Asian Journal of Probability and Statistics

Volume 23, Issue 4, Page 26-42, 2023; Article no.AJPAS.103987
ISSN: 2582-0230

An Extension of the Chen Distribution: Properties, Simulation Study and Applications to Data

Joseph Acquah [a], Benjamin Odoi [a], Abdulzeid Yen Anafo [a*] and Bosson-Amedenu Senyefia [a]

[a] *Department of Mathematical Sciences, University of Mines and Technology, 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/v23i4510

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/103987

Received: 29/05/2023
Accepted: 01/08/2023
Published: 07/08/2023

Original Research Article

Abstract

In this study, a new three parameter extension of the Chen distribution was proposed and called the New Extended Chen distribution. Some statistical properties of the proposed distribution are presented. The proposed distribution exhibits varied complex and hazard shapes. Parameters of the distribution are estimated using the maximum likelihood estimation method and a simulation study is conducted to evaluate the performance of the estimators. The New Extended Chen distribution is applied to two real data set and compared to other modifications of the Chen distribution to emphasise the applicability of the the distribution.

Keywords: Chen distribution; estimation; hazard; new extended family.

2010 Mathematics Subject Classification: 53C25; 83C05; 57N16.

**Corresponding author: E-mail: ayanafo@umat.edu.gh;*

1 Introduction

In the field of univariate statistical analysis, the modelling of monotonic hazard rates may be done using Weibull, gamma, lognormal, exponential, and other underlying distributions which are examples of conventional distributions. However, the limitations in some of these distributions makes it challenging in modelling data set that exhibits flexible characteristics such as bathtub and upside down bathtub features for their hazard functions. Majority of these traditional models mentioned only show hazard rates that are monotonically increasing, constant and decreasing. In most data sets generated from the field of finance, climatology, geology, hydrology, ecology, health sciences, reliability, life testing and risk analysis, the most important and realistic shape is the bathtub-shaped, increasing and decreasing. Although, there exist many statistical distributions with some level of flexibility, the continues production of data sets from various fields do not fit to some of these developed distributions.

The methods of generating new distributions, has been an area of interest to researchers. There are several methods used to generate distributions notable amongst them is the method of using an underlining family or generator. Some important families include odd exponentiated half-logistic-G [1], Topp Leone odd Lindley-G [2], Marshall–Olkin alpha power-G [3], transmuted transmuted-G [4], Weibull Marshall–Olkin [5], Extended odd Fréchet-G [6], rT-X Family [7] and [8] families. Recently, [9] established a new approach of introducing an additional shape parameter to modify the flexibility of existing distributions, New Extended Family NE-F with CDF defined by

$$G(x;\theta,a) = F(x;a)e^{\bar{\theta}F(x;a)} \tag{1.1}$$

and PDF,

$$g(x;\theta,\xi) = f(x;\xi)e^{\bar{\theta}F(\bar{x};\xi)}\left[1-\bar{\theta}F(x;\xi)\right] \tag{1.2}$$

where $x \in R$

Chen [10] a novel two-parameter distribution with an increasing failure rate function or bathtub. The advantage of this distribution can be associated with the exact confidence intervals and joint confidence region for the parameter [11]. In order to increase the flexibility and increase the number of alternative hazard shapes for the distribution, the family of distribution proposed by [12] was applied by [13] to propose the exponentiated Chen distribution. The Kumaraswamy Chen distribution was obtained [13] to increase the flexibility of the Chen by applying the Kumaraswamy family. Subsequently The Marshall-Olkin technique was applied by [11] in order to develop the Marshall-Olkin Chen distribution. The Chen-geometric and Marshall-Olkin Chen distributions can be viewed as similar distributions with the same parameters. Subsequently, Another compounding distribution was proposed by [14], who studied the Chen-logarithmic distribution and also extended the parameter space of the logarithmic distribution. In recent times many versions of the Chen distribution has been proposed and can be found in [15], [16], [17] and [18].

The purpose of this study is to developed a new flexible extension of the Chen distribution in light of the aforementioned limitations.[10] using the New Extended family (NE-F) introduced by [9]. The advantage of using the Extended Family is from the fact that the NE-F introduces an additional shape parameter θ to the existing two parameter Chen distribution which increases the flexibility by increasing its skewness. The study of the proposed distribution is motivated by the following:

- To develop a distribution which exhibits unimodal, and varying skewness and kurtosis.
- To develope a new extension of the Chen distribution that can be used as a better substitute for other other extensions of the Chen distribution in modelling lifetime data sets.

The remainder of the paper is organized as follows. In Section 2, the model formulation of the new lifetime distribution is discussed together followed by the study of its properties, including the shapes of the probability density function (PDF) and hazard function (HF). Section 3 presents the graphical shapes of the PDF and hazard function of the New Extended Chen distribution. Section 4 presents some statistical properties of the New Extended Chen. The maximum likelihood methods are presented in Section 5. In Sections 6, we obtain the best method of estimating the parameters of the New Extended Chen distribution by obtaining the maximum likelihood estimates through a simulation study. The New Extended Chen distribution is applied to two data sets in Section 7 and the results are compared with different known modified Chen distributions. Finally some conclusions draw from the study are presented in section 8.

2 Model Formulation

Let X be a random variable following a Chen distribution [10] with CDF and PDF given by

$$f(x) = 1 - e^{d(1-e^{x^b})} \quad x \geq 0 \tag{2.1}$$

and probability density function given as,

$$f(x) = bdx^{b-1}e^{d(1-e^{x^b})x^b}. \tag{2.2}$$

where $b > 0$ is the shape parameter and d is the scale parameter. The CDF of the NEC distribution is defined by substituting the CDF of the Chen distribution in Equation 2.1 into the NE-F CDF in Equation 1.1. Then, the New Extended Chen Distribution (NEC) distribution with parameters $b > 0, d > 0$ and $\theta > 0$ is given by

$$F(x; b, d, \theta) = 1 - \exp\left[d\left(1 - e^{x^b}\right) + \bar{\theta}\left(1 - e^{d\left(1-e^{x^b}\right)}\right)\right], x \geq 0, \tag{2.3}$$

where $b > 0$, $\theta > 0$ are the shape parameters and $d > 0$ is the scale parameter. The PDF of the NEC distribution is subsequently derived by differentiating the CDF which is given by;

$$f(x) = bdx^{b-1}\exp\left[x^b e^{d\left(1-e^{x^b}\right)} + \bar{\theta}\left(1 - \left(1 - e^{d\left(1-e^{x^b}\right)}\right)\right)\right]\left[1 - \bar{\theta}\left(1 - e^{\left(d\left(1-e^{x^b}\right)\right)}\right)\right] \tag{2.4}$$

with survival and hazard functions derived as shown in Equations (7) and (8) respectively

$$S(x) = \exp\left[d\left(1 - e^{x^b}\right) + \bar{\theta}\left(1 - e^{d\left(1-e^{x^b}\right)}\right)\right] \tag{2.5}$$

$$h(x) = \frac{bdx^{b-1}\exp\left[x^b e^{d\left(1-e^{x^b}\right)} + \bar{\theta}\left(1 - \left(1 - e^{d\left(1-e^{x^b}\right)}\right)\right)\right]\left[1 - \bar{\theta}\left(1 - e^{\left(d\left(1-e^{x^b}\right)\right)}\right)\right]}{\exp\left[d\left(1 - e^{x^b}\right) + \bar{\theta}\left(1 - e^{d\left(1-e^{x^b}\right)}\right)\right]}. \tag{2.6}$$

3 Shapes of the PDF and Hazard Function

The PDF and hazard plots of the NEC distribution are displayed for some parameter values in Fig. 1(a) and 1(b) respectively. The theoretical analysis of Fig. 1 displays various shapes of the NEC distribution. In Fig. 1(a) the various parameter values that were generally used for the density plots are unimodal. It can be also be observed that the density plots exhibits both right and left skewed shapes. In addition, for some parameter values, the density exhibits a reverse J (exponential decreasing) and increasing shapes. Afterwards, the hazard plots in Fig. 1(b) exhibits both right and left skewed shapes. Also the hazard plots exhibits upside down bathtub shape.

Lemma 3.1. *The PDF of the NEC distribution can be expressed in a series representation using the generalised PDF series expansion of the NE-F family introduced in [9]. Thus ,*

$$f(x) = \sum_{i,j=0}^{\infty}\sum_{m=0}^{k}\sum_{n=0}^{\infty}\sum_{p=0}^{n}\omega_{jkmnp}x^{b(p+1)-1}e^{-px^b}, \tag{3.1}$$

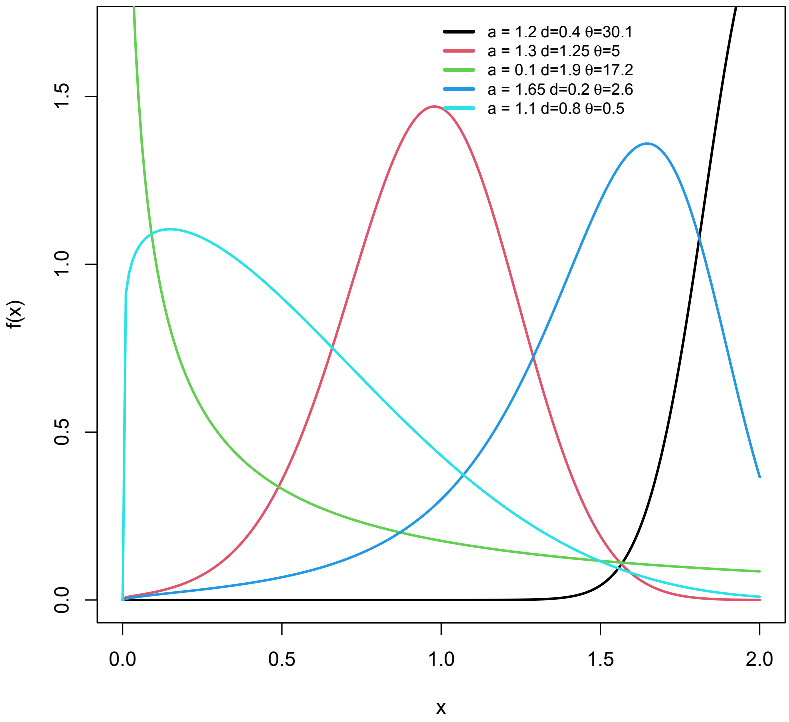

Fig. 1. PDF plots of the NEC distribution

where $x > 0, d > 0, b > 0$ *and* $\theta > 0$ *and*

$$\omega = \binom{j}{k}\binom{k}{m}\binom{n}{p}\frac{\bar{\theta}(-1)^{k+m+q+n}(k+1)(n+1)^p d^{(n+1)}(m+1)^n b}{j!n!p!}$$

The proof of Lemma 2.1 is straightforward using series expansions techniques and, therefore, are omitted. The New Extended family PDF introduced by [9] can be expressed as

$$g(x;\theta,\xi) = \sum_{i,j=0}^{\infty}\frac{\bar{\theta}(-1)^k(k+1)}{j!}\binom{j}{k}f(x;\xi)(F(x;\xi)^k). \tag{3.2}$$

Substituting the Equations ... and ..... into Equation 3.2,

$$f(x) = \sum_{i,j=0}^{\infty}\frac{\bar{\theta}(-1)^k(k+1)}{j!}\binom{j}{k}bdx^{b-1}e^{d(1-e^{x^b})x^b}\left\{1-e^{d(1-e^{x^b})}\right\}^k, \tag{3.3}$$

Considering the expansion,

$$e^x = \sum_{m=0}^{\infty}\frac{x^m}{m!}$$

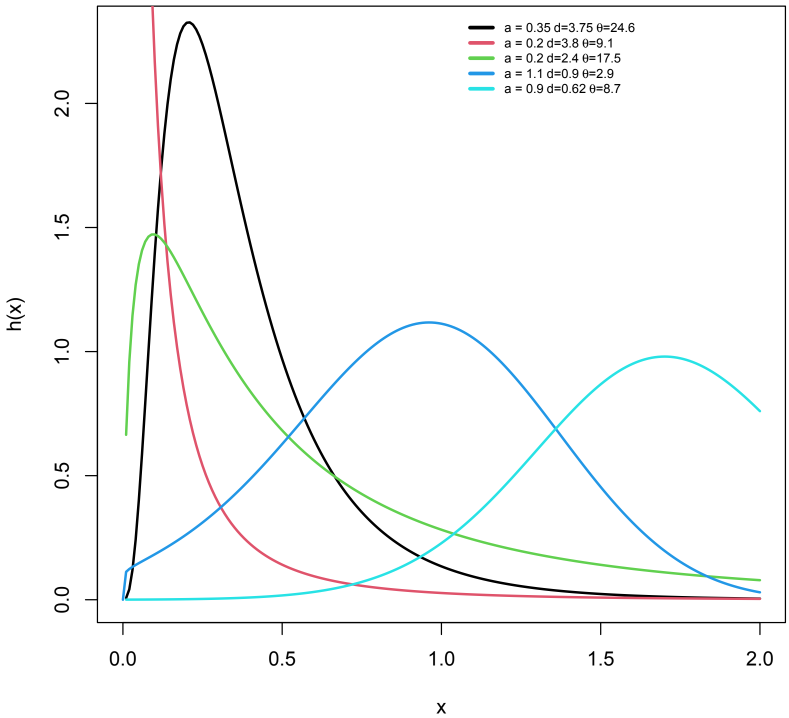

Fig. 2. Hazard plots of the NEC distribution

then,

$$e^{d(1-e^{x^b})x^b} = \sum_{m=0}^{\infty} \frac{d^m(1-e^{x^b})^m x^{bm}}{m!} \tag{3.4}$$

Consider the power series,

$$(1-z)^t = \sum_{i=0}^{\infty}(-1)^t \binom{t}{i} z^t$$

where $|z| < 1$, then the equation

$$(1-e^{x^b})^m = \sum_{p=0}^{m}(-1)^m \binom{m}{p} e^{mx^b}. \tag{3.5}$$

Considering $\left\{1-e^{d(1-e^{x^b})}\right\}^k$, we apply the power series equation, thus

$$\left\{1-e^{d(1-e^{x^b})}\right\}^k = \sum_{q=o}^{k}(-1)^q \binom{q}{k} e^{kd(1-e^{x^b})} \tag{3.6}$$

further,

$$e^{kd(1-e^{x^b})} = \sum_{n=0}^{\infty} \frac{k^n d^n}{n!}(1-e^{x^b})^n \tag{3.7}$$

then,

$$(1-e^{x^b})^n = \sum_{h=0}^{n}(-1)^n \binom{n}{h} e^{nx^b}. \tag{3.8}$$

Inserting Equations 3.4,3.5, 3.6, 3.7 and 3.8 into Equation 3.3,

$$f(x) = \sum_{i,j,m=0}^{\infty} \sum_{p=0}^{m}\sum_{q=0}^{k} \frac{\bar{\theta}(-1)^{k+m+q+n}(k+1)d^{(m+n+1)}b}{j!m!n!} \binom{n}{h}\binom{q}{k}\binom{m}{p} x^{bm+b-1} e^{(m+n)x^b}$$

The proof is complete.

4 Statistical Properties

In this section, some statistical properties of the NEC distribution are presented to authenticate the uniqueness of the proposed distribution.

4.1 Quantile, median and quartiles

Suppose X has the NEC distribution, then the Quantile function is given by

$$Q(u) = \frac{Log[(1+d-\theta - Log[1-u] + W[-e^{(-1+\theta)}(-1+\theta)(-1+u)])}{d}^{1/b} \tag{4.1}$$

where $u \in (0,1)$ and $W[\cdot]$ is the Lambert W function. The first quartile, the median, and the upper quartile can be obtained by using the quantile function by setting $p = 0.25, 0.5, 0.75$. In addition, the quantile function can be applied to generate random data using the quantile of the NEC distribution.

4.2 Moments

The moments of a random variable, if they exist, are important tools for deriving statistical measures such as standard deviation and variance of data sets, and they aid in identifying the varied pictorial shapes of the population.

Let X follow the NEC distribution, then the r^{th} moment can be derived as

$$\mu_r^{'} = \int_0^{\infty} x^r f(x)dx. \tag{4.2}$$

Substituting Equation (3.1) into Equation (4.2), we obtain

$$\mu_r^{'} = \sum_{i,j=0}^{\infty}\sum_{m=0}^{k}\sum_{n=0}^{\infty}\sum_{p=0}^{n} \omega_{jkmnp} \int_0^{\infty} x^{r+b(p+1)-1} e^{-px^b} dx. \tag{4.3}$$

Let $u = px^b$, as $x \to 0$, $u \to 0$ and as $x \to \infty$, $u \to \infty$, which implies $x = (\frac{u}{p})^{1/b}$ and $dx = \frac{du}{pbx^{b-1}}$. By substituting x and dx we obtain,

$$\mu_r^{'} = \sum_{i,j=0}^{\infty}\sum_{m=0}^{k}\sum_{n=0}^{\infty}\sum_{p=0}^{n} \omega_{jkmnp} \int_0^{\infty} \frac{u^{\frac{bp+r+1}{b}+1-1} e^{-u}}{p^{\frac{bp+b+r+1}{b}} b} du.$$

Using the identity

$$\Gamma(\alpha) = \int_0^\infty y^{\alpha-1} e^{-y} \mathrm{d}y,$$

we obtain the r^{th} moment of the NEC distribution as

$$\mu_r^{'} = \sum_{i,j=0}^{\infty} \sum_{m=0}^{k} \sum_{n=0}^{\infty} \sum_{p=0}^{n} \omega_{jkmnp} \frac{\Gamma(\frac{bp+r+1}{b}+1)}{p^{\frac{bp+b+r+1}{b}} b}. \tag{4.4}$$

The first four moments of the random variable X are given by;

$$E[X] = \sum_{i,j=0}^{\infty} \sum_{m=0}^{k} \sum_{n=0}^{\infty} \sum_{p=0}^{n} \omega_{jkmnp} \frac{\Gamma(\frac{bp+2}{b}+1)}{p^{\frac{bp+b+2}{b}} b}$$

$$E[X^2] = \sum_{i,j=0}^{\infty} \sum_{m=0}^{k} \sum_{n=0}^{\infty} \sum_{p=0}^{n} \omega_{jkmnp} \frac{\Gamma(\frac{bp+3}{b}+1)}{p^{\frac{bp+b+3}{b}} b}$$

$$E[X^3] = \sum_{i,j=0}^{\infty} \sum_{m=0}^{k} \sum_{n=0}^{\infty} \sum_{p=0}^{n} \omega_{jkmnp} \frac{\Gamma(\frac{bp+4}{b}+1)}{p^{\frac{bp+b+4}{b}} b}$$

and

$$E[X^4] = \sum_{i,j=0}^{\infty} \sum_{m=0}^{k} \sum_{n=0}^{\infty} \sum_{p=0}^{n} \omega_{jkmnp} \frac{\Gamma(\frac{bp+5}{b}+1)}{p^{\frac{bp+5}{b}} b}.$$

4.3 Incomplete moment

The incomplete moment is used to complete certain measures such as Lorenz curve, Gini, and pietra measures of inequality across two populations. In this study we obtain the incomplete moment of the NEC using the relation

$$m_r(y) = E(X^r | X \leq y) = \int_0^y x^r f(x) \mathrm{d}x. \tag{4.5}$$

Substituting Equation 3.1 into Equation 4.5 we obtain

$$m_r(y) = \sum_{i,j=0}^{\infty} \sum_{m=0}^{k} \sum_{n=0}^{\infty} \sum_{p=0}^{n} \omega_{jkmnp} \int_0^y x^{r+b(p+1)-1} e^{-px^b}. \tag{4.6}$$

We again let $u = px^b$, as $x \to 0$, $u \to 0$ and as $x \to y$, $u \to py^b$ which implies $x = (\frac{u}{p})^{1/b}$ and $dx = \frac{du}{pbx^{b-1}}$. By substituting x and dx we obtain

$$m_r(y) = \sum_{i,j=0}^{\infty} \sum_{m=0}^{k} \sum_{n=0}^{\infty} \sum_{p=0}^{n} \omega_{jkmnp} \int_{py^b}^{\infty} \frac{u^{\frac{bp+r+1}{b}+1-1} e^{-u}}{p^{\frac{bp+b+r+1}{b}+1} b} du.$$

Using the identity

$$\Gamma(\alpha, x) = \int_x^\infty t^{\alpha-1} e^{-t} \mathrm{d}t.$$

The incomplete moment of the NEC distribution can be defined as

$$m_r(y) = \sum_{i,j=0}^{\infty} \sum_{m=0}^{k} \sum_{n=0}^{\infty} \sum_{p=0}^{n} \omega_{jkmnp} \frac{\Gamma(\frac{bp+r+1}{b}, py^b)}{p^{\frac{bp+b+r+1}{b}} b}.$$

4.4 Moment generating function

The moment generating function (MGF) if it exist is a special functions used to compute the moments. Given a random variable X having the NEC distribution is given by The MGF of Y is defined as,

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

Using the results obtained in Equation (13) the MGF of the NEC distribution reduces to

$$M_X(t) = \sum_{i,j=0}^{\infty}\sum_{m=0}^{k}\sum_{n=0}^{\infty}\sum_{p=0}^{n}\sum_{j=0}^{i}\omega_{jkmnp}\frac{t^r\Gamma(\frac{bp+r+1}{b}+1)}{r!p^{\frac{bp+b+r+1}{b}}b}.$$

4.5 Beforroni and lorenz curve

Measures of income disparity using Lorenz and Bonferroni curves are broadly relevant and applicable to different fields as reliability, demography, medicine, and insurance. In section, We derive the Lorenz and Bonferroni curves for the NEC distribution in this section. The Lorenz curve for a random variable having the NEC distribution is derived from the definition of the Lorenz curve given by,

$$L_G(x) = \frac{1}{\mu}\int_0^y xf(x)dx$$

the Lorenz curve is simply the product of the first incomplete moment and the reciprocal of the mean of the random variable. Hence, the Lorenz curve of the NEC distribution is given by,

$$L_{NEC}(x) = \sum_{i,j=0}^{\infty}\sum_{m=0}^{k}\sum_{n=0}^{\infty}\sum_{p=0}^{n}\omega_{jkmnp}\frac{\Gamma(\frac{bp+2}{b}+1)}{\mu p^{\frac{bp+b+2}{b}}b} \tag{4.7}$$

The Befforoni curve on the other hand is defined as,

$$B_(G)(x) = \frac{L_G(x)}{F(x)}. \tag{4.8}$$

From Equation (17) and (18) the Befforoni curve of the NEC distribution is given by,

$$B_G(x) = \sum_{i,j=0}^{\infty}\sum_{m=0}^{k}\sum_{n=0}^{\infty}\sum_{p=0}^{n}\omega_{jkmnp}\frac{\Gamma(\frac{bp+2}{b})}{F(x)\mu p^{\frac{bp+b+2}{b}}b}.$$

5 Maximum Likelihood Estimation Method

In this section, we employ the maximum likelihood estimation (MLE) method to obtain the estimators of the parameters b, d and θ of the NEC distribution. The method of maximum likelihood estimation (MLE) is one of the most used estimation technique used for parameter estimation due to its desirable characteristics. Let $x_1, \ldots x_n$ be a random sample of size n from Equation (6), then the log-likelihood function of the NEC distribution is given by;

$$\begin{aligned}\ln L\left(x; b, d, \theta\right) =& n\ln b + n\ln d + (b-1)\ln\sum_{i=1}^{n} x_i + \sum_{i=1}^{n} x_i^b + \sum_{i=1}^{n} e^{d(e^{x_i^b}-1)} - (1-\theta)\sum_{i=1}^{n}(1-e^{d(1-e^{x_i^b}}))+(1-\theta)\\ &\sum_{i=1}^{n}\ln(1-e^{(d(1-e^{x_i^b}))})\end{aligned} \tag{5.1}$$

For uniformity, the first partial derivatives of Equation 5.1 are denoted by L_b, L_d and L_θ, By setting $L_b = 0$, $L_\theta = 0$ and $L_d = 0$, we obtain Equations (20), (21) and (22).

$$
\begin{aligned}
L_b = \frac{n}{b} + \sum_{i=1}^{n} x_i^b \ln(x_i) + \sum_{i=1}^{n} de^{d(-1+e^{x_i^b})+x_i^b} + \ln(x_i) + \sum_{i=1}^{n} de^{d(-1+e^{{x_i}^{b}})+x_i^b}(1-\theta)x_i^b \ln(x_i) \\
+ \sum_{i=1}^{n} \frac{de^{d(-1+e^{x_i^b})+x_i^b}(1-\theta)x_i^b(x_i)}{1-e^{d(-1+e^{x_i^b})})} = 0,
\end{aligned} \tag{5.2}
$$

$$
L_\theta = \sum_{i=1}^{n} 1 - e^{(d(-1+e^{x_i^b}))} + \sum_{i=1}^{n} \ln(1 - e^{(d(1-e^{x_i^b}))}) = 0, \tag{5.3}
$$

$$
L_d = \frac{n}{d} + \sum_{i=1}^{n} x_i^b e^{d(-1+e^{x_i^b})}(e^{x_i^b - 1}) + \sum_{i=1}^{n} e^{d(-1+e^{x_i^b})} e^{({x_i}^{b}-1)(\theta-1)} \frac{e^{d(-1+e^{x_i^b})+x_i^b}(1-\theta)x_i^b x_i}{1-e^{d(-1+e^{x_i^b})})} = 0, \tag{5.4}
$$

The estimators of the NEC distribution can be obtained by solving equation (20), (21) and (22) in relation to a, θ and d simultaneously using a numerical procedure.

6 Simulation Study

We examine the behavior of the NEC estimators provided in section 4 for the NEC distribution's parameters. To verify the validity of these estimators, the average bias (AB) and mean square error (MSE) are computed. In the study we consider sample sizes $n = 5, 100, 250, 300, 350$ and 500 were considered. The R software was used to estimate the proposed models parameters. The experiment is replicated for $10,000$ times to compare the AB and MSE of th four estimation techniques suggested in section 4. The AB and MSE are computed using,

$$
AB = \frac{1}{N}\sum_{i=0}^{N}\left(\hat{\theta}_i - \theta\right) \qquad MSE = \frac{1}{N}\sum_{i=1}^{N}\left(\hat{\theta}_i - \theta\right)^2.
$$

where $r = 10,000$ and $\hat{\theta_i}$ is the estimation parameters of the NEC distribution that is $\theta = (b, d, \theta)$. The simulation results for $\theta = (0.9, 9.8, 5.9)$, $\theta = (9.2, 0.5, 3.4)$ and $\theta = (9.1, 2.3, 5.6)$ are presented in Table 1. The selection of the best estimation method will be made having a minimum estimate of MSEs and decreasing ABs. The R software is used to derive the simulation results. The results of ABs and MSEs for the MLE are presented in Tables 1. It is observed that the ABs decreases as the samples sizes increases and also the MSEs decrease with the increase in the sample size. It is sufficient enough to conclude that the estimators are unbiased asymptotically. In general, these results suggest that the estimation of parameters was performed consistently. Table 1 displays the simulation results of the ER-Kum distribution. It is clear that the estimators are consistent since both the RMSEs and ABs decreases with increasing sample size.

7 Data Analysis

In this section, we offer two examples using actual data sets that demonstrate the versatility of the NEC distribution. The computations are performed using the R software. The NEC distribution is compared with the Chen-Logarithmic [14], Exponentiated Chen [19], Modified Extended Chen (MEC) [15], Transmuted Chen [20] and Marshall-Olkin Chen [11] distribution. The competing distributions were compared using different goodness of fit criteria and information criteria such as;

i. Kolmogorov-Smirnov (KS) test

$$
KS = \{\ |G_{exp}(x_i) - G_{obs}(x_i)|, |G_{exp}(x_i) - G_{obs}(x_i)|\}
$$

where $G(x)$ is the CDF of the random variable X and $i = 1, 2, \ldots, n$.

Table 1. Simulation results

		$b = 0.9, d = 9.8, \theta = 5.9$		$b = 9.2, d = 0.5, \theta = 3.4$		$b = 9.1, d = 2.3, \theta = 5.6$	
Parameter	**n**	**AB**	**RMSE**	**AB**	**RMSE**	**AB**	**RMSE**
a	20	0.0568	0.0120	6.0932	3.4541	0.2614	0.0062
	50	0.0411	0.0111	1.0981	1.0009	0.2602	0.0026
	100	0.0303	0.0100	0.0019	0.9111	0.2601	0.0017
	150	0.0144	0.0050	0.0200	0.1113	0.2597	0.0014
	200	0.0065	0.0034	0.0136	0.0123	0.2596	0.0010
	250	0.0052	0.0020	0.0100	0.0032	0.2498	0.0008
	300	0.0021	0.0014	0.0078	0.0027	0.2390	0.0006
	400	0.0015	0.0003	0.0064	0.0008	0.2100	0.0005
	500	0.0012	0.0001	0.0011	0.0003	0.2111	0.0002
b	20	-0.9240	1.2309	-8.9120	5.9021	0.0012	1.0981
	50	-0.1092	3.9821	-0.9871	1.9081	-0.0007	0.9011
	100	0.2309	1.0091	-0.9970	1.4524	-0.0014	0.6202
	150	0.1267	0.2309	9.0971	1.2109	-0.0039	0.5400
	200	0.4509	0.0091	3.0871	1.1102	-0.0017	0.3206
	250	0.3201	0.0083	2.9811	1.0872	0.0014	0.2111
	300	0.2134	0.0068	1.0987	0.0271	0.0010	0.1111
	400	0.1009	0.0032	0.0389	0.0251	0.0009	0.1190
	500	0.0981	0.0004	0.0010	0.0012	0.0001	0.0010
θ	20	-0.9240	1.2309	-8.9120	5.9021	0.0012	1.0981
	50	-0.1092	3.9821	-0.9871	1.9081	-0.0007	0.9011
	100	0.2309	1.0091	-0.9970	1.4524	-0.0014	0.6202
	150	0.1267	0.2309	9.0971	1.2109	-0.0039	0.5400
	200	0.4509	0.0091	3.0871	1.1102	-0.0017	0.3206
	250	0.3201	0.0083	2.9811	1.0872	0.0014	0.2111
	300	0.2134	0.0068	1.0987	0.0271	0.0010	0.1111
	400	0.1009	0.0032	0.0389	0.0251	0.0009	0.1190
	500	0.0981	0.0004	0.0010	0.0012	0.0001	0.0010

ii. Anderson-Darling Test (A)

$$A = A^2\Big(1 + \frac{0.75}{n} + \frac{2.25}{n^2}\Big).$$

where $A^2 = -n - \frac{1}{n}\sum_{i=1}^{n}\{(2i-1)\log(u_i) + (2n+1-2i)\log(1-u_i)\}$ and n is the sample size

iii. Akaike Information Criterion (AIC)

$$AIC = -2\log\Delta(\hat{\theta}) + 2k.$$

where k is the number of parameters and Δ be the maximum value of the model's likelihood function

iv. Bayesian Information Criterion (BIC)

$$BIC = -2\log\Delta(\hat{\theta}) + k\log(n),$$

where $\Delta(\cdot)$ is the number of estimated parameters, k is the value of the Log-likelihood function and n is the sample size.

v. The corrected Akiake Information Criterion (AICc)

$$AICc = AIC + \frac{2k(k+1)}{n-k-1},$$

where k is the number of estimated parameters in the model.

7.1 Guinea pig data

In this section, the NEC distribution is applied to the guinea pigs data set [21, 22]. The data shows how long 72 guinea pigs survived after contracting virulent tubercle bacilli, measured in days. Also, the NEC distribution's suitability is evaluated in relation to a few alternative modifications of the Chen distribution.

Table 2. ML estimates and standard errors for Guinea pigs data

Model	ML estimates			Standard error		
	b	**d**	θ	**a**	**b**	θ
NEC	0.4091	1.2108	11.7199	0.0634	0.2692	5.3382
Chen	0.2081	0.7592	–	0.0342	0.0431	–
Chen-logarithmic	0.2082	0.7584	1.0081	0.1312	0.0940	1.3951
Exponentiated Chen	0.9951	0.4441	7.2090	0.3061	0.0801	4.0951
MEC	0.9493	1.21864	1.21749	0.27382	0.13345	0.53810
Transmuted Chen	0.1171	0.8090	0.7531	0.0252	0.0452	0.2032
Marshall-Olkin Chen	0.0033	1.1311	0.0161	0.0011	0.0431	0.0062

From Table 2, the NEC distribution performs better than other candidate distributions, studeid in the paper in terms of AIC, BIC, CAIC, KS, CVM and AD. Decision was taken based on the fact that NEC has the lowest values of the AIC, BIC, CAIC and the highest p-value of AD, CVM and KS. This confirms that the NEC distribution is more robust compared to the other six models studied.

Table 3. Goodness-of-fit statistics for Guinea pigs data

Model	**AIC**	**BIC**	**CAIC**	**KS (P-Value)**	**CVM**	**AD**
NEC	193.174	200.0039	193.5269	0.0824 (0.7114)	0.0831	0.5010
Chen	212.4820	217.0360	212.6560	0.1650 (0.040)	0.3670	2.1300
Chen-logarithmic	214.4820	221.312	214.8350	0.1650 (0.040)	0.3670	2.1300
Exponentiated Chen	194.3720	201.2020	194.7250	0.0900 (0.601)	0.0940	0.5852
MEC	242.1338	248.9638	242.4868	0.1962 (0.0078)	0.7665	4.7705
Transmuted Chen	211.2352	218.0651	211.5881	0.1581 (0.055)	0.3361	1.9501
Marshall-Olkin Chen	201.6521	208.4821	202.0053	0.1371 (0.134)	0.1992	1.1530

Fitting the distributions considered for this data, the histogram with their respective fitting distribution is revealed as shown in Fig. 3. It can be observed visually that the NEC distribution is more robust compared to the data compared to other models.

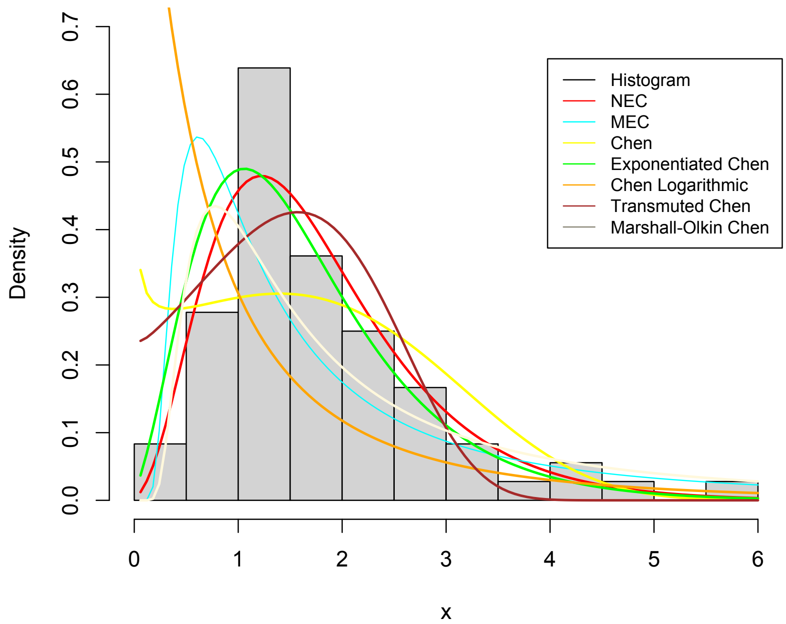

Fig. 3. Fitted PDFs for Guinea Pig data.

It can also be observed from Fig. (2) that amongst the CDF plots of all the distributions considered, the NEC distribution fits the Guinea Pig data better than the other seven models as it follows the pattern closely.

7.2 Failure times of component data

The failure times data contains failure times of 50 components (per 1000 h) as reported in [23]. The essence of failure times of component data is to determine the stability of components studied over time. In the field of engineering identifying the right model in analysing these data is essential in understanding the reliability of the components. In this section, we show that the NEC model is a better model for fitting failure times compared to the other extensions of the Chen models.

The log-likelihood, information criterion, and goodness-of-fit statistics for asteroid densities are presented in Table 5.

It is evident that the EMEC had the greatest L value and the lowest AIC, CAIC and BIC values among the seven models stated.

The NEC distribution, on the other hand, has the smallest W, A, and KS values as well as the biggest associated p-values.

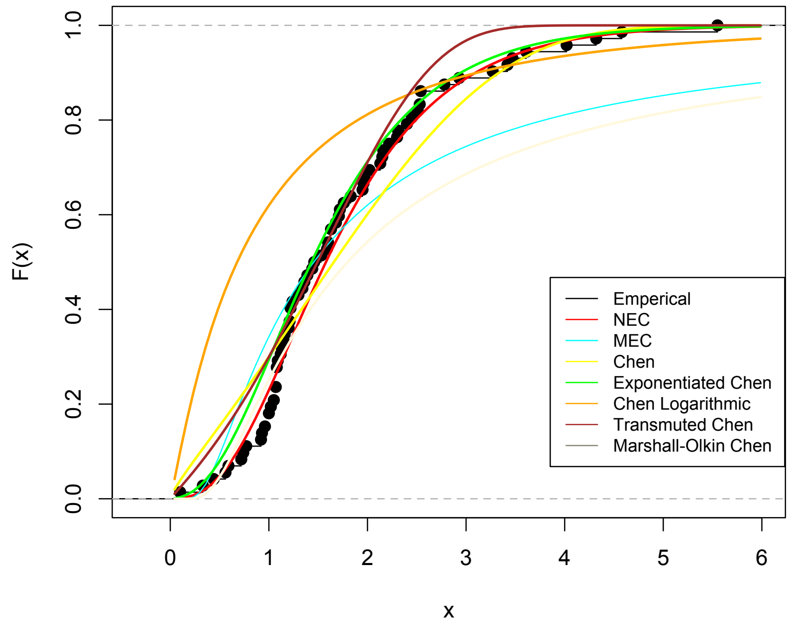

Fig. 4. Fitted CDFs for Guinea Pig data

Table 4. ML estimates and standard errors for failure times of component data.

Model	ML estimates			Standard error		
	b	**d**	θ	**a**	**b**	θ
NEC	0.2077	0.865	12.4205	0.0231	0.1870	4.5864
Chen	0.1114	0.3550	–	0.0155	0.0123	–
Chen-logarithmic	0.1102	0.9014	1.1901	0.1012	0.0140	2.3951
Exponentiated Chen	8.0951	1.3441	5.2090	0.2061	0.1801	3.0951
MEC	0.7797	1.1879	10.6965	0.1643	0.1084	5.2352
Transmuted Chen	0.2170	0.9091	0.8131	0.0341	0.0982	0.1132
Marshall-Olkin Chen	0.0077	1.1333	0.9061	0.0011	0.1211	0.2322

In light of this information, the NEC distribution clearly outperforms the other seven alternative models in terms of fitting asteroid densities.

Table 5. Goodness-of-fit statistics for failure times of component data

Model	AIC	BIC	CAIC	KS (P-Value)	CVM	AD
NEC	826.5814	835.1375	826.7749	0.0406 (0.9841)	0.0333	0.2184
Chen	866.3251	872.0291	866.4211	0.1429 (0.0107)	0.6879	4.3879
Chen-logarithmic	840.0911	851.9781	840.2391	0.0591 (0.8871)	0.4056	0.4511
Exponentiated Chen	851.2675	854.9011	851.3781	0.0810 (0.8911)	0.8170	0.4331
MEC	871.5463	880.1023	871.7398	0.1147 (0.0688)	0.56881	3.7862
Transmuted Chen	881.7611	889.8221	881.8191	0.2391 (0.0990)	0.3381	1.9501
Marshall-Olkin Chen	869.9101	878.2311	869.0911	0.4311 (0.1566)	0.2312	1.8951

In Fig. 5, we show the estimated PDFs of all fitted distributions. It is clear that, the NEC distribution matches the pattern of the asteroid densities data better than the other seven models, as shown in Fig. 5, which supports the results in Table 5.

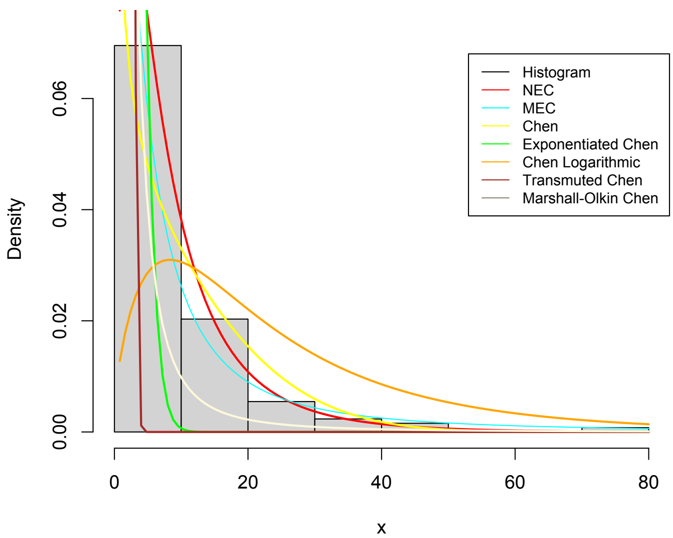

Fig. 5.Fitted PDFs for failure times of component data

Similarly, Fig. 5 shows a plot of all competing distributions versus the emperical CDF of the observed data. Fig. 5 shows the results of the analysis, which suggest that the NEC distribution is better suited to the data than the other competing distributions.

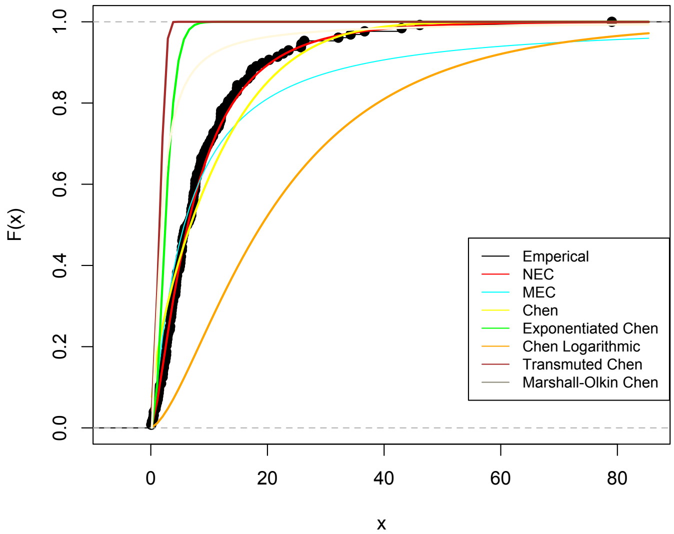

Fig. 6. Fitted CDFs for failure times of component data

8 Conclusion

An extension of the Chen distribution with three parameters is presented in this paper, and certain statistical features, including the Beferroni and Lorenz curves, the ordinary and incomplete moments, and generating functions are explored. The study also used the Maximum Likelihood estimation approach to estimate the model parameters, and a simulation study was used to confirm their performance. To highlight the potentiality of the suggested model, two applications to real data sets are compared to rival variants of the Chen distribution. In both applications, the NEC distribution clearly revealed to be suitable for modelling survival and failure times of data sets when compared using various goodness-of-fit tests.

Code Availability

Codes used during the study was done through the use of R programming language.

Competing Interests

Authors have declared that no competing interests exist.

References

[1] Afify AZ, Altun E, Alizadeh M, Ozel G, Hamedani GG. The odd exponentiated half-logistic-G family: properties, characterizations and applications. Chilean Journal of Statistics. 2017;8(2):65-91.

[2] Reyad H, Alizadeh M, Jamal F, Othman S. The Topp Leone odd Lindley-G family of distributions: Properties and applications. Journal of Statistics and Management Systems. 2018;21(7):1273-97.

[3] Eghwerido JT, Oguntunde PE, Agu FI. The alpha power Marshall-Olkin-G distribution: properties, and applications. Sankhya A. 2021:1-26.

[4] Nofal ZM, Afify AZ, Yousof HM, Cordeiro GM. The generalized transmuted-G family of distributions. Communications in Statistics-Theory and Methods. 2017;46(8):4119-36.

[5] Korkmaz MÇ, Cordeiro GM, Yousof HM, Pescim RR, Afify AZ, Nadarajah S. The Weibull Marshall–Olkin family: Regression model and application to censored data. Communications in Statistics-Theory and Methods. 2019;48(16):4171-94.

[6] Nasiru S. Extended odd Fréchet-G family of distributions. Journal of Probability and Statistics. 2018 Dec 2;2018:1-2.

[7] Ampadu CB, Anafo AY. The New rT-X Family of Distributions: Some Properties with Applications. Earthline Journal of Mathematical Sciences. 2019;2(2):409-32.

[8] Yousof H, Mansoor M, Alizadeh M, Afify A, Ghosh I. The Weibull-G Poisson family for analyzing lifetime data. Pakistan Journal of Statistics and Operation Research. 2020:131-48.

[9] Khosa SK, Afify AZ, Ahmad Z, Zichuan M, Hussain S, Iftikhar A. A new extended-f family: properties and applications to lifetime data. Journal of Mathematics. 2020;2020:1-9.

[10] Chen Z. A new two-parameter lifetime distribution with bathtub shape or increasing failure rate function. Statistics & Probability Letters. 2000;49(2):155-61.

[11] Cordeiro GM, Lemonte AJ, Ortega EM. The Marshall-Olkin family of distributions: Mathematical properties and new models. Journal of Statistical Theory and Practice. 2014;8:343-66.

[12] 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):4049-64.

[13] Nadarajah S, Cordeiro GM, Ortega EM. General results for the Kumaraswamy-G distribution. Journal of Statistical Computation and Simulation. 2012;82(7):951-79.

[14] Pappas V, Adamidis K, Loukas S. A three-parameter lifetime distribution. Adv. Applic. Statist. 2011;20: 159–167.

[15] Anafo AY, Brew L, Nasiru S. The Modified Extended Chen Distribution: Properties and Application to Rainfall Data. Appl. Math. 2022;16(5):711-28.

[16] Joshi S, Pandit PV. Estimation of stress strength reliability in s-out-of-k system for a two parameter inverse Chen distribution. Journal of Computer and Mathematical Sciences. 2018;9(12):1898-906.

[17] Reis LD, Cordeiro GM, Lima MC. The Gamma-Chen distribution: a new family of distributions with applications. Spanish Journal of Statistics. 2020;2(1):23-40.

[18] Joshi RK, Kumar V. Logistic Chen Distribution with Properties and Applications. International Journal of Mathematics Trends and Technology. 2021;67(1):141-51.

[19] Dey S, Kumar D, Ramos PL, Louzada F. Exponentiated Chen distribution: Properties and estimation. Communications in Statistics-Simulation and Computation. 2017;46(10):8118-39.

[20] Khan MS, King R, Hudson I. A new three parameter transmuted Chen lifetime distribution with application. Journal of Applied Statistical Science. 2013;21(3):239.

[21] Bjerkedal T. Acquisition of Resistance in Guinea Pies infected with Different Doses of Virulent Tubercle Bacilli. American Journal of Hygiene. 1960;72(1):130-48.

[22] Dey S, Nassar M, Kumar D. Alpha power transformed inverse Lindley distribution: A distribution with an upside-down bathtub-shaped hazard function. Journal of Computational and Applied Mathematics. 2019 Mar 1;348:130-45.

[23] Murthy DP, Xie M, Jiang R. Weibull models. John Wiley Sons; 2004.

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/103987