF of a random variable X having the OCE distribution is given by;

$$M_x(t) = \frac{\lambda\beta}{\alpha}e^\lambda\sum_{i=0}^\infty\sum_{j=0}^\infty\sum_{k=0}^\infty\sum_{m=0}^\infty\sum_{r=0}^\infty\left(\omega_{ijkm}\frac{t^r}{r!}\int_0^\infty x^{m+r}e^{-kx/\alpha}dx\right) \tag{2.36}$$

Proof. The MGF is defined as;

$$M_x(t) = E(e^{tx}) = \int_0^\infty e^{tx}f(x)dx \tag{2.37}$$

Using the Taylor series expression, the MGF can be rewritten as;

$$M_x(t) = \int_0^\infty\sum_{r=0}^\infty\frac{t^r x^r}{r!}f(x)dx \tag{2.38}$$

$$M_x(t) = \sum_{r=0}^\infty\frac{t^r}{r!}\int_0^\infty x^r f(x)dx = \sum_{r=0}^\infty\frac{t^r}{r!}\mu_r' \tag{2.39}$$

Substituting the r^{th} moment of the OCE distribution yields;

$$M_x(t) = \frac{\lambda\beta}{\alpha}e^\lambda\sum_{i=0}^\infty\sum_{j=0}^\infty\sum_{k=0}^\infty\sum_{m=0}^\infty\sum_{r=0}^\infty\left(\omega_{ijkm}\frac{t^r}{r!}\int_0^\infty x^{m+r}e^{-kx/\alpha}dx\right) \tag{2.40}$$

This therefore completes the proof.

2.5 Characteristic Function

Characteristic functions are particularly useful in handling heavy-tailed random variables for which the corresponding moment-generating functions do not exist.

Proposition 2.6

The characteristic function of the OCE distribution for a random variable X is given by;

$$C_x(t) = \frac{\lambda\beta}{\alpha} e^{\lambda} \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \sum_{m=0}^{\infty} \sum_{r=0}^{\infty} \left(\omega_{ijkm} \frac{(it)^r}{r!} \int_0^{\infty} x^{m+r} e^{-kx/\alpha} dx \right) \tag{2.41}$$

Proof. The characteristic function of X is defined as;

$$C_x(t) = E(e^{itx}) = \int_0^{\infty} e^{itx} f(x) dx, \qquad where, i = \sqrt{-1}$$

Using Taylor's expansion, the characteristic function can be rewritten as;

$$C_x(t) = \int_0^{\infty} \sum_{r=0}^{\infty} \frac{(it)^r x^r}{r!} f(x) dx \tag{2.42}$$

$$C_x(t) = \sum_{r=0}^{\infty} \frac{(it)^r}{r!} \int_0^{\infty} x^r f(x) dx = \sum_{r=0}^{\infty} \frac{(it)^r}{r!} \mu_r' \tag{2.43}$$

Substituting the r^{th} moment of the OCE distribution yields;

$$C_x(t) = \frac{\lambda\beta}{\alpha} e^{\lambda} \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \sum_{m=0}^{\infty} \sum_{r=0}^{\infty} \left(\omega_{ijkm} \frac{(it)^r}{r!} \int_0^{\infty} x^{m+r} e^{-kx/\alpha} dx \right) \tag{2.44}$$

This completes the proof.

2.6 Order Statistics

The most common application of order statistics is in statistical theory. Order statistics is concerned with the properties and applications of these ordered random variables, as well as the functions that incorporate them. The order statistics for the OCE distribution is given by;

$$f_{p:n} = \lambda\beta \sum_{i=0}^{p-1} \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \sum_{m=0}^{\infty} D_{ijklm} \frac{n!}{(p-1)!(n-p)!} x^m e^{-lx/\alpha} \tag{2.45}$$

where,

$$D_{ijklm} = \binom{\beta(k+1)-1}{l} \binom{p-1}{i} \frac{(-1)^{i+j+1} [\lambda(n-p+i+1)]^j (j+1)^k \beta^m (k+1)^m}{j!k!m!\alpha^{m+1}} e^{\lambda(n-p+i+1)} \tag{2.46}$$

Proof. The pdf of the p^{th} order statistics is given by;

$$f_{p:n} = U_{r:n} [F(x)]^{p-1} [1 - F(x)]^{n-p} f(x), \quad r = 1,2,3, \dots, n \tag{2.47}$$

where,

$$U_{r:n} = \frac{n!}{(p-1)!(n-p)!} \tag{2.48}$$

By further simplification, equation (2.47) becomes;

$$f_{p:n} = U_{r:n} [F(x)]^{p-1} [S(x)]^{n-p} f(x) \tag{2.49}$$

$$f_{p:n} = U_{r:n} \left[1 - \exp\left(\lambda - \lambda \exp\left(e^{x/\alpha} - 1\right)^{\beta} \right) \right]^{p-1} \left[\exp\left(\lambda - \lambda \exp\left(e^{x/\alpha} - 1\right)^{\beta} \right) \right]^{n-p} f(x) \tag{2.50}$$

Applying the binomial expansion to the above yields;

$$f_{p:n} = U_{r:n} \sum_{i=0}^{p-1} \binom{p-1}{i} (-1)^i \left[\exp\left(\lambda - \lambda \exp\left(e^{x/\alpha} - 1\right)^{\beta}\right)\right]^{n-p+i} f(x) \quad (2.51)$$

$$f_{p:n} = U_{r:n} \sum_{i=0}^{p-1} \binom{p-1}{i} (-1)^i [\mathrm{S(x)}]^{n-p+i} f(x) \quad (2.52)$$

$$[\mathrm{S(x)}]^{n-p+i} f(x) = \left[\exp\left(\lambda - \lambda \exp\left(e^{x/\alpha} - 1\right)^{\beta}\right)\right]^{n-p+i} \left[\frac{\lambda\beta}{\alpha} e^{-x/\alpha}\left(e^{-x/\alpha}\right)^{-(\beta+1)}\left(1 - e^{-x/\alpha}\right)^{(\beta-1)} \phi_{(x,\alpha)}\right] \quad (2.53)$$

Applying Taylor's and binomial expansions on equation 2.53, the resulting equation is given by;

$$[\mathrm{S(x)}]^{n-p+i} f(x) = \lambda\beta \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \sum_{m=0}^{\infty} \binom{\beta(k+1)-1}{l} \frac{(-1)^{j+l}[\lambda(n-p+i+1)]^j (j+1)^k \beta^m (k+1)^m}{e^{-\lambda(n-p+i+1)} j!k!m!\alpha^{m+1}} x^m e^{-lx/\alpha} \quad (2.54)$$

Substituting equation (2.54) into equation (2.52), equation (2.55) is obtained as;

$$f_{p:n} = \lambda\beta \sum_{i=0}^{p-1} \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \sum_{m=0}^{\infty} \binom{\beta(k+1)-1}{l} \binom{p-1}{i} \frac{n!(-1)^{i+j+l}[\lambda(n-p+i+1)]^j (j+1)^k \beta^m (k+1)^m}{e^{-\lambda(n-p+i+1)}(p-1)!(n-p)!j!k!m!\alpha^{m+1}} x^m e^{-lx/\alpha}$$
(2.55)

Hence the proof.

2.7 Mean Residual Life

The expected remaining life, $X - x$, given that the item has survived to time X, is the mean residual life (MRL). Thus, in life-testing situations, the MRL is the expected additional lifetime, given that a component has survived until time X. Since the MRL function represents the expected remaining life, x must be subtracted, resulting in;

$$h(x) = E[X - x | X > x] = \frac{\int_x^{\infty} y f(y) dy}{S(x)} - x \quad (2.56)$$

$$h(x) = \lambda\beta e^{\lambda} \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \sum_{m=0}^{\infty} \frac{\omega_{ijkm}}{exp\left[\lambda\left(1 - exp\left(e^{x/\alpha} - 1\right)^{\beta}\right)\right]} \int_x^{\infty} y^{m+1} e^{-\frac{ky}{\alpha}} dy - x \quad (2.57)$$

2.8 Method of Estimations

In this section, some estimations for the OCE are computed. The maximum likelihood estimators, least square estimators, percentile estimators, maximum product of spacings estimators and the minimum distances estimators are discussed.

2.8.1 Maximum Likelihood Estimators

$$L(x; \alpha, \beta, \lambda) = \prod_{i=1}^{n} f(x) \quad (2.58)$$

$$L(x; \alpha, \beta, \lambda) = \left(\frac{\lambda\beta}{\alpha}\right)^n e^{\sum_{i=1}^{\infty} \frac{\beta x_i}{\alpha}} \sum_{i=1}^{n} \left(1 - e^{-\frac{x_i}{\alpha}}\right)^{\beta-1} e^{\sum_{i=1}^{n}\left(e^{\frac{x_i}{\alpha}} - 1\right)^{\beta}} exp\left(\lambda \sum_{i=1}^{n}\left(1 - exp\left(e^{\frac{x_i}{\alpha}} - 1\right)^{\beta}\right)\right) \quad (2.59)$$

$$l = log\ \ L(x; \alpha, \beta, \lambda) = n\, log\left(\frac{\lambda\beta}{\alpha}\right) + \sum_{i=1}^{\infty} \frac{\beta x_i}{\alpha} + (\beta - 1) \sum_{i=1}^{n} log\left(1 - e^{-\frac{x_i}{\alpha}}\right) + \lambda \sum_{i=1}^{n}\left(1 - exp\left(e^{\frac{x_i}{\alpha}} - 1\right)^{\beta}\right)$$
(2.60)

For ease of notation, the first partial derivative of equation (2.60) with respect to α, β, λ is denoted by $ml_\alpha, ml_\beta, ml_\lambda$ respectively. Now, equating the first partial derivatives to zero results in equation (61);

$$ml_\alpha = -\frac{n}{\alpha} - \sum_{i=1}^{n} \frac{\beta x_i}{\alpha^2} - (\beta - 1) \sum_{i=1}^{n} \frac{x e^{-\frac{x_i}{\alpha}}}{\alpha^2\left(1 - e^{-\frac{x_i}{\alpha}}\right)} + \lambda \sum_{i=1}^{n} \left(\frac{\left(e^{\frac{x_i}{\alpha}} - 1\right)^{\beta} \beta x e^{\frac{x_i}{\alpha}} exp\left(\left(e^{\frac{x_i}{\alpha}} - 1\right)^{\beta}\right)}{\alpha^2\left(e^{\frac{x_i}{\alpha}} - 1\right)} \right) = 0 \quad (2.61)$$

$$ml_\beta = \frac{n}{\beta} + \sum_{i=1}^{n} log\left(1 - e^{-\frac{x_i}{\alpha}}\right) - \lambda \sum_{i=1}^{n} \left(\left(e^{\frac{x_i}{\alpha}} - 1\right)^{\beta}\right) exp\left(\left(e^{\frac{x_i}{\alpha}} - 1\right)^{\beta}\right) ln\left(e^{\frac{x_i}{\alpha}} - 1\right) = 0 \quad (2.62)$$

$$ml_\lambda = \frac{n}{\lambda} + \sum_{i=1}^{n} \left(1 - exp\left(e^{\frac{x_i}{\alpha}} - 1\right)^{\beta}\right) = 0 \quad (2.63)$$

The maximum likelihood estimates $\hat{\alpha}, \hat{\beta}$ and $\hat{\lambda}$ of α, β and λcan be obtained by solving equations 2.61 - 2.63.

2.8.2 Least Squares Estimators

The least-square estimators and weighted least-square estimators were proposed by [13] to estimate the parameters of Beta distributions. Suppose $F(x_i)$ denotes the OCE distribution function of the ordered random variables $X_1 < X_2 < \ldots < X_n$ where $\{X_1, X_2, \ldots, X_n\}$ is a random sample of size n from a distribution function $F(x; \alpha, \beta, \lambda)$. The least-square estimators $\hat{\alpha}_{lse}, \hat{\beta}_{lse}$ and $\hat{\lambda}_{lse}$of α, β and λ can be obtained by minimizing

$$LSE = \sum_{i=1}^{n} \left[F(x_{i:1}|\alpha, \beta, \lambda) - \frac{i}{n+1}\right]^2 \quad (2.64)$$

with respect to α, β and $\lambda$and equating them to zero where $F(x_{i:1}|\alpha, \beta, \lambda)$ is the cdf of the OCE. Consequently, they can be obtained by solving the following equations,

$$\frac{\partial LSE}{\partial \alpha} = -2 \sum_{i=1}^{n} \left[1 - exp\left(\lambda\left(1 - exp\left(\left(e^{\frac{x_{i:n}}{\alpha}} - 1\right)^{\beta}\right)\right)\right) - \frac{i}{n+1}\right] \mu_1(x_{i:n}|\alpha, \beta, \lambda) = 0$$

$$\frac{\partial LSE}{\partial \beta} = 2 \sum_{i=1}^{n} \left[1 - exp\left(\lambda\left(1 - exp\left(\left(e^{\frac{x_{i:n}}{\alpha}} - 1\right)^{\beta}\right)\right)\right) - \frac{i}{n+1}\right] \mu_2(x_{i:n}|\alpha, \beta, \lambda) = 0$$

$$\frac{\partial LSE}{\partial \lambda} = -2 \sum_{i=1}^{n} \left[1 - exp\left(\lambda\left(1 - exp\left(\left(e^{\frac{x_{i:n}}{\alpha}} - 1\right)^{\beta}\right)\right)\right) - \frac{i}{n+1}\right] \mu_3(x_{i:n}|\alpha, \beta, \lambda) = 0$$

where,

$$\mu_1(x_{i:n}|\alpha, \beta, \lambda) = \frac{\lambda \beta x \left(e^{\frac{x_{i:n}}{\alpha}} - 1\right)^{\beta} exp\left(\left(e^{\frac{x_{i:n}}{\alpha}} - 1\right)^{\beta}\right) exp\left(\lambda\left(1 - exp\left(\left(e^{\frac{x_{i:n}}{\alpha}} - 1\right)^{\beta}\right)\right)\right)}{\alpha^2\left(1 - e^{-\frac{x_{i:n}}{\alpha}}\right)} \quad (2.65)$$

$$\mu_2(x_{i:n}|\alpha, \beta, \lambda) = \lambda \left(e^{\frac{x_{i:n}}{\alpha}} - 1\right)^{\beta} ln\left(e^{\frac{x_{i:n}}{\alpha}} - 1\right) exp\left(\left(e^{\frac{x_{i:n}}{\alpha}} - 1\right)^{\beta}\right) exp\left(\lambda\left(1 - exp\left(\left(e^{\frac{x_{i:n}}{\alpha}} - 1\right)^{\beta}\right)\right)\right) \quad (2.66)$$

$$\mu_3(x_{i:n}|\alpha, \beta, \lambda) = \left(1 - exp\left(\left(e^{\frac{x_{i:n}}{\alpha}} - 1\right)^{\beta}\right)\right) exp\left(\lambda\left(1 - exp\left(\left(e^{\frac{x_{i:n}}{\alpha}} - 1\right)^{\beta}\right)\right)\right) \quad (2.67)$$

The weighted least squares estimators, on the other hand, can be obtained by minimizing the equation;

$$WLSE = \sum_{i=1}^{n} \frac{(n+1)^2(n+2)}{i(n-i+1)} \left[F(x_{i:1}|\alpha,\beta,\lambda) - \frac{i}{n+1}\right]^2 \tag{2.68}$$

with respect to α, β and λ. The estimators can be obtained by solving the system of equations in equation 2.69;

$$\frac{\partial LSE}{\partial \alpha} = -2\sum_{i=1}^{n} \frac{(n+1)^2(n+2)}{i(n-i+1)} \left[1 - exp\left(\lambda\left(1 - exp\left(\left(e^{\frac{x_{i:n}}{\alpha}} - 1\right)^{\beta}\right)\right)\right) - \frac{i}{n+1}\right] \mu_1(x_{i:n}|\alpha,\beta,\lambda) = 0$$

$$\frac{\partial LSE}{\partial \beta} = 2\sum_{i=1}^{n} \frac{(n+1)^2(n+2)}{i(n-i+1)} \left[1 - exp\left(\lambda\left(1 - exp\left(\left(e^{\frac{x_{i:n}}{\alpha}} - 1\right)^{\beta}\right)\right)\right) - \frac{i}{n+1}\right] \mu_2(x_{i:n}|\alpha,\beta,\lambda) = 0$$

$$\frac{\partial LSE}{\partial \lambda} = -2\sum_{i=1}^{n} \frac{(n+1)^2(n+2)}{i(n-i+1)} \left[1 - exp\left(\lambda\left(1 - exp\left(\left(e^{\frac{x_{i:n}}{\alpha}} - 1\right)^{\beta}\right)\right)\right) - \frac{i}{n+1}\right] \mu_3(x_{i:n}|\alpha,\beta,\lambda) = 0$$

(2.69)

2.8.3 Percentile Estimators

If the data come from a closed-form distribution function, the unknown parameters can be estimated by fitting a straight line to the theoretical points from the distribution function and the sample percentile points [14] proposed this method, which has been used for Weibull distribution and generalized exponential distribution. Let j be the jth order statistics, that is, $X_1 < X_2 < \ldots < X_n$. If p_j denotes some estimates of $F(x_j; \alpha, \beta, \lambda)$ the estimates, $\hat{\alpha}, \hat{\beta}$ and $\hat{\lambda}$ can be obtained by minimizing with respect to α, β and λ and equating them to zero in the equation below;

$$PE = \sum_{j=1}^{n} \left(x_j - \alpha\, log\left[1 + \left(log\left(1 - \frac{1}{\lambda} log(1-p)\right)\right)^{\frac{1}{\beta}}\right]\right)^2 \tag{2.70}$$

The estimators can be obtained by solving the following system in equation (2.71);

$$\frac{\partial PE}{\partial \alpha} = -2\sum_{j=1}^{n} \left(x_j - \alpha\, log\left[1 + \left(log\left(1 - \frac{1}{\lambda} log(1-p)\right)\right)^{\frac{1}{\beta}}\right]\right)^2 \Omega_1(p;\alpha,\beta,\lambda) = 0,$$

$$\frac{\partial PE}{\partial \beta} = 2\sum_{j=1}^{n} \left(x_j - \alpha\, log\left[1 + \left(log\left(1 - \frac{1}{\lambda} log(1-p)\right)\right)^{\frac{1}{\beta}}\right]\right)^2 \Omega_2(p;\alpha,\beta,\lambda) = 0 \tag{2.71}$$

$$\frac{\partial PE}{\partial \lambda} = -2\sum_{j=1}^{n} \left(x_j - \alpha\, log\left[1 + \left(log\left(1 - \frac{1}{\lambda} log(1-p)\right)\right)^{\frac{1}{\beta}}\right]\right)^2 \Omega_3(p;\alpha,\beta,\lambda) = 0,$$

Where

$$\Omega_1(p;\alpha,\beta,\lambda) = ln\left(1 + ln\left(1 - \frac{ln(1-p)}{\lambda}\right)^{\frac{1}{\beta}}\right) \tag{2.72}$$

$$\Omega_2(p;\alpha,\beta,\lambda) = \frac{\alpha\, ln\left(1 - \frac{ln(1-p)}{\lambda}\right)^{\frac{1}{\beta}} ln\left(ln\left(1 - \frac{ln(1-p)}{\lambda}\right)\right)}{\beta^2\left(1 + ln\left(1 - \frac{ln(1-p)}{\lambda}\right)^{\frac{1}{\beta}}\right)} \tag{2.73}$$

$$\Omega_3(p;\alpha,\beta,\lambda) = \frac{\alpha\, ln\left(1-\frac{ln(1-p)}{\lambda}\right)^{\frac{1}{\beta}} ln(1-p)}{\beta\lambda^2\left(1-\frac{ln(1-p)}{\lambda}\right) ln\left(1-\frac{ln(1-p)}{\lambda}\right)\left(1+ln\left(1-\frac{ln(1-p)}{\lambda}\right)^{\frac{1}{\beta}}\right)} \quad (2.74)$$

These estimators are referred to as percentile estimators.

2.9 Maximum Product of Spacings Estimators

Ranneby [15] proposed the maximum product spacing (MPS) method as an alternative to MLE for estimating the unknown parameters of continuous univariate distributions [8] and independently developed the MPS method as an approximation to the Kullback-Leibler measure of information. Cheng [16] demonstrated that this method is as efficient as MLE estimators and consistent under more general conditions, which motivated our choice. Define the uniform spacings of a random sample from the OCE distribution as follows, using the same notation as in the percentile estimator:

$$D_i(\alpha,\beta,\lambda) = F(x_{i:n}|\alpha,\beta,\lambda) - F(x_{i-1:n}|\alpha,\beta,\lambda),\ i = 1,2,\dots,n,$$

where $F(x_{0:n}|\alpha,\beta,\lambda) = 0$ and $F(x_{n+1:n}|\alpha,\beta,\lambda) = 1$. Clearly, $\sum_{i=1}^{n+1} D_i(\alpha,\beta,\lambda) = 1$.

The maximum product of spacings estimator $\hat{\alpha}_{MPS}, \hat{\beta}_{MPS}$ and $\hat{\lambda}_{MPS}$ of the parameters α,β and λ are obtained by maximizing with respect to α,β and λ, the geometric mean of spacings:

$$G(\alpha,\beta,\lambda) = [\prod_{i=1}^{n+1} D_i(\alpha,\beta,\lambda)]^{\frac{1}{n+1}} \quad (2.75)$$

or, equivalently, by minimizing the function;

$$H(\alpha,\beta,\lambda) = \frac{1}{n+1}\sum_{i=1}^{n+1} log\, D_i\,(\alpha,\beta,\lambda) \quad (2.76)$$

The estimators can be obtained by solving the non-linear system in equation 2.77;

$$\frac{\partial}{\partial\alpha}H(\alpha,\beta,\lambda) = \frac{1}{n+1}\sum_{i=1}^{n+1}\frac{1}{D_i(\alpha,\beta,\lambda)}[\mu_1(x_{i:n}|\alpha,\beta,\lambda) - \mu_1(x_{i-1:n}|\alpha,\beta,\lambda)] = 0$$

$$\frac{\partial}{\partial\beta}H(\alpha,\beta,\lambda) = \frac{1}{n+1}\sum_{i=1}^{n+1}\frac{1}{D_i(\alpha,\beta,\lambda)}[\mu_2(x_{i:n}|\alpha,\beta,\lambda) - \mu_2(x_{i-1:n}|\alpha,\beta,\lambda)] = 0 \quad (2.77)$$

$$\frac{\partial}{\partial\beta}H(\alpha,\beta,\lambda) = \frac{1}{n+1}\sum_{i=1}^{n+1}\frac{1}{D_i(\alpha,\beta,\lambda)}[\mu_3(x_{i:n}|\alpha,\beta,\lambda) - \mu_3(x_{i-1:n}|\alpha,\beta,\lambda)] = 0$$

2.10 Minimum Distances Estimators

Three minimization-based estimation methods for α,β and λ with respect to α,β and λ of the goodness-of-fit statistics are presented. This statistical class is based on the difference between the cumulative distribution function estimate and the empirical distribution function [10].

2.10.1 Cramer-von-mises

Luciano [17] provided empirical evidence that the bias of the estimator is smaller than that of the other minimum distance estimators, which motivated our choice of Cramér-von-Mises type minimum distance estimators. Thus, the Cramer-Von-Mises estimators, $\hat{\alpha}_{CVME}, \hat{\beta}_{CVME}$ and $\hat{\lambda}_{CVME}$ of the parameters α,β and λ can be obtained by minimizing with respect to α,β and λ, the function;

$$C(\alpha,\beta,\lambda) = \frac{1}{12n} + \sum_{i=1}^{n}\left[F(x_{i:1}|\alpha,\beta,\lambda) - \frac{2i-1}{2n}\right]^2 \quad (2.78)$$

These estimators can also be obtained by solving the system of nonlinear equations;

$$\frac{\partial}{\partial\alpha}C(\alpha,\beta,\lambda) = -2\sum_{i=1}^{n}\left[F(x_{i:1}|\alpha,\beta,\lambda) - \frac{2i-1}{2n}\right]^2 \mu_1(x_{i:n}|\alpha,\beta,\lambda) = 0$$

$$\frac{\partial}{\partial\beta}C(\alpha,\beta,\lambda) = 2\sum_{i=1}^{n}\left[F(x_{i:1}|\alpha,\beta,\lambda) - \frac{2i-1}{2n}\right]^2 \mu_2(x_{i:n}|\alpha,\beta,\lambda) = 0 \quad (2.79)$$

$$\frac{\partial}{\partial\lambda}C(\alpha,\beta,\lambda) = -2\sum_{i=1}^{n}\left[F(x_{i:1}|\alpha,\beta,\lambda) - \frac{2i-1}{2n}\right]^2 \mu_3(x_{i:n}|\alpha,\beta,\lambda) = 0$$

where, $\mu_1(x_{i:n}|\alpha,\beta,\lambda)$, $\mu_2(x_{i:n}|\alpha,\beta,\lambda)$ and $\mu_3(x_{i:n}|\alpha,\beta,\lambda)$ are the same as have been through this study.

2.10.2 Anderson-Darling and Right-Tail Anderson-Darling estimators

The Anderson-Darling test [18] is an alternative to other statistical tests for detecting deviations from normality in sample distributions. The AD test, in particular, converges very quickly toward the asymptote [18-20]. The Anderson-Darling Estimators, $\hat{\alpha}_{ADE}, \hat{\beta}_{ADE}$ and $\hat{\lambda}_{ADE}$ of the parameters α, β and λ can be obtained from minimizing with respect to α, β and λ the function;

$$A(\alpha,\beta,\lambda) = -n - \frac{1}{n}\sum_{i=1}^{n}(2i-1)[log\,F\,(x_{i:n}|\alpha,\beta,\lambda) + log\,S\,(x_{n+1-i:n}|\alpha,\beta,\lambda)] \quad (2.80)$$

These estimators can also be obtained by solving the nonlinear equations;

$$\sum_{i=1}^{n}(2i-1)\left[\frac{\mu_1(x_{i:n}|\alpha,\beta,\lambda)}{F(x_{i:n}|\alpha,\beta,\lambda)} + \frac{\mu_1(x_{n+1-i:n}|\alpha,\beta,\lambda)}{S(x_{n+1-i:n}|\alpha,\beta,\lambda)}\right] = 0$$

$$\sum_{i=1}^{n}(2i-1)\left[\frac{\mu_2(x_{i:n}|\alpha,\beta,\lambda)}{F(x_{i:n}|\alpha,\beta,\lambda)} + \frac{\mu_2(x_{n+1-i:n}|\alpha,\beta,\lambda)}{S(x_{n+1-i:n}|\alpha,\beta,\lambda)}\right] = 0, \quad (2.81)$$

$$\sum_{i=1}^{n}(2i-1)\left[\frac{\mu_3(x_{i:n}|\alpha,\beta,\lambda)}{F(x_{i:n}|\alpha,\beta,\lambda)} + \frac{\mu_3(x_{n+1-i:n}|\alpha,\beta,\lambda)}{S(x_{n+1-i:n}|\alpha,\beta,\lambda)}\right] = 0$$

The right-tail Anderson-Darling estimators $\hat{\alpha}_{RTADE}, \hat{\beta}_{RTADE}$ and $\hat{\lambda}_{RTADE}$ of the parameters α, β and λ are obtained by minimizing, with respect to α, β and λ, the function;

$$R(\alpha,\beta,\lambda) = \frac{n}{2} - 2\sum_{i=1}^{n}F(x_{i:n}|\alpha,\beta,\lambda) - \frac{1}{n}\sum_{i=1}^{n}(2i-1)\,log\,S\,(x_{n+1-i:n}|\alpha,\beta,\lambda) \quad (2.82)$$

These estimators can also be obtained by solving the nonlinear equations:

$$\frac{\partial}{\partial\alpha}R(\alpha,\beta,\lambda) = 2\sum_{i=1}^{n}\mu_1(x_{i:n}|\alpha,\beta,\lambda) - \frac{1}{n}\sum_{i=1}^{n}(2i-1)\,log\,\frac{\mu_1(x_{n+1-i:n}|\alpha,\beta,\lambda)}{S(x_{n+1-i:n}|\alpha,\beta,\lambda)} = 0,$$

$$\frac{\partial}{\partial\beta}R(\alpha,\beta,\lambda) = -2\sum_{i=1}^{n}\mu_2(x_{i:n}|\alpha,\beta,\lambda) - \frac{1}{n}\sum_{i=1}^{n}(2i-1)\,log\,\frac{\mu_2(x_{n+1-i:n}|\alpha,\beta,\lambda)}{S(x_{n+1-i:n}|\alpha,\beta,\lambda)} = 0, \quad (2.83)$$

$$\frac{\partial}{\partial\lambda}R(\alpha,\beta,\lambda) = 2\sum_{i=1}^{n}\mu_3(x_{i:n}|\alpha,\beta,\lambda) - \frac{1}{n}\sum_{i=1}^{n}(2i-1)\,log\,\frac{\mu_3(x_{n+1-i:n}|\alpha,\beta,\lambda)}{S(x_{n+1-i:n}|\alpha,\beta,\lambda)} = 0,$$

3 Shapes of CDF, Density and Hazard Function

Fig. 1 shows the cdf of the OCE distribution for different parameter values. It can be observed that as *x* gets closer to zero the CDF approaches zero and as *x* gets bigger the CDF approaches one. This satisfies the condition of a CDF, which states that the limit of a CDF as x approaches zero must be zero and the limit of CDF as x approaches infinity must be 1.

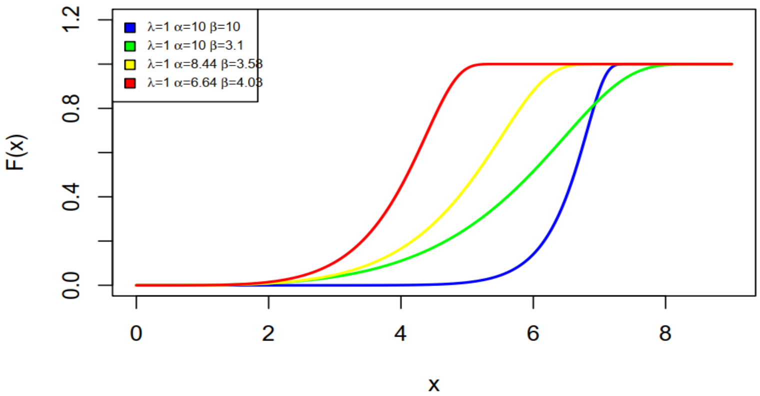

Fig. 1. CDF of the OCE Distribution

Fig. 2 also shows the density plots of the OCG distribution. The density plots for various parameter values exhibit different kinds of shapes; most of the plots are uni-modal in shape with different degrees of kurtosis. For the various parameter values that were used, it can be observed that some of the plots exhibit a right-skewed shape, whereas a significant number also showed a left-skewed shape. In addition, for some parameter values, the density plot exhibits a higher altitude.

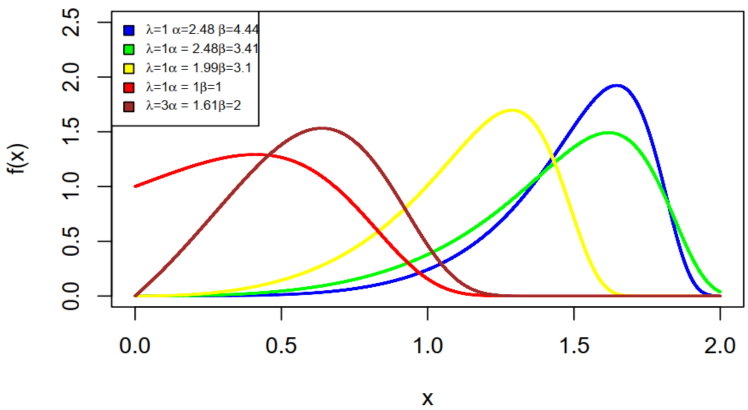

Fig. 2. PDF of the OCE distribution

Fig. 3 depicts the various shapes of the survival function for the OCE distribution for various parameter values. It shows that the survival function decays very quickly as it approaches one for some parameter values, indicating low reliability. Furthermore, the survival function decays slowly as it approaches one for some parameter values, indicating high reliability.

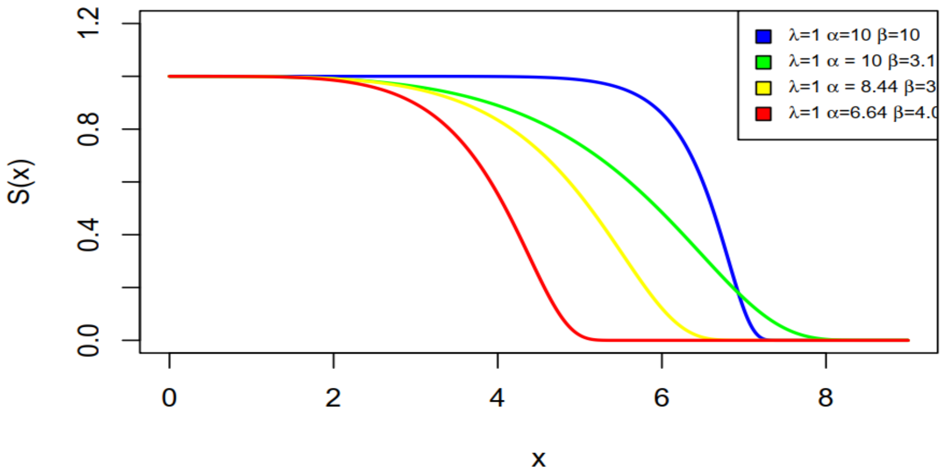

Fig. 3. Survival function plots for the OCE distribution

Fig. 4 shows the plots of the hazard function of the OCE distribution for different parameter values. It can be observed that the hazard function can be a constant and can also increase exponentially. It can also be a bathtub and also decrease exponentially. Furthermore, for a few selected parameter values, the hazard plot shows an upside-down bathtub shape, which is an indication that the OCE distribution is appropriate for modeling lifetime data.

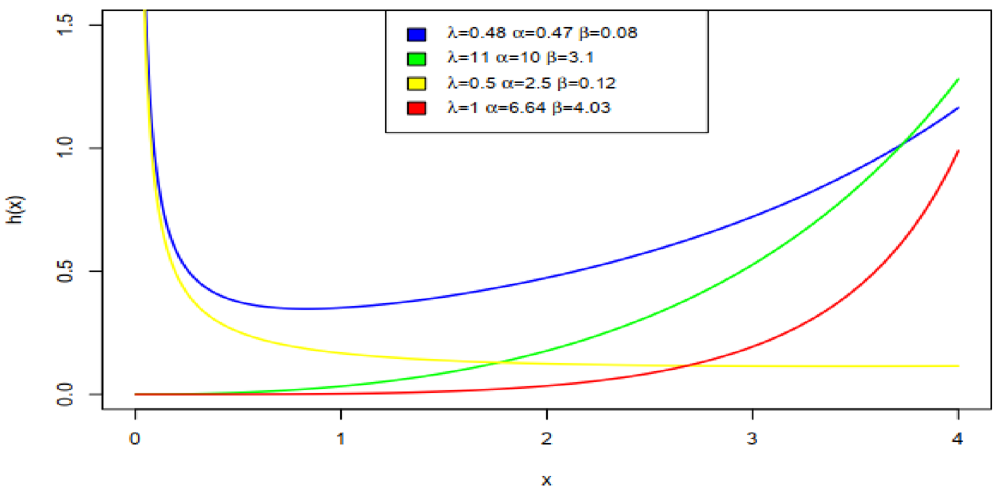

Fig. 4. Hazard function diagram for OCG distribution

3.1 Simulation

The various estimators are compared using Monte-Carlo's simulations by analyzing the behaviour of the OCE estimators. To verify these estimators, the Root Mean Square Error (RMSE) and the Average Bias (AB) are computed. The sample sizes of 50, 100, 200 and 400 were considered in the study and the R software was used to estimate the proposed model's parameters. The experiment is replicated 10,000 times to compare the AB and RMSE of the three estimation techniques suggested in this study. The AB and RMSE are computed as:

$$AB = \frac{1}{N}\sum_{i=0}^{N}\left(\hat{\theta}_i - \theta\right) \tag{3.1}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=0}^{N}\left(\hat{\theta}_i - \theta\right)^2} \tag{3.2}$$

where $\hat{\theta}$ is the estimation parameters of the OCE distribution that is θ =(α, β, λ). The simulation results for θ =(3.5, 0.5, 0.01) and θ =(4.5, 1.5, 0.01) are presented in the Tables 1 and 2 respectively. The best estimation method depends on an estimator having a minimum estimate and a decreasing Abs.

Table 1 displays the AB and RMSEs of the MLE, WLSE and ADE for the OCE distribution for the increasing dataset size, n and fixed parameter values for $\alpha, \beta$and λ. It shows that both the AB and RMSE decrease as the size of the dataset increases. It features the rank which is computed by summing the superscript of the AB and RMSE values for the three estimators at each dataset size. The superscript is assigned from the smallest to the largest on each row with the smallest having the superscript (1) and the largest having the superscript (3). It also reveals that the MLE is the best estimator for the OCE distribution.

Table 1. Simulations results for α=3.5, β=0.5, λ=0.01

n	Estimate	Parameters	MLE	WLSE	ADE
50	AB	α	20.5958^{1}	24.7002^{2}	65.2848^{3}
		β	10.5919^{2}	4.9797^{1}	18.5748^{3}
		λ	2.1450^{2}	3.8740^{3}	1.6052^{1}
	RMSE	α	12.0804^{1}	20.7419^{3}	19.9743^{2}
		β	6.6878^{3}	3.9174^{2}	3.6718^{1}
		λ	1.2502^{1}	2.1701^{3}	1.7933^{2}
	$\sum$ Rank		10^{1}	14^{3}	12^{2}
100	AB	α	4.0177^{1}	4.0457^{2}	17.9100^{3}
		β	0.9998^{1}	1.0418^{2}	4.4146^{3}
		λ	0.3733^{2}	0.5627^{3}	0.3493^{1}
	RMSE	α	8.0243^{1}	11.8207^{2}	11.9647^{3}
		β	3.6070^{3}	2.1780^{2}	2.1383^{1}
		λ	0.9621^{3}	0.5098^{2}	0.4146^{1}
	$\sum$ Rank		11^{3}	13^{3}	12^{2}
200	AB	α	3.3355^{2}	1.5215^{1}	10.1920^{3}
		β	0.2535^{1}	0.3461^{2}	1.4996^{3}
		λ	0.0519^{1}	0.1041^{3}	0.0546^{2}
	RMSE	α	5.0842^{1}	6.8689^{3}	5.9682^{2}
		β	0.9828^{2}	1.1996^{3}	0.9425^{1}
		λ	0.1101^{2}	0.3171^{3}	0.0861^{1}
	$\sum$ Rank		9^{1}	15^{3}	12^{2}
400	AB	α	3.5095^{1}	4.0043^{3}	3.6503^{2}
		β	0.5157^{1}	0.6509^{3}	0.5660^{2}
		λ	0.0479^{2}	0.0356^{1}	0.0628^{3}
	RMSE	α	3.5095^{1}	4.0043^{3}	3.6503^{2}
		β	0.5157^{1}	0.6509^{3}	0.5660^{2}
		λ	0.0479^{2}	0.0356^{1}	0.0628^{3}
	$\sum$ Rank		9^{1}	13^{2}	14^{3}

Table 2 shows the AB and RMSE values for the three estimators. It can be observed that the MLE ranks the smallest making it the best estimator amongst the three estimators for the OCE distribution.

Table 2. Simulations results for α=4.5, β=1.5, λ=0.01

n	Est	Parameters	MLE	WLSE	ADE
50	AB	α	10.3145^2	3.5823^1	38.7946^3
		β	11.3173^2	7.3072^1	16.7089^3
		λ	1.6453^1	1.6761^2	1.7959^3
	RMSE	α	9.2441^1	9.6783^3	9.4266^2
		β	8.0022^3	5.3529^2	4.3517^1
		λ	0.6762^2	0.7028^3	0.4122^1
	$\sum$ Rank		11^1	12^2	13^3
100	AB	α	2.6851^1	3.0319^2	17.8483^3
		β	4.0150^1	4.1666^2	7.1050^3
		λ	0.5524^3	0.3554^2	0.2065^1
	RMSE	α	6.0794^1	7.0082^3	6.6568^2
		β	3.5515^2	3.7384^3	2.4931^1
		λ	0.3448^1	0.4277^3	0.3523^2
	$\sum$ Rank		9^1	15^3	12^2
200	AB	α	1.5291^1	1.5650^2	8.5623^3
		β	2.2142^2	1.7979^1	3.2025^3
		λ	0.1261^2	0.2160^3	0.0463^1
	RMSE	α	4.1772^1	4.1968^2	4.5482^3
		β	2.8713^3	2.3054^2	1.5478^1
		λ	0.1195^2	0.2278^3	0.0354^1
	$\sum$ Rank		11^1	13^3	12^2
400	AB	α	0.4071^1	0.8911^2	4.0691^3
		β	0.6547^1	0.8004^2	1.5173^3
		λ	0.0594^3	0.0447^2	0.0151^1
	RMSE	α	2.6971^2	2.5477^1	2.7622^3
		β	1.0476^2	1.3615^3	1.0179^1
		λ	0.0625^2	0.1026^3	0.0104^1
	$\sum$ Rank		11^1	13^3	12^2

A closer look at Tables 1 and 2 indicates that the MLE performs better than the other estimators. The RMSE appears to decrease with increasing number of data points. This is the same for the average bias (AB) for the three estimators and thus the MLE can be used in the application of OCE distribution.

3.2 Application

The OCE was compared to the Odd Chen Weibull (OCW), Odd Chen Rayleigh (OCR), Rayleigh, Cauchy, Generalized Inverse Weibull and the Modified Extended Chen by applying them to real-life data and observing their AIC and the p-values for their parameter.

Dataset I: This is the unscheduled downtime for milling processing at a mining company. This dataset is in hours and contains 29 data points.

7.6, 12.1, 20.7, 23.8, 53.7, 22.4, 2.91, 54.8, 27.1, 28, 30, 12.2, 28.7, 21.3, 7.1, 7.7, 8.2, 5.8, 14.9, 19.2, 36.3, 74.5, 22.4, 35.5, 13.3, 11.4, 9.1, 29.1, 26.9

Dataset II: Survival times of patients with acute myelogenous leukaemia. This dataset was picked from [1]. It contains 33 data points

65, 156, 100, 134, 16, 108, 121, 4, 39, 143, 56, 26, 22, 1, 1, 5, 65, 56, 65, 17, 7, 16, 22, 3, 4, 2, 3, 8, 4, 3, 30, 4, 43

Table 3. Descriptive statistics of the two above-mentioned datasets

Data	n	min.	Q_1	median	mean	Q_3	max	Skew.	Kurto.	Var	Sd
I	29	2.91	11.40	21.30	22.99	28.70	74.50	1.31	1.60	268	16.4
II	33	1	4.00	22.00	40.88	65.00	156.00	1.11	-0.06	2181	46.7

From Tables 4 and 5, it can be observed that the OCE distribution has the highest log likelihood of -116.3209 and -153.3405. It also has the smallest AIC value of 238.6419 in Table 4 and 312.6810 in Table 5 as compared to the other fitted distributions. The p-values of the OCE in most cases are lower than the other distributions, proving that the OCE can perform better than the other distributions.

Table 4. Comparison criteria for data set I

Distribution	-L	AIC	CVM (P-value)	AD (P-value)	KS(P-value)
OCE	**-116.3209**	**238.6419**	**0.0499 (0.8806)**	**0.3586 (0.8877)**	**0.1109 (0.8683)**
OCW	-116.5975	241.1951	0.0579 (0.8310)	0.4151 (0.8325)	0.1220 (0.7809)
OCR	-116.3662	238.7323	0.0513 (0.8719)	0.3691 (0.8778)	0.1132 (0.8517)
Rayleigh	-118.4599	238.9199	0.2321 (0.2133)	1.5678 (0.1613)	0.1781 (0.3166)
Cauchy	-123.2769	250.5539	0.0995 (0.5905)	0.7613 (0.5085)	0.1511 (0.522)
GIW	-119.3165	244.6331	0.1304 (0.4577)	0.8139 (0.4700)	0.1566 (0.476)
MEC	-116.9755	239.9510	0.0800 (0.6958)	0.4572 (0.7893)	0.1351 (0.6651)

Table 5. Comparison Criteria for Dataset II

Distribution	-L	AIC	CVM (P-Value)	AD (P-value)	KS (P-value)
OCE	**-153.3405**	**312.6810**	**0.0932 (0.6221)**	**0.6345 (0.6148)**	**0.1344 (0.5907)**
OCW	-153.3508	314.7016	0.0932 (0.6219)	0.6351 (0.6142)	0.1342 (0.5921)
OCR	-153.4761	312.9522	0.0946 (0.615)	0.6417 (0.6082)	0.1354 (0.5804)
Rayleigh	-188.6356	379.2713	1.9215 (1.14e-05)	23.5960 (1.818e-05)	0.4255 (1.296e-05)
Cauchy	-172.8878	349.7756	0.6688 (0.0146)	4.2108 (0.0070)	0.2931 (0.0069)
GIW	-155.9985	317.9971	0.1444 (0.4088)	0.8935 (0.4174)	0.1490 (0.4561)
MEC	-154.3179	314.6358	0.0908 (0.6348)	0.6242 (0.6242)	0.1326 (0.6078)

**p-values are in bracket*

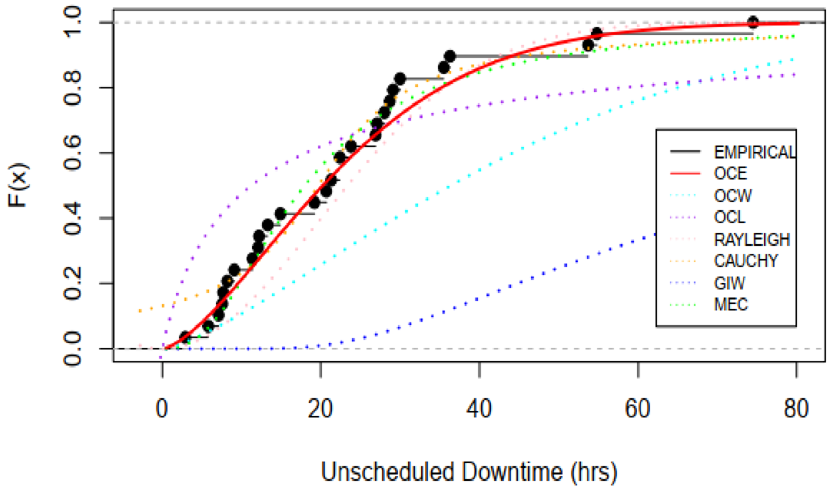

Fig. 5. CDF Plot for the OCE and the selected distributions

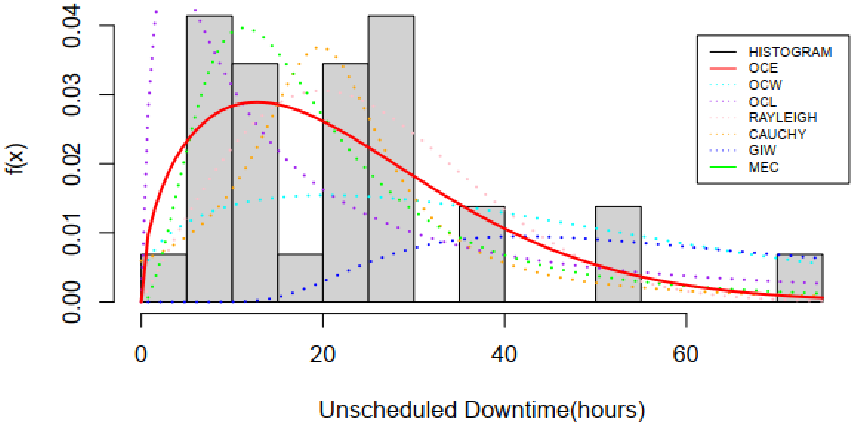

Fig. 6. PDF Plots for the OCE and the selected distributions

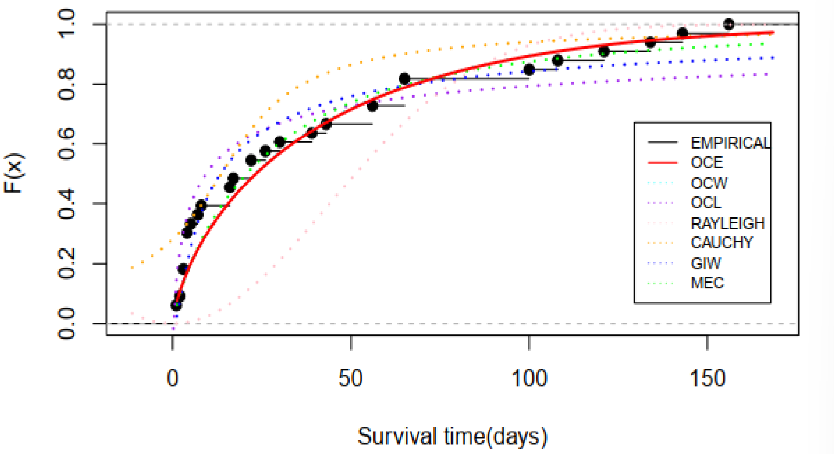

Fig. 7. CDF Plot for the OCE and the selected distributions

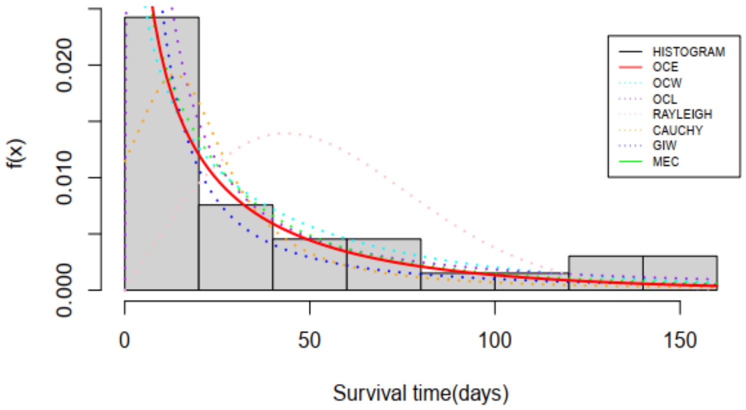

Fig. 8. PDF Plot for the OCE and the selected distributions

Fig. 5 and 6 show plots of some selected distributions. It can also be observed in both Fig. 5 and 6 that the OCE distribution is a better fit for the Unscheduled Downtime dataset (dataset I) than the other selected distributions. This confirms the observations in Table 2.

Furthermore, Figs. 7 and 8 show a plot of the selected distributions and it can be observed that the OCE distribution is a better fit for the Unscheduled Downtime dataset than the other selected distributions.

4 Conclusion

In this study, a new continuous probability distribution with a flexible hazard rate is introduced and discussed for the Oden Chen Distribution. Its properties were investigated accordingly. The parameter estimates were carried out using the maximum likelihood method.

The hazard function displays different shapes of failure rates, both monotonic and non-monotonic. Datasets of varying kinds can mostly fit well with distributions with the type of hazard function. The generated distribution exhibits varied shapes of density and failure rates for various combinations of parameter values, making it particularly adaptable in modelling datasets. Additionally, the outcomes of conducted simulations demonstrate that the maximum likelihood estimators of the distribution are consistent. Two real-life data were fitted with this distribution, and the results revealed that the OCE distribution performed better than the Odd Chen Weibull, Odd Chen Rayleigh, Rayleigh, Cauchy, Generalised Inverse Weibull and the Modified Extended Chen distributions.

Competing Interests

Authors have declared that no competing interests exist.

References

[1] Anafo AY, Brew L, Nasiru S. The modified Extended Chen Distribution: Properties and application to rainfall data. Appl. Math. 2022;16(5):711-728.

[2] Luguterah A. Odd generalized exponential Rayleigh distribution. Advances and Applications in Statistics. 2016;48(1):33-48.

[3] Chaubey YP, Zhang R. An extension of Chen's family of survival distributions with bathtub shape or increasing hazard rate function. Communications in Statistics-Theory and Methods. 2015;44(19):4049-4064.

[4] Bourguignon M, Silva RB, Cordeiro GM. The weibull-G family of probability distributions. Journal of Data Science. 2014;53-68.

[5] Tahir MH, Cordeiro GM, Alizadeh M, Mansoor M, Zubair M, Hamedani GG. The odd generalized exponential family of distributions with applications. Journal of Statistical Distributions and Applications. 2015;2(1):1-28.

[6] Khan M, Sharma, A. Generalized order statistics from Chen distribution and its characterization. J. of Stat. Appl. and Prob. 2016;5(1):123-128.

[7] Dey S, Kumar D, Ramos PL, Louzada F. Exponentiated Chen distribution: Properties and estimation. Communications in Statistics-Simulation and Computation. 2017;46(10):8118-8139.

[8] Dzivor ND, Otoo H, Wiah EN. Exponentiated janardan distribution: A modification of janardan distribution and application to life time data. International Journal of Probability and Statistics. 2021;10(1):9-16.

[9] Dzivor ND, Otoo H, Wiah EN. Kumaraswamy-Janardan distribution: A generalized Janardan distribution with application to real data. Asian J. of Prob. and Stats. 2021;15(4):111-122.

[10] Anzagra L, Sarpong S, Nasiru S. Odd Chen-G family of distributions. Annals of Data Science. 2022;9(2):369-391.

[11] Eliwa MS, El-Morshedy, M. and Ali, S. Exponentiated odd Chen-G family of distributions: statistical properties, Bayesian and non-Bayesian estimation with applications. Journal of Applied Statistics. 2021; 48(11): 1948-1974.

[12] Swain, J.J., Venkatraman S, Wilson JR. Least-squares estimation of distribution functions in Johnson's translation system. Journal of Statistical Computation and Simulation. 1988;29(4):271-297.

[13] Kao JH. A graphical estimation of mixed Weibull parameters in life-testing of electron tubes. Technometrics. 1959;1(4):389-407.

[14] Cheng RCH, Amin NAK. Estimating parameters in continuous univariate distributions with a shifted origin. Journal of the Royal Statistical Society: Series B (Methodological).1983;45(3):394-403.

[15] Ranneby B. The maximum spacing method. An estimation method related to the maximum likelihood method. Scandinavian Journal of Statistics. 1984;93-112.

[16] Cheng RCH, Amin NAK. Maximum likelihood estimation of parameters in the inverse Gaussian distribution, with unknown origin. Technometrics. 1981;23(3):257-263.

[17] Luciano JV, Belloch A, Algarabel S, Tomás JM, Morillo C, Lucero M. Confirmatory factor analysis of the white bear suppression inventory and the thought control questionnaire: A comparison of alternative models. European Journal of Psychological Assessment. 2006;22(4):250-258.

[18] Anderson TW, Darling DA. Asymptotic theory of certain "Goodness of Fit" criteria based on stochastic processes. The Annals of Mathematical Statistics. 1952;193-212.

[19] Macdonald P. Estimation procedure for mixtures of distributions. Journal of The Royal Statistical Society Series B-Statistical Methodology. 1971;33(2):326-350.

[20] Stephens MA. EDF statistics for goodness of fit and some comparisons. Journal of the American Statistical Association. 1974;69(347):730-737.

Peer-review history:
The peer review history for this paper can be accessed here (Please copy paste the total link in your browser address bar)
https://www.sdiarticle5.com/review-history/106376