Pakistan Journal of Statistics and Operation Research

A new probability distribution: properties, copulas and applications in medicine and engineering

Mohamed K. A. Refaie[1,*] Nadeem Shafique Butt[2] Emadeldin I. A. Ali[3,4]

* Corresponding Author

[1]Agami High Institute of Administrative Sciences, Alexandria, Egypt; refaie_top@yahoo.com
[2]Department of Family and Community Medicine, King Abdul Aziz University, Jeddah, Kingdom of Saudi Arabia; nshafique@kau.edu.sa
[3]Department of Economics, College of Economics and Administrative Sciences, Al Imam Mohammad Ibn Saud Islamic University, Saudi Arabia; EIALI@IMAMU.EDU.SA
[4]Department of Mathematics, Statistics, and Insurance, Faculty of Business, Ain Shams University, Egypt; i_emadeldin@yahoo.com

Abstract
In this work, we construct a three-parameter Chen modification that is flexible. The "J shape", "monotonically increasing", "U shape," and "upside down (reversed bathtub)" hazard rate forms are all supported by the new Chen extension's hazard rate. We derive pertinent statistical features. A few distributions of the bivariate kind are generated. For evaluating the model parameters, we took the maximum likelihood estimation approach into consideration. Maximal likelihood estimators are evaluated via graphical simulations. To demonstrate the applicability of the new approach, two genuine data sets are taken into consideration and examined. The Akaike Information criterion, Bayesian Information criterion, Cramer-von Mises criterion, Anderson-Darling criterion, Kolmogorov-Smirnov test, and its related p-value are used to evaluate the new model with a variety of popular competing models.

Keywords: Chen Distribution; Maximum Likelihood; Simulations; Hazard Rate; Farlie-Gumbel-Morgenstern Copula; Clayton Copula.

Mathematical Subject Classification: 62N01; 62N02; 62E10.

1.Introduction and motivation

The Chen distribution is a continuous probability distribution that has been used in physics, engineering, and finance among other disciplines. The Chen distribution can describe a wide range of phenomena, including non-normal and skewed data, thanks to its adaptable shape. It is advantageous in circumstances where other distributions would not adequately match the data because of its flexibility. Because the moments of the Chen distribution have closed-form expressions, it is simpler to calculate statistical characteristics like mean, variance, skewness, and kurtosis. The Chen distribution is computationally efficient for simulations and other computational applications because it provides a straightforward and effective algorithm for producing random variations. The distribution of stock returns, interest rates, and other financial variables has been modelled in finance using the Chen distribution. It is excellent for modelling intricate financial phenomena due to its flexibility and effectiveness. The Chen distribution has numerous intriguing theoretical characteristics, including connections to the incomplete gamma function and the beta distribution. It is a topic of interest in mathematical research because of these characteristics. Let X be a non-negative random variable (RV) with Chen (C) distribution (see Chen (2000)) with cumulative distribution function (CDF) given by

$$G_{a,b}(x) = 1 - exp\{a[1 - exp(x^b)]\}, \quad (1)$$

where $x > 0, a > 0$ and $b > 0$. The C distribution has "increasing" and bathtub-shaped" hazard rate function (HRF). The HRF of the C model may have a "bathtub" shape when $b < 1$ and has "increasing" failure rate function when $b \geq 1$. Chen (2000) analyzed and compared the C model with many other relevant models. Recently, Chaubey and Zhang (2015) investigated a new probability density function (PDF) called the Exp-C model. Chaubey and Zhang (2015) studied the problem of estimating the Exp-C parameters. For the Exp-C model, Dey et al. (2017) sudied many statistical properties and presented some useful applications. They discussed a variety of estimation techniques, including the maximum likelihood estimation (MXLE), percentile estimation (PE), ordinary least square and weighted least square estimation (OLSE & WLSE), maximum product of spacings estimation (MPSE), Cramér-von Mises estimation (CVMSE), Anderson-Darling and right-tail Anderson-Darling estimation (ADE & RTADE), and more. The Burr type X Chen (BXC) distribution, a new iteration of the Chen distribution, will be derived in this work using the Burr X generator (BX-G). The HRF of the new BXC can be "J-HRF", "monotonically increasing HRF", "decreasing-constant-increasing HRF (bathtub HRF)", and "upside down HRF (reversed bathtub HRF)", as shown in Figure 2.

The Chen distribution is a generalization of the beta distribution in which the beta distribution's shape parameters are permitted to have negative values. The standard beta distribution is what happens when the shape parameters of the Chen distribution are set to 1/2 and -1/2, respectively. The normal distribution's mean and variance are permitted to be non-zero and non-positive in the Chen distribution, which is a generalization of the normal distribution. The Chen distribution decreases to the typical normal distribution when the shape parameters are set to 0 and -1/2, respectively. The gamma distribution, which is frequently used to model positive random variables, and the Chen distribution are linked. Specifically, the square root of a gamma-distributed random variable can be used to produce the Chen distribution. The CDF of the Burr type X G (BX-G) is defined as

$$F_{\vartheta,\underline{\xi}}(x) = \left\{1 - exp\left[-\boldsymbol{\mho}^2_{\underline{\xi}}(x)\right]\right\}^{\vartheta}, \quad (2)$$

where $\boldsymbol{\mho}_{\underline{\xi}}(x) = \frac{G_{\underline{\xi}}(x)}{\bar{G}_{\underline{\xi}}(x)}$. Inserting (1) into (2), the CDF of the BXC distribution can be written as

$$F_{\underline{\boldsymbol{\Psi}}}(x) = \left\{1 - exp\left[-\boldsymbol{\mho}^2_{a,b}(x)\right]\right\}^{\vartheta}, \quad (3)$$

where

$$\boldsymbol{\mho}^2_{a,b}(x) = \left[\frac{1 - \boldsymbol{q}_{a,b}(x)}{\boldsymbol{q}_{a,b}(x)}\right]^2.$$

and $\boldsymbol{q}_{a,b}(x) = exp\{a[1 - exp(x^b)]\}$. The relevant CDF of the BXC model can be expressed as

$$f_{\underline{\boldsymbol{\Psi}}}(x) = 2\vartheta ab \frac{x^{b-1}\, exp(x^b)\, exp\left[-\boldsymbol{\mho}^2_{a,b}(x)\right]\left[1 - \boldsymbol{q}_{a,b}(x)\right]}{\boldsymbol{q}^2_{a,b}(x)\left\{1 - exp\left[-\boldsymbol{\mho}^2_{a,b}(x)\right]\right\}^{1-\vartheta}}. \quad (4)$$

where $\underline{\boldsymbol{\Psi}} = (\vartheta, a, b)$. In order to explore the applicability and the flexibility of the new BXC PDF, we present Figure 1 which show the wide importance and flexibility of the new model. Based on Figure 1, we see that the new BXC PDF can be right-skewed and left-skewed with different useful PDF shapes. Analogously, Figure 2 is allocated to explore the new HRF. Based on Figure 2, we note that the new BXC HRF can be

i- "**J**-HRF" (see Figure 2 (top left)).
ii- "monotonically increasing HRF" (see Figure 2 (top right)).
iii- "decreasing-constant-increasing HRF (bathtub HRF)" (see Figure 2 bottom left).
iv- "upside-down HRF (reversed bathtub HRF)" (see Figure 2 bottom right)).

We are excited and motivated to define and analyze the BXC for the following main reasons:
i- The new PDF in (4) can be "right skewed" and "left skewed" with different shapes (see Figure 1).
ii- The HRF of the BXC model can be "**J**-shape", "monotonically increasing shape", "decreasing-constant-increasing shape (bathtub shape)" and "upside down shape (reversed bathtub shape)" (see Figure 2).

iii- In reliability analysis, the BXC model may be chosen as the best model, especially in modeling bimodal asymmetric real data which have right heavy tailed and the bimodal asymmetric data which have left tailed.

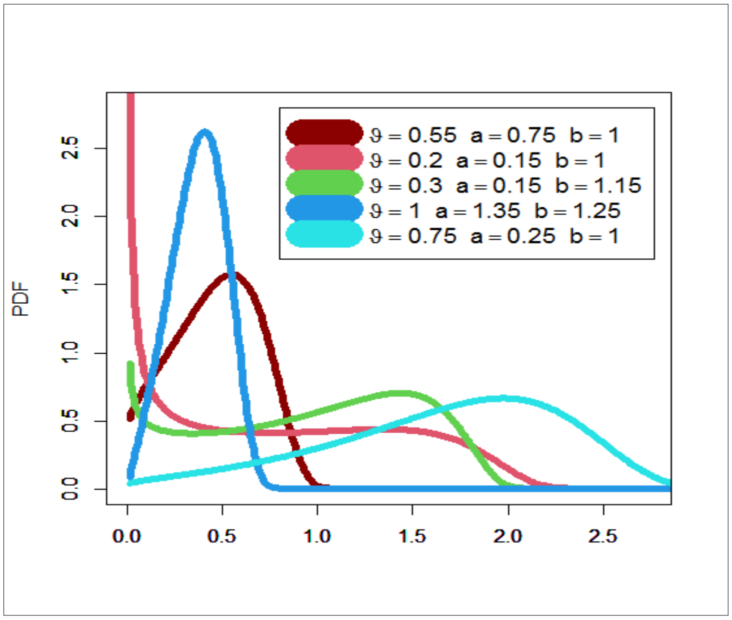

Figure 1: Various PDF plots for the new PDF.

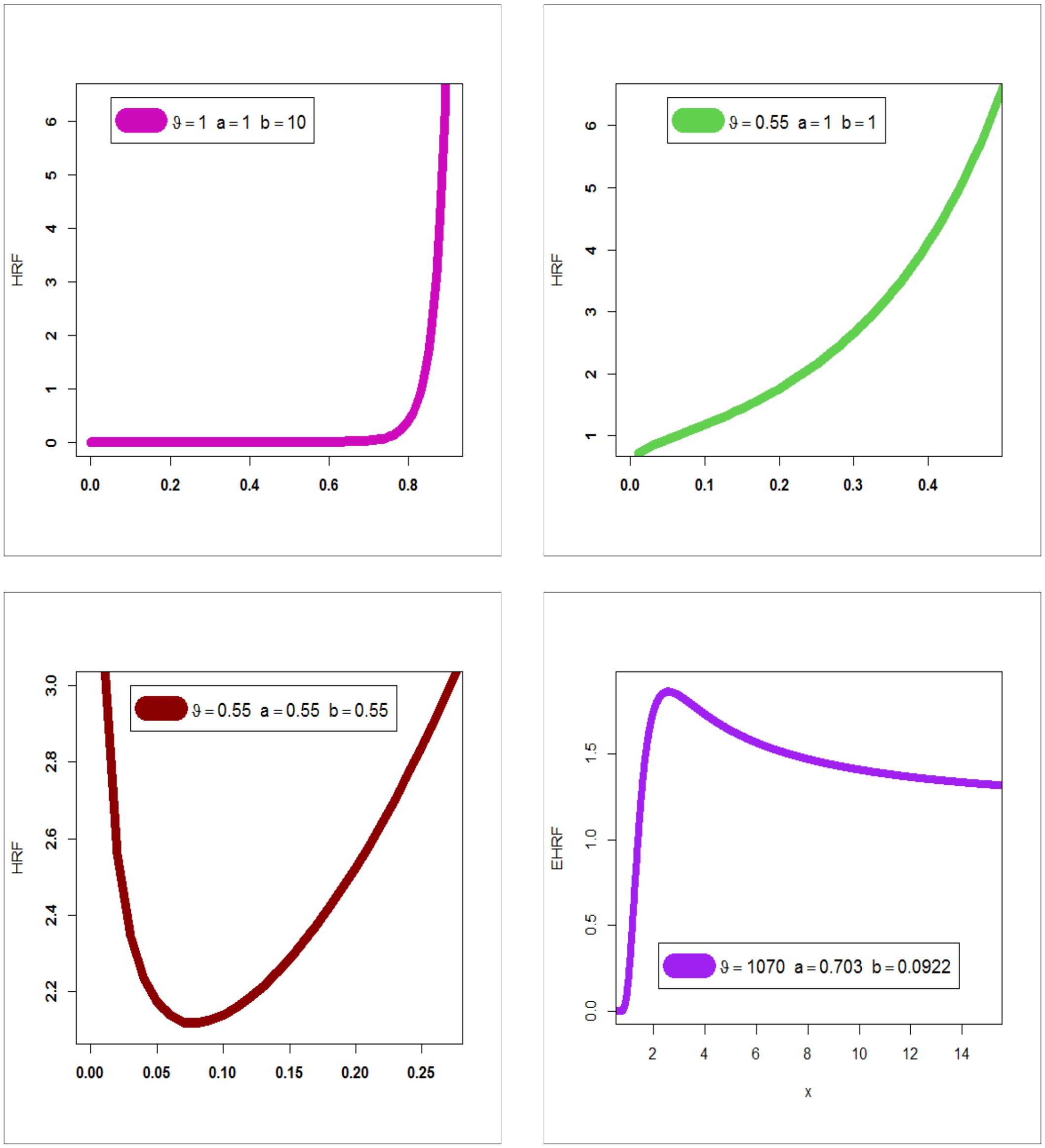

Figure 2: Some HRF plots.

2.Simple derivations for the new PDF

Consider the power series

$$\left(1-\frac{\xi_1}{\xi_2}\right)^{\xi_3} = \sum_{i_1=0}^{+\infty}\left(\frac{\xi_1}{\xi_2}\right)^{i_1}\frac{(-1)^{i_1}\Gamma(1+\xi_3)}{i_1!\,\Gamma(1+\xi_3-i_1)}\Big|_{\left|\frac{\xi_1}{\xi_2}\right|<1,\ \xi_3>0}. \tag{5}$$

Applying (5) to (4) we have

$$f_{\underline{\boldsymbol{\Psi}}}(x) = 2\vartheta a b x^{b-1}\, exp(x^b)\,\frac{1-\boldsymbol{q}_{a,b}(x)}{\boldsymbol{q}_{a,b}^2(x)}\sum_{i_1=0}^{+\infty}\frac{(-1)^{i_1}\Gamma(\vartheta)}{i_1!\,\Gamma(\vartheta-i_1)}exp\left[-(i_1+1)\boldsymbol{\mathcal{O}}_{a,b}^2(x)\right]. \tag{6}$$

Consider and apply the well-known power series to the quntity $exp\left[-(i_1+1)\boldsymbol{\mathcal{O}}_{a,b}^2(x)\right]$, equation (6) becomes

$$f_{\underline{\Psi}}(x) = 2\vartheta abx^{b-1}\sum_{i_1,i_2=0}^{+\infty}\frac{(-1)^{i_1+i_2}(i_1+1)^{i_2}\Gamma(\vartheta)\left[1-\boldsymbol{q}_{a,b}(x)\right]^{2i_2+1}}{i_1!\,i_2!\,\Gamma(\vartheta-i_1)\,exp(x^b)\,\boldsymbol{q}_{a,b}^{2i_2+2}(x)}. \quad (7)$$

Consider and apply the following series expansion

$$\left(1-\frac{\xi_1}{\xi_2}\right)^{-\xi_3} = \sum_{i_3=0}^{+\infty}\left(\frac{\xi_1}{\xi_2}\right)^{i_3}\frac{\Gamma(\xi_3+i_3)}{i_3!\,\Gamma(\xi_3)}\Big|_{\left|\frac{\xi_1}{\xi_2}\right|<1,\ \xi_3>0}. \quad (8)$$

Applying the expansion in (8) to (7) for the term $\boldsymbol{q}_{a,b}^{2i_2+3}(x)$, equation (7) becomes

$$f_{\underline{\Psi}}(x) = \sum_{i_2,i_3=0}^{+\infty}\boldsymbol{\zeta}_{i_2,i_3}\,\pi_{\boldsymbol{\varsigma}^{\cdot}}(x)\big|_{\boldsymbol{\varsigma}^{\cdot}=2i_2+i_3+2}, \quad (9)$$

where

$$\boldsymbol{\zeta}_{i_2,i_3} = \frac{2\vartheta(-1)^{i_2}\Gamma(\vartheta)\Gamma(2i_2+i_3+2)}{i_2!\,i_3!\,\Gamma(2i_2+2)\boldsymbol{\varsigma}^{\cdot}}\sum_{i_1=0}^{+\infty}(-1)^{i_1}\frac{(i_1+1)^{i_2}}{i_1!\,\Gamma(\vartheta-i_1)}$$

and

$$\pi_{\boldsymbol{\varsigma}^{\cdot}}(x) = \boldsymbol{\varsigma}^{\cdot}g_{a,b}(x)\left[G_{a,b}(x)\right]^{\boldsymbol{\varsigma}^{\cdot}-1}.$$

Equation (9) refers to that the new density can be simplified as a mixture of Exp-C densities. Similarly, the CDF of the BX-G family can also be expressed as a mixture of EC CDFs given by

$$F_{\underline{\Psi}}(x) = \sum_{i_2,i_3=0}^{+\infty}\boldsymbol{\zeta}_{i_2,i_3}\ \Pi_{\boldsymbol{\varsigma}^{\cdot}}(x), \quad (10)$$

where $\Pi_{\boldsymbol{\varsigma}^{\cdot}}(x) = \left[G_{a,b}(x)\right]^{\boldsymbol{\varsigma}^{\cdot}}$ is the CDF of the EC family with power parameter $\boldsymbol{\varsigma}^{\cdot}$. Due to Dey et al. (2017), the shape of the PDF of the Exp-C distribution may be characterized as follows: for $\boldsymbol{\varsigma}^{\cdot} < 1,\ 1 > b$, $g_{a,b}(x)$ is a decreasing density, for $1 < \boldsymbol{\varsigma}^{\cdot},\ b > 1$, $g_{a,b}(x) = dG_{a,b}(x)/dx$ is a unimodal density and for $\boldsymbol{\varsigma}^{\cdot} < 1, b > 1$ and $1 < \boldsymbol{\varsigma}^{\cdot}$, $1 > b$, $g_{a,b}(x)$ may be unimodal or decreasing density. Chaubey and Zhang (2015) presented a proof that the failure behavior of the Exp-C distribution are, respectively, bathtub ($\boldsymbol{\varsigma}^{\cdot} < 1$, $b < 1$), increasing ($1 < \boldsymbol{\varsigma}^{\cdot}$, $b > 1$), increasing or bathtub ($\boldsymbol{\varsigma}^{\cdot} < 1, b > 1$ and $1 < \boldsymbol{\varsigma}^{\cdot}, b < 1$).

3.Mathematical and statistical properties

Understanding probability distributions requires an understanding of mathematical properties. Mathematical functions called probability distributions describe the likelihood of various outcomes of a random variable. They are employed in a variety of fields, including finance, physics, engineering, and many more, and play a significant role in statistics and probability theory.

3.1 Moments and generating function

Numerous fields, including physics, engineering, economics, and finance use moments and incomplete moments as crucial statistical tools. They are mathematical operations that provide details about a probability distribution's form, location, and variability. Following Dey et al. (2017), we can extract the following two theorems:

***Theorem 1*:**

Let X be a RV having the Exp-C distribution with parameters a, b and power parameter $\varsigma\cdot$. Then using the transformation

$$t = \left[G_{a,b}(x)\right]^{\frac{1}{\varsigma^{\cdot}}},$$

the r^{th} ordinary moment of X is given by

$$\mu_r' = E[X^r] = b\sum_{p,\jmath=0}^{+\infty} \varsigma_p\left(\frac{r}{b}\right)\varsigma_\jmath\left(\frac{r}{b}+p\right)\frac{(-1)^{\frac{2r}{b}+p}}{a^{\frac{2r}{b}+p}[b(\varsigma+p+\jmath)+r]},$$

where $\varsigma_p\left(\frac{r}{b}\right)$ is the coefficient of $\left[\frac{1}{a}log(1-t)\right]^{\frac{2r}{b}+p}$ in the expansion of $\left\{\sum_{h_1=1}^{+\infty}\frac{1}{h_1}\left[\frac{1}{a}log(1-t)\right]\right\}^{\frac{r}{b}}$ and $\varsigma_\jmath\left(\frac{r}{b}+p\right)$ is the coefficient of $t^{p+\jmath+\frac{r}{b}}$ in the expansion of $\left(\sum_{h_2=1}^{+\infty}\frac{t^{h_2}}{h_2}\right)^{\frac{r}{b}+p}$ (see Balakrishnan and Cohen (2014) for more details).

***Theorem 2*:**
Let X be a RV having the Exp-C distribution. Then, the r^{th} conditional moment can be derived as

$$E(X^r|X>x) = \varsigma b\sum_{p,\jmath=0}^{+\infty}\frac{\varsigma_p\left(\frac{r}{b}\right)\varsigma_\jmath\left(\frac{r}{b}+p\right)(-1)^{\frac{2r}{b}+p}\,q_{a,b}(x)}{a^{\frac{2r}{b}+p}[b(\varsigma+p+\jmath)+r]\left\{1-\left[1-q_{a,b}(x)\right]^{\varsigma}\right\}}.$$

Based on Theorem 1, the r^{th} ordinary moment of the BXC model can then be expressed as

$$\mu_r' = E[X^r] = \varsigma^{\cdot} b\sum_{i_2,i_3,p,\jmath=0}^{+\infty}\zeta_{i_2,i_3}\,\varsigma_p^{\cdot}\left(\frac{r}{b}\right)\varsigma_\jmath^{\cdot}\left(\frac{r}{b}+p\right)\frac{(-1)^{\frac{2r}{b}+p}}{a^{\frac{2r}{b}+p}[b(\varsigma^{\cdot}+p+\jmath)+r]}. \quad (11)$$

In particular,

$$\mu_1' = E[X] = \varsigma^{\cdot} b\sum_{i_2,i_3,p,\jmath=0}^{+\infty}\zeta_{i_2,i_3}\,\varsigma_p^{\cdot}\left(\frac{1}{b}\right)\varsigma_\jmath^{\cdot}\left(\frac{1}{b}+p\right)\frac{(-1)^{\frac{2}{b}+p}}{a^{\frac{2}{b}+p}[b(\varsigma^{\cdot}+p+\jmath)+1]},$$

$$\mu_2' = E[X^2] = \varsigma^{\cdot} b\sum_{i_2,i_3,p,\jmath=0}^{+\infty}\zeta_{i_2,i_3}\,\varsigma_p^{\cdot}\left(\frac{2}{b}\right)\varsigma_\jmath^{\cdot}\left(\frac{2}{b}+p\right)\frac{(-1)^{\frac{2}{b}+p}}{a^{\frac{2}{b}+p}[b(\varsigma^{\cdot}+p+\jmath)+2]},$$

$$\mu_3' = E[X^3] = \varsigma^{\cdot} b\sum_{i_2,i_3,p,\jmath=0}^{+\infty}\zeta_{i_2,i_3}\,\varsigma_p^{\cdot}\left(\frac{3}{b}\right)\varsigma_\jmath^{\cdot}\left(\frac{3}{b}+p\right)\frac{(-1)^{\frac{6}{b}+p}}{a^{\frac{6}{b}+p}[b(\varsigma^{\cdot}+p+\jmath)+3]},$$

and

$$\mu_4' = E[X^4] = \varsigma^{\cdot} b\sum_{i_2,i_3,p,\jmath=0}^{+\infty}\zeta_{i_2,i_3}\,\varsigma_\jmath^{\cdot}\left(\frac{4}{b}\right)\varsigma_\jmath^{\cdot}\left(\frac{4}{b}+p\right)\frac{(-1)^{\frac{8}{b}+p}}{a^{\frac{8}{b}+p}[b(\varsigma^{\cdot}+p+\jmath)+4]}.$$

The variance (V(Y)), cumulants, n^{th} central moment, skewness (S(Y)), kurtosis (K(Y)) and Index of dispersion or the variance to mean ratio (ID(Y)) measures can be calculated from the ordinary moments using well-known relationships.

3.2 Conditional moments

For the increasing failure rate models, it is also of interest to know what $E(X^r|X>x)$ is. It can be easily seen that

$$E(X^r|X>x) = \varsigma^{\cdot} b\sum_{i_2,i_3,p,\jmath=0}^{+\infty}\frac{\zeta_{i_2,i_3}\,\varsigma_p^{\cdot}\left(\frac{r}{b}\right)\varsigma_\jmath^{\cdot}\left(\frac{r}{b}+p\right)(-1)^{\frac{2r}{b}+p}q_{a,b}(x)}{a^{\frac{2r}{b}+p}[b(\varsigma^{\cdot}+p+\jmath)+r]\left\{1-\left[1-q_{a,b}(x)\right]^{\varsigma^{\cdot}}\right\}}. \quad (12)$$

In particular,

$$E(X|X>x) = \varsigma^{\cdot} b\sum_{i_2,i_3,p,\jmath=0}^{+\infty}\frac{\zeta_{i_2,i_3}\,\varsigma_p^{\cdot}\left(\frac{1}{b}\right)\varsigma_\jmath^{\cdot}\left(\frac{1}{b}+p\right)(-1)^{\frac{2}{b}+p}q_{a,b}(x)}{a^{\frac{2}{b}+p}[b(\varsigma^{\cdot}+p+\jmath)+1]\left\{1-\left[1-q_{a,b}(x)\right]^{\varsigma^{\cdot}}\right\}},$$

$$E(X^2|X>x)=\varsigma^{\cdot} b\sum_{i_2,i_3,p,\jmath=0}^{+\infty}\frac{\zeta_{i_2,i_3}\,\varsigma_p^{\cdot}\left(\frac{2}{b}\right)\varsigma_{\jmath}^{\cdot}\left(\frac{2}{b}+p\right)(-1)^{\frac{4}{b}+p}\,q_{a,b}(x)}{a^{\frac{4}{b}+p}[b(\varsigma^{\cdot}+p+\jmath)+2]\left\{1-\left[1-q_{a,b}(x)\right]^{\varsigma^{\cdot}}\right\}},$$

$$E(X^3|X>x)=\varsigma^{\cdot} b\sum_{i_2,i_3,p,\jmath=0}^{+\infty}\frac{\zeta_{i_2,i_3}\,\varsigma_p^{\cdot}\left(\frac{3}{b}\right)\varsigma_{\jmath}^{\cdot}\left(\frac{3}{b}+p\right)(-1)^{\frac{3}{b}+p}\,q_{a,b}(x)}{a^{\frac{3}{b}+p}[b(\varsigma^{\cdot}+p+\jmath)+3]\left\{1-\left[1-q_{a,b}(x)\right]^{\varsigma^{\cdot}}\right\}},$$

and

$$E(X^4|X>x)=\varsigma^{\cdot} b\sum_{i_2,i_3,p,\jmath=0}^{+\infty}\frac{\zeta_{i_2,i_3}\,\varsigma_p^{\cdot}\left(\frac{4}{b}\right)\varsigma_{\jmath}^{\cdot}\left(\frac{4}{b}+p\right)(-1)^{\frac{8}{b}+p}\,q_{a,b}(x)}{a^{\frac{8}{b}+p}[b(\varsigma^{\cdot}+p+\jmath)+4]\left\{1-\left[1-q_{a,b}(x)\right]^{\varsigma^{\cdot}}\right\}}.$$

3.3 Mean residual lifetime

The residual life function, often referred to as the remaining life function, is a function that is used in dependability theory to characterize the likelihood that a product will fail after a specific amount of time has passed, assuming that it has lasted up to that point. The conditional survival probability, assuming the object has already survived up to a particular point, can be defined as the residual life function. The mean residual life (MRL) can be defined as

$$M_1=E(X-x|X>x)=\frac{1}{1-F_{\underline{\Psi}}(x)}\left[\int_x^{+\infty} y\,f_{\underline{\Psi}}(y)dy\right]-x.$$

Then using (12), we get

$$M_1=\varsigma^{\cdot} b\sum_{i_2,i_3,p,\jmath=0}^{+\infty}\frac{\zeta_{i_2,i_3}\,\varsigma_p^{\cdot}\left(\frac{1}{b}\right)\varsigma_{\jmath}^{\cdot}\left(\frac{1}{b}+p\right)(-1)^{\frac{2}{b}+p}\,q_{a,b}(x)}{a^{\frac{2}{b}+p}[b(\varsigma^{\cdot}+p+\jmath)+1]\left[1-\left(1-q_{a,b}(x)\right)^{\varsigma^{\cdot}}\right]}-x.$$

3.4 Mean past lifetime

The mean past lifetime (MPL) of the component can be defined as

$$P_1=E(x-X|X\le x)=x-\frac{1}{F_{\underline{\Psi}}(x)}\int_0^t y\,f_{\underline{\Psi}}(y)dy.$$

Then using (11) and (12), we get

$$P_1=x-\varsigma^{\cdot} b\sum_{i_2,i_3,p,\jmath=0}^{+\infty}\frac{\zeta_{i_2,i_3}\,\varsigma_p^{\cdot}\left(\frac{1}{b}\right)\varsigma_{\jmath}^{\cdot}\left(\frac{1}{b}+p\right)(-1)^{\frac{2}{b}+p}}{a^{\frac{2}{b}+p}[b(\varsigma^{\cdot}+p+\jmath)+1]}\left(1-q_{a,b}(x)\right)^{\frac{b(p+\jmath)+1}{b}}.$$

Finally, developing and evaluating statistical models requires a comprehension of these mathematical features. We build models that faithfully capture the behavior of real-life occurrences by employing probability distributions with well-defined mathematical features.

3.5 Numerical analysis for some properties

Skewness and kurtosis coefficients provide a numerical basis for comparing different datasets. By comparing the skewness and kurtosis values, you can assess the relative shape and distribution characteristics of multiple datasets. This analysis can be helpful in various fields, such as finance, economics, biology, and social sciences. It's important to note that while skewness and kurtosis provide valuable insights, they should be used in conjunction with other statistical measures and domain knowledge to gain a comprehensive understanding of the data. Additionally, the interpretation of skewness and kurtosis may vary depending on the specific context and the underlying data generating process. Table 1 below gives numerical results for the mean, variance (V(X)), skewness (S(X)), kurtosis (K(X)) and dispersion index (DisIx (X)). Based on Table 1, we note that:

i-The skewness of the BXC distribution can range in the interval (0.4287,4890.765).

iii-The spread for the BXC kurtosis is much larger ranging from 2.494433 to 26541861.

iii-ID (X) more than 1 which recommend the BXC model for modeling the over-dispersed real data sets.

Table 1: Mean, variance, skewness and kurtosis.

ϑ	a	b	E(X)	V(X)	S(X)	K(X)	ID(X)
0.01	0.35	0.5	0.951667	1641.974	100.4695	14275.09	1725.367

0.25			23.47723	40430.34	20.20585	579.3061	1722.109
5			383.6581	640166.7	4.892992	35.86464	1668.587
20			1080.718	1682322	2.749077	12.72226	1556.671
50			1872.575	2671425	1.889141	7.136960	1426.605
100			2641.222	3463667	1.354422	4.804903	1311.388
200			3495.955	4283805	0.800976	3.388077	1225.361
2	0.01	0.1	1.069857	7310.186	88.95714	8399.703	6832.862
	0.1		2469.187	4887737	1.132327	3.70872	1979.493
	0.2		138.4765	26766.7	2.404077	11.64097	193.2942
	0.3		17.96500	577.7038	2.885243	15.89334	32.15718
	0.4		3.646305	28.47126	3.291995	20.21016	7.808249
	0.5		0.969634	2.3134	3.648973	24.5676	2.385848
5	0.1	0.01	0.000363	1.965695	4890.765	26541861	5415.138
		0.05	1.693889	11429.94	70.58933	5303.069	6747.75
		0.10	3885.444	5786855	0.4287445	2.494433	1489.368
		0.125	826.0476	184714.0	0.9397281	4.199913	223.6118
		0.15	264.7239	13317.82	0.6932051	3.571194	50.30835

4.Copula

When modelling bivariate or multivariate data, the copula is a crucial statistical concept. The marginal distributions of two or more variables are connected to their joint distribution via this function. Copulas' adaptability and capacity to model intricate interrelationships between variables have made them more and more common in recent years. The following list highlights the significance and application of copulas in statistics and bivariate data modelling:

I. A fundamental topic in many branches of statistics and data science, dependence between variables is modelled using copulas. Copulas make it easier to represent complicated interdependencies between variables that are difficult to capture by straightforward correlation measurements because they preserve the marginal distributions of each variable while modelling the joint distribution of variables.

II. In finance, copulas are used to represent the relationship between the returns on various assets. This is crucial for portfolio optimization, which aims to build an asset mix that maximizes returns while lowering risk. We can more accurately evaluate a portfolio's risk and create more effective portfolios by utilizing copulas to model the interdependence between assets.

III. Copulas can be used to calculate the risk of severe events like market collapses or natural disasters and are also utilized in risk management. Copulas can give a more precise assessment of the likelihood that such events will occur by modelling the dependence structure between variables, which is crucial for risk management and insurance.

IV. Copulas can also be used to generate data, which is advantageous when gathering data is difficult or expensive. We can simulate new data sets with dependence structures identical to the original data, enabling us to produce fresh data for analysis and testing. To do this, we model the dependence structure between variables using a copula (see Morgenstern (1956), Gumbel (1958) and Gumbel (1960), Rodriguez-Lallena and Ubeda-Flores (2004)), Pougaza and Djafari (2011), Elgohari et al. (2021), Shehata and Yousof (2021, 2022), Shehata et al. (2021, 2022) and Elgohari and Yousof (2020a,b,c))

First, we will consider the joint CDF of the FGM family, where

$$H_\zeta(\nu,\tau)=\nu\tau(1+\zeta\nu^*\tau^*)|_{\nu^*=1-\nu,\tau^*=1-\tau},$$

and the marginal function $\nu=F_1$, $\tau=F_2$, $\zeta\in(-1,1)$ is a dependence parameter and for every $\nu,\tau\in(0,1)$, $H(\nu,0)=H(0,\tau)=0$ which is "grounded minimum" and $H(\nu,1)=\nu$ and $H(1,\tau)=\tau$ which is "grounded maximum", $H(\nu_1,\tau_1)+H(\nu_2,\tau_2)-H(\nu_1,\tau_2)-H(\nu_2,\tau_1)\geq 0$.

4.1 Via FGM family

A copula is continuous in ν and τ ; actually, it satisfies the stronger Lipschitz condition, where

$$|H(\nu_2,\tau_2)-H(\nu_1,\tau_1)|\le|\nu_2-\nu_1|+|\tau_2-\tau_1|.$$

For $0\le\nu_1\le\nu_2\le 1$ and $0\le\tau_1\le\tau_2\le 1$, we have

$$Pr(\nu_1\le\nu\le\nu_2,\tau_1\le\tau\le\tau_2)=H(\nu_1,\tau_1)+H(\nu_2,\tau_2)-H(\nu_1,\tau_2)-H(\nu_2,\tau_1)\ge 0.$$

Then, setting $\nu^*(w_1)=1-F_{\underline{\Psi}_1}(w_1)|_{[\nu^*=(1-\nu)\in(0,1)]}$ and $\tau^*(w_2)=1-F_{\underline{\Psi}_2}(w_2)|_{[\tau^*=(1-\tau)\in(0,1)]}$, we can then obtain the joint CDF of the BXC using the FGM family

$$H_\zeta(x,y)=\left\{1-exp\left[-O^2_{a_1,b_1}(x)\right]\right\}^{\vartheta_1}\left\{1-exp\left[-O^2_{a_2,b_2}(y)\right]\right\}^{\vartheta_2}$$
$$\times\left\{1+\zeta\left[\begin{matrix}\left(1-\left\{1-exp\left[-O^2_{a_1,b_1}(x)\right]\right\}^{\vartheta_1}\right)\\ \left(1-\left\{1-exp\left[-O^2_{a_2,b_2}(y)\right]\right\}^{\vartheta_2}\right)\end{matrix}\right]\right\}.$$

The joint PDF can then be derived from $h_\zeta(x,y)=1+\zeta\nu^*\tau^*|_{(\nu^*=1-2\nu\text{ and }\tau^*=1-2\tau)}$.

4.2 Via modified FGM family

As a generalization for the FGM, the modified FGM copula is given by

$$H_\zeta(\nu,\tau)=\nu\tau\left[1+\zeta B_{(\nu)}A_{(\tau)}\right]|_{\zeta\in(-1,1)}$$

or

$$H_\zeta(\nu,\tau)=\nu\tau+\zeta\varpi_\nu\dot{A}_\tau|_{\zeta\in(-1,1)},$$

where $\varpi_\nu=\nu B_{(\nu)}$, and $\dot{A}_\tau=\tau A_{(\tau)}$ and $B_{(\nu)}$ and $A_{(\tau)}$ are two continuous functions on $(0,1)$ with $B(0)=B(1)=A(0)=A(1)=0$.

Let

$$c_1(\varpi_\nu)=inf\left\{\varpi_\nu\ :\ \frac{\partial}{\partial\nu}\varpi_\nu\right\}|_{\Omega_{1,\nu}}<0,$$
$$c_2(\varpi_\nu)=sup\left\{\varpi_\nu\ :\ \frac{\partial}{\partial\nu}\varpi_\nu\right\}|_{\Omega_{1,\nu}}<0,$$
$$c_3(\dot{A}_\tau)=inf\left\{\dot{A}_\tau\ :\ \frac{\partial}{\partial\tau}\dot{A}_\tau\right\}|_{\Omega_{2,\tau}}>0,$$

and

$$c_4(\dot{A}_\tau)=sup\left\{\dot{A}_\tau\ :\ \frac{\partial}{\partial\tau}\dot{A}_\tau\right\}|_{\Omega_{2,\tau}}>0.$$

Then,

$$1\le min\left\{c_1(\varpi_\nu)c_2(\varpi_\nu),c_3(\dot{A}_\tau)c_4(\dot{A}_\tau)\right\}<\infty,$$

where

$$\nu\frac{\partial}{\partial\nu}B_{(\nu)}=\frac{\partial}{\partial\nu}\varpi_\nu-B_{(\nu)},$$

$$\Omega_{1,\nu}=\left\{\nu\ :\ \nu\in(0,1)|\frac{\partial}{\partial\nu}\varpi_\nu\ \text{ exists}\right\}$$

and

$$\Omega_{2,\tau}=\left\{\tau\ :\ \tau\in(0,1)|\frac{\partial}{\partial\tau}\dot{A}_\tau\ \text{ exists}\right\}.$$

Then, we can consider the following three type of the modified FGM copula.

TYPE-I

Consider the following functional form for both $B_{(\nu)}$ and $A_{(\tau)}$. Then, the B-BXC-FGM (Type-I) can be derived from

$$H_\zeta(x,y)=\left\{1-exp\left[-O^2_{a_1,b_1}(x)\right]\right\}^{\vartheta_1}\left\{1-exp\left[-O^2_{a_2,b_2}(y)\right]\right\}^{\vartheta_2}$$
$$+\zeta\left[\begin{matrix}\left\{1-exp\left[-O^2_{a_1,b_1}(x)\right]\right\}^{\vartheta_1}\left(1-\left\{1-exp\left[-O^2_{a_1,b_1}(x)\right]\right\}^{\vartheta_1}\right)\\ \left\{1-exp\left[-O^2_{a_2,b_2}(y)\right]\right\}^{\vartheta_2}\left(1-\left\{1-exp\left[-O^2_{a_2,b_2}(y)\right]\right\}^{\vartheta_2}\right)\end{matrix}\right]|_{\zeta\in(-1,1)}.$$

TYPE-II

Let $B_{(\nu)}$ and $A_{(\tau)}$ be two functional form satisfying all the conditions stated earlier where $B^*_{(\nu)}|_{(\zeta_1>0)} = \nu^{\zeta_1}(1-\nu)^{1-\zeta_1}$ and $A^*_{(\tau)}|_{(\zeta_2>0)} = \tau^{\zeta_2}(1-\tau)^{1-\zeta_2}$. Then, the corresponding B-BXC-FGM **(**Type-**II)** can be derived from $H_{\zeta,\zeta_1,\zeta_2}(\nu,\tau) = \nu\tau\left[1+\zeta B^*_{(\nu)}\,A^*_{(\tau)}\right]$. Thus

$$H_{\zeta,\zeta_1,\zeta_2}(x,y) = \left\{1-exp\left[-O^2_{a_1,b_1}(x)\right]\right\}^{\vartheta_1}\left\{1-exp\left[-O^2_{a_2,b_2}(y)\right]\right\}^{\vartheta_2} \times\left[1+\zeta\begin{pmatrix}\left\{1-exp\left[-O^2_{a_1,b_1}(x)\right]\right\}^{\vartheta_1\zeta_1}\\ \left\{1-exp\left[-O^2_{a_2,b_2}(y)\right]\right\}^{\vartheta_2\zeta_2}\\ \left(1-\left\{1-exp\left[-O^2_{a_1,b_1}(x)\right]\right\}^{\vartheta_1}\right)^{1-\zeta_1}\\ \left(1-\left\{1-exp\left[-O^2_{a_2,b_2}(y)\right]\right\}^{\vartheta_2}\right)^{1-\zeta_2}\end{pmatrix}\right].$$

TYPE-III

Let $\omega(\nu) = \nu[log(1+\nu^*)]$ and $\psi(\tau) = \tau[log(1+\tau^*)]$ for all $B_{(\nu)}$ and $A_{(\tau)}$ which satisfies all the conditions stated earlier. In this case, one can also derive a closed form expression for the associated CDF of the B-BXC-FGM (Type-**III**) from $H_\zeta(\nu,\tau) = \nu\tau\left(1+\zeta\omega(\nu)\,\psi(\tau)\right)$. Then

$$H_\zeta(x,y) = \left\{1-exp\left[-O^2_{a_1,b_1}(x)\right]\right\}^{\vartheta_1}\left\{1-exp\left[-O^2_{a_2,b_2}(y)\right]\right\}^{\vartheta_2} \times\left[1+\zeta\begin{pmatrix}\left\{1-exp\left[-O^2_{a_1,b_1}(x)\right]\right\}^{\vartheta_1}\\ \left\{1-exp\left[-O^2_{a_2,b_2}(y)\right]\right\}^{\vartheta_2}\\ \left[log\left(2-\left\{1-exp\left[-O^2_{a_1,b_1}(x)\right]\right\}^{\vartheta_1}\right)\right]\\ \left[log\left(2-\left\{1-exp\left[-O^2_{a_2,b_2}(y)\right]\right\}^{\vartheta_2}\right)\right]\end{pmatrix}\right].$$

4.3 B-BXC and M-BXC type via Clayton copula

The Clayton copula can be considered as $H(\tau_1,\tau_2) = [(1/\tau_1)^\zeta + (1/\tau_2)^\zeta - 1]^{-\zeta^{-1}}|_{\zeta\in(0,\infty)}$. Setting $\tau_1 = F_{\underline{\Psi}_1}(\nu)$ and $\tau_2 = F_{\underline{\Psi}_2}(w)$, the B-BXC type can be derived from $H(\tau_1,\tau_2) = H(F_{\underline{\Psi}_1}(\tau_1), F_{\underline{\Psi}_1}(\tau_2))$. Then

$$H(x,y) = \begin{pmatrix}\left\{1-exp\left[-O^2_{a_1,b_1}(x)\right]\right\}^{\vartheta_1}\\ +\left\{1-exp\left[-O^2_{a_2,b_2}(y)\right]\right\}^{\vartheta_2}\\ -1\end{pmatrix}^{-\zeta^{-1}}|_{\zeta\in(0,\infty)}$$

Similarly, the M-BXC can be derived from by generalizing the above result.

4.4 B-BXC type via Renyi's entropy

Using the theorem of Pougaza and Djafari (2011), we have

$$H(x,y) = yF_{\underline{\Psi}_1}(x) + xF_{\underline{\Psi}_2}(y) - xy,$$

the associated B-BXC can be derived from

$$H(x,y) = y\left\{1-exp\left[-O^2_{a_1,b_1}(x)\right]\right\}^{\vartheta_1} + x\left\{1-exp\left[-O^2_{a_2,b_2}(y)\right]\right\}^{\vartheta_2} - xy.$$

5. Parameter estimation

Once a distribution has been selected, its parameters can be calculated via maximum likelihood estimation using the data. This entails determining the parameter values that maximize the distribution-given probability of observing the data. The distribution can be used to draw conclusions about the population after the parameters have been estimated, such as determining survival probabilities, hazard rates, and median survival times. The method of maximum likelihood is the most frequently used method of parameter estimation. Its success stems from its many desirable properties including consistency, asymptotic efficiency, invariance property as well as its intuitive appeal. Let

 DOI: http://dx.doi.org/10.18187/pjsor.v19i2.3633

$x_1, \ldots, x_n$ be a random sample of size n from (4), then the log-likelihood function of (4) without constant terms is given by

$$\ell(\underline{\mathbf{\Psi}}) = n\, log\, 2 + n\, log\, \vartheta + n\, log\, a + n\, log\, b + (b-1)\sum_{i=1}^{n} log\, x_i + \sum_{i=1}^{n} x_i^b - \sum_{i=1}^{n} \boldsymbol{\mho}_{a,b}^{-2}(x_i)$$
$$-(1-\vartheta)\sum_{i=1}^{n}\{1 - exp[-\boldsymbol{\mho}_{a,b}^{-2}(x_i)]\} - 2\sum_{i=1}^{n} log[\boldsymbol{q}_{a,b}(x)] + \sum_{i=1}^{n} log[1 - \boldsymbol{q}_{a,b}(x_i)].$$

The components of the score vector, $U(\underline{\boldsymbol{\Psi}}) = \frac{\partial \ell(\underline{\boldsymbol{\Psi}})}{\partial \underline{\boldsymbol{\Psi}}} = \left(U(\vartheta), U(a), U(b)\right)^T$. The maximum likelihood estimates of the parameters ϑ, a and b are obtained by solving the following nonlinear systems of equations

$$\frac{\partial}{\partial \vartheta}\ell(\underline{\boldsymbol{\Psi}}) = 0, \frac{\partial}{\partial}\ell(\underline{\boldsymbol{\Psi}}) = 0, \frac{\partial}{\partial a}\ell(\underline{\boldsymbol{\Psi}}) = 0, \frac{\partial}{\partial b}\ell(\underline{\boldsymbol{\Psi}}) = 0.$$

6. Simulation study

Simulation studies are a common approach for assessing the performance of estimation methods in statistics. The use of simulation studies has become increasingly popular in recent years due to their ability to provide a controlled and rigorous evaluation of different estimation methods under various scenarios. In this context, this essay aims to highlight the statistical importance and motivations behind simulation studies for assessing estimation methods. Graphically and using the biases and mean squared errors (MSEs), we can perform simulation experiments to assess the finite sample behavior of the maximum likelihood estimations (MXLEs). The assessment was based on N =10000 replication for all $n|_{(n=200,400,\ldots,10000)}$. For a number of reasons, simulation studies are a crucial tool in statistical analysis.

I. By producing data under particular circumstances and evaluating the methods' accuracy in estimating the genuine parameters, simulation studies enable statisticians to assess the effectiveness of various statistical techniques. This gives a way to evaluate the approaches' reliability and applicability to various kinds of data.

II. By creating data that reflects the null or alternative hypothesis and comparing the outcomes with the observed data, simulation studies can be used to test hypotheses. This enables researchers to evaluate the strength of their tests and gauge the statistical significance of their findings.

III. To choose the right sample size for a study, simulation studies might be employed. Researchers can identify the sample size that will produce the most accurate and trustworthy results by producing data with various sample sizes and evaluating the effectiveness of various statistical methods.

The following algorithm is considered:

I. Generate N =1000 samples of size $n|_{(n=200,400,\ldots,10000)}$ from the BXC distribution using (4);

$$X_U = \left(ln\left\{1 - \frac{1}{a}\left[ln\left(1 - \left\{\left[-ln\left(1 - U^{\frac{1}{\vartheta}}\right)\right]^{-\frac{1}{2}} + 1\right\}^{-1}\right)\right]\right\}\right)^{\frac{1}{b}}$$

II. Compute the MXLEs for the N =10000 samples,

III. Compute the standard errors (SEs) of the MXLEs for the 10000 samples,

IV. The standard errors (SEs) were computed by inverting the observed information matrix.

V. Compute the biases and mean squared errors given for $\underline{\Theta} = \theta, a, b$. We repeated these steps for $n|_{(n=200,400,\ldots,10000)}$ with $\vartheta = 1,2,..,1000, b = 1,2,..,1000, a = 1,2,..,1000$, so computing biases $\left(\text{Bias}_{\underline{\Theta}}(n)\right)$, mean squared errors (MSEs) $(\text{MSE}_h(n))$ for $\underline{\Theta} = \vartheta, a, b$ and $n|_{(n=200,400,\ldots,10000)}$.

Figure 3 gives the biases (left) and MSEs (right) for the parameters ϑ, b and b respectively. The left plots from show how the three biases vary with respect to the sample size n . The right plots show how the four MSEs vary with respect to sample size n. The broken line in red in Figure 3 (left plots) corresponds to the biases being 0 . From Figure 3 (left plots), the biases for each parameter tends to zero as $n \to \infty$. From Figure 3 (right plots), the MSEs for each parameter decrease to zero as $n \to \infty$.

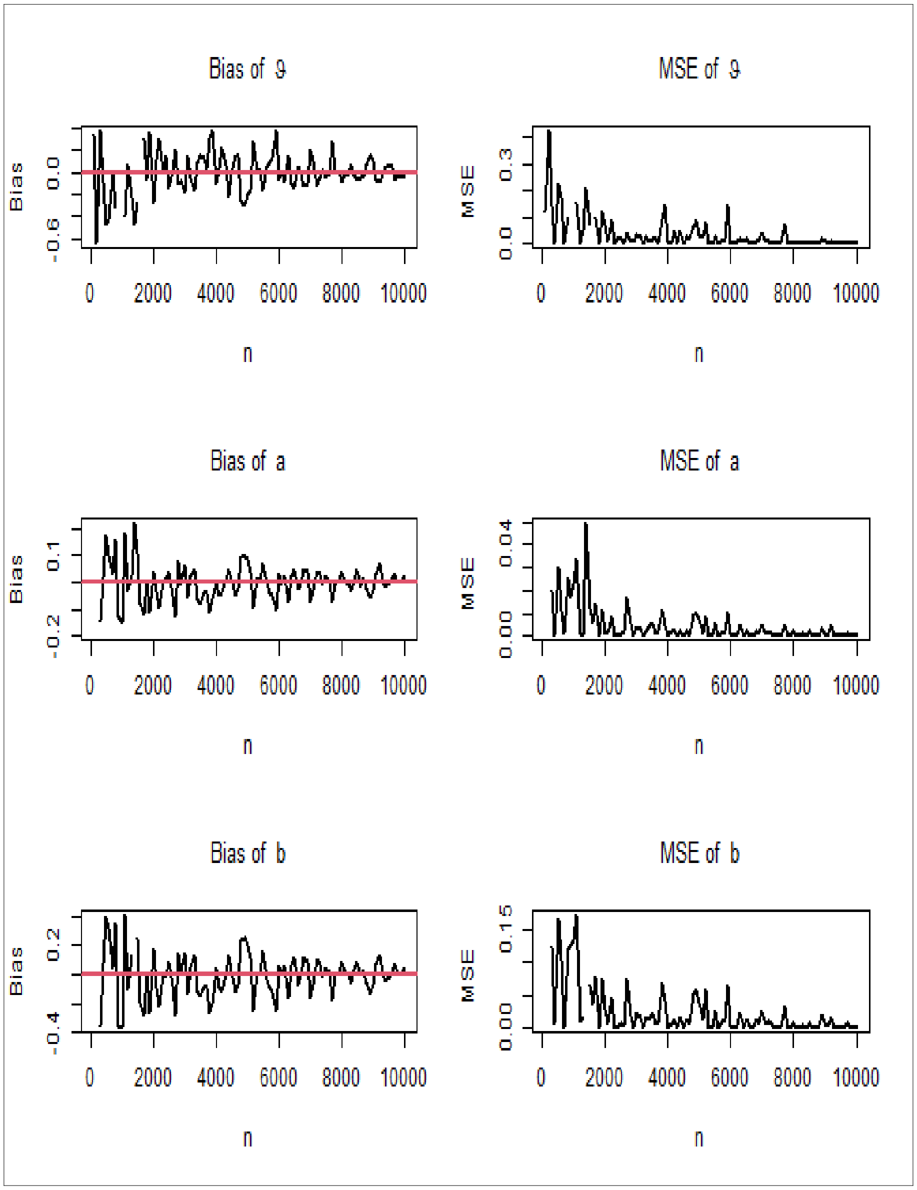

Figure 3: Biases (left) and MSEs (right) for the parameters ϑ, a, b.

7. Applications

In statistics, comparing competing probability distributions is an important task that helps researchers to understand the underlying data generating process and to make inferences about the population parameters. However, the choice of the appropriate probability distribution for a given dataset can be challenging, especially when there are several competing distributions. In this context, the importance of real data modeling for comparing different probability distributions cannot be overstated.

The Chen distribution has been used in quality control to model the distribution of defects in a manufacturing process. Its ability to model non-normal and skewed data makes it useful in situations where the defect data is not normally distributed. This distribution can be used to estimate the process capability and to optimize the manufacturing process. The Chen distribution has been used in engineering applications such as fatigue analysis, where it can be used to model the number of cycles to failure of a component subjected to cyclic loading. Its flexibility and efficiency in generating random variates make it useful in simulations and other computational applications. The Chen distribution has been used in reliability engineering to model the time-to-failure of components in a system. Its ability to model non-normal and skewed data is useful in situations where the failure data is not normally distributed. This distribution can be used to estimate the reliability and availability of a system and to optimize its maintenance. The Chen distribution has been applied to engineering tasks like fatigue analysis, where it can be used to predict how many cycles will pass before a component that has been subjected to cyclic loading will fail. It is helpful in simulations and other computer applications due to its adaptability and effectiveness in producing random variations. In reliability engineering, the Chen distribution has been used to simulate the time to failure of system components. When the failure data is not regularly distributed, it can be helpful because of its capacity to represent skewed and non-normal data. This distribution can be used to predict a system's availability and reliability and to plan for the best possible maintenance.

In this section, two real data sets are considered and analyzed to show the applicability of the BXC model. The new model is compared with many common competitive models under the Akaike Information Criterion (AICR), the Bayesian Information Criterion (BICR), the Cramér-von Mises (CVMS), the Anderson--Darling (AD) and Kolmogorov--Smirnov (KS) (corresponding p-value). The following competitive models are considered: the Gamma-Chen distribution (GC) (Alzaatreh et al. (2014)), the Beta-Chen distribution (BC) (Eugene et al. (2002)), Marshall-Olkin Chen distribution (MOC) (Jose (2011)), the Kumaraswamy Chen distribution (KC) (Cordeiro and de Castro (2011)), the Transmuted Chen (TC) (Khan et al. (2013)), the Transmuted Exponentiated Chen (TEC) (Khan et al. (2016)), the Extended Chen (EC) and the standard Chen distribution have been selected for comparison in three examples. Some other useful Chen extension can be found in Yousof et al. (2022), Korkmaz et al. (2022), Ibrahim et al. (2022) and Ali et al. (2022). The parameters of models have been estimated by the MXLE method. Moreover, Chen extension and its generalizations have many applications in insurance and actuarial sciences, see for example, Mohamed et al. (2022a,b,c), Yousof et al. (2023), Emam et al. (2023a,b) and Ibrahim et al. (2023).

Data 1: The relief times of twenty patient’s data

This subsection is allocated for studying the data set of Gross and Clark (1975) on the relief times of twenty patients receiving an analgesic: 1.1, 1.40, 1.3, 1.70, 1.9, 1.80, 1.6, 2.20, 1.7, 2.70, 4.1, 1.80, 1.5, 1.20, 1.40, 3, 1.7, 2.3, 1.60, 2.00. Figure 4 gives the box plot, quantile-quantile (Q-Q) plot, TTT plot and NKDE plot for the relief times data. The box plot shows that relief times data has one extreme value. The Q-Q plots ensure the results obtained by the box plot. The dashed line in all the Q-Q plots refers to the safe boundaries for the standard errors. The TTT plot shows that the HRFs for the relief times data is "monotonically increasing". The NKDE plot shows that the KDE is "asymmetric bimodal density" with right tail. Estimated PDF (EPDF) estimated CDF(ECDF), probability-probability (P-P), estimate HRF (EHRF) and Kaplan-Meier plots for the relief times data are listed in Figure 5. Table 2 lists statistics for comparing model under the relief times data. Based on Table 2, it is noted that the BXC model provides the best results: AICR=37.51, BICR=40.49, CVMS=0.040, AD=0.232, K. S=0.116 and P-Value=0.9524. Table 3 gives the MXLEs (SEs) for the relief times data.

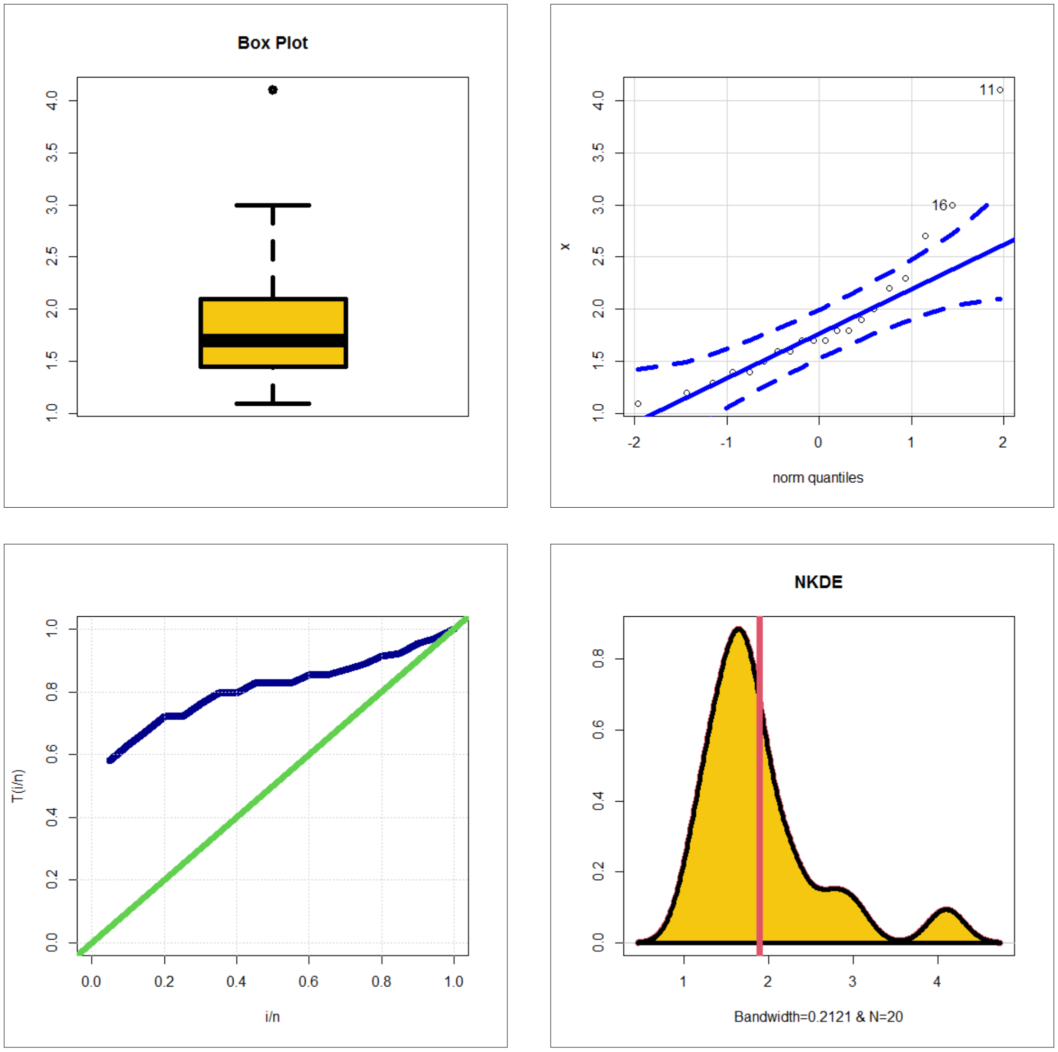

Figure 4: Box plot, Q-Q plot, TTT plot and NKDE plot for the relief times data.

Table 2: Results for comparing competing model under the relief times data.

Model	AICR	BICR	CVMS	AD	K.S	P-Value
BXC	37.511	40.49	0.040	0.232	0.116	0.9524
GC	46.396	51.36	0.049	0.288	0.998	<0.010
EC	38.156	41.18	0.057	0.305	0.138	0.8644
KC	41.069	44.03	0.058	0.310	0.149	0.8208
TC	53.693	56.65	0.279	1.579	0.234	0.2439
TEC	39.584	43.57	0.047	0.289	0.129	0.9493
MOC	45.821	47.89	0.147	0.849	0.159	0.7743
BC	41.545	44.47	0.067	0.347	0.157	0.7692
Chen	54.133	55.14	0.290	1.663	0.240	0.2066

Table 3: The MXLEs (SEs) for the relief times data.

Model	MXLE (Standard errors)			
BXC(ϑ,a,b)	1.097×10^{3}	7.029×10^{-2}	9.22×10^{-2}	
	(1.6×10^{3})	(5.3×10^{-2})	(2.4×10^{-2})	
GC(ϑ,λ,a,b)	7.59144	1.98813	5.00233	0.5349
	(2.0955)	(0.4653)	(1.0743)	(0.0035)
TEC(ϑ,λ,a,b)	300.0144	0.50233	2.43029	0.34435
	(587.043)	(0.5643)	(1.0888)	(0.1190)
KC(ϑ,λ,a,b)	160.074	0.49432	2.2143	0.5200
	(222.41)	(0.5198)	(0.7556)	(0.2149)
BC(ϑ,λ,a,b)	85.874	0.4812	2.0134	0.5555
	(103.130)	(0.5132)	(0.699)	(0.2020)
EC(ϑ,a,b)	250.0143	2.4010	0.3743	
	(407.52)	(0.890)	(0.1001)	
MOC(ϑ,a,b)	400.014	2.3202	0.4332	
	(488.061)	(0.6399)	(0.089)	
TC(ϑ,a,b)	0.74555	0.07142	1.0223	
	(0.28403)	(0.03434)	(0.090)	
Chen(a,b)	0.13880	0.94534		
	(0.05111)	(0.09435)		

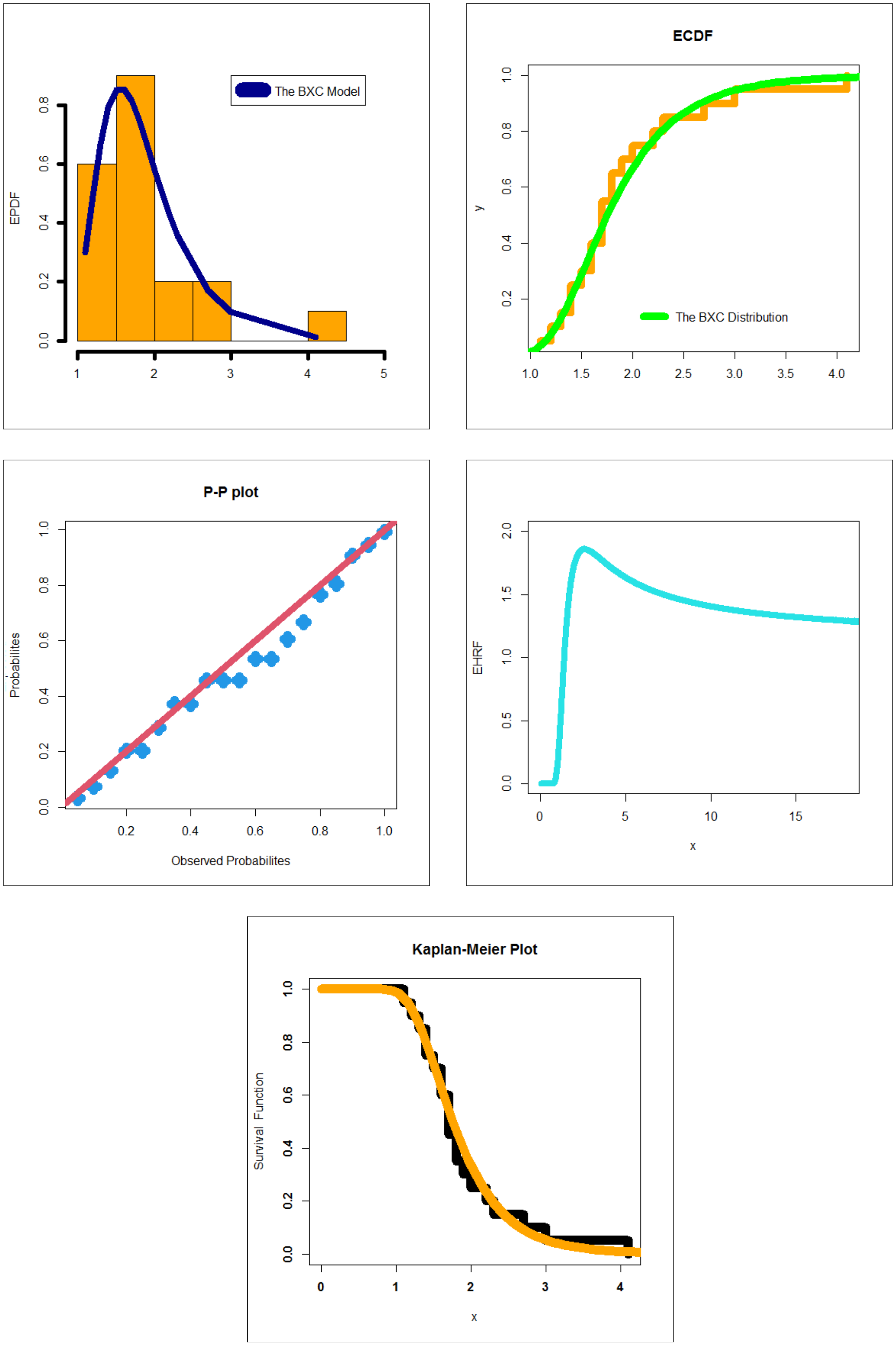

Figure 5: EPDF, ECDF, P-P, EHRF and Kaplan-Meier plot for the relief times data.

Data 2: The Minimum Flow data

This subsection is related to the study of minimum flow data which was presented by Cordeiro et al. (2012) that include 38 observations. The data set is the following: 43.860, 44.970, 46.27, 51.290, 61.190, 61.20, 67.80, 69.00, 71.840, 77.310,85.39, 86.590, 86.66, 88.160, 96.030, 102.00, 108.29, 113.00, 115.14, 116.71, 126.86, 127.00, 127.14, 127.29, 128.00, 134.140, 136.140, 140.430, 146.430, 146.43,148.00, 148.43, 150.86, 151.29, 151.43, 156.14, 163.00, 186.43. Like the previous application examples, we have Tables 5 and 6. As is clear, the BXC is selected as the best model with all criteria. Figure 6 gives the box plot, Q-Q plot, TTT plot, NKDE plot and Kaplan-Meier plot for the minimum flow data. The box plot shows that minimum flow data has one extreme value. The Q-Q plots ensure the results obtained by the box plot. The TTT plot shows that the HRFs for the relief times data is "monotonically increasing". The NKDE plot shows that the KDE of minimum flow data is asymmetric bimodal" density. Table 4 gives statistics for comparing model under the minimum flow data. Table 5 lists the MXLEs (and their corresponding standard errors (SEs)) for the minimum flow data. Based on Table 2, it is noted that the BXC model provides the best results: AICR=389.44, BICR=394.35, CVMS=0.05, AD=0.44, K. S=0.11 and P-Value=0.762. EPDF, ECDF, P-P and EHRF plots for the minimum flow data are displayed in Figure 7.

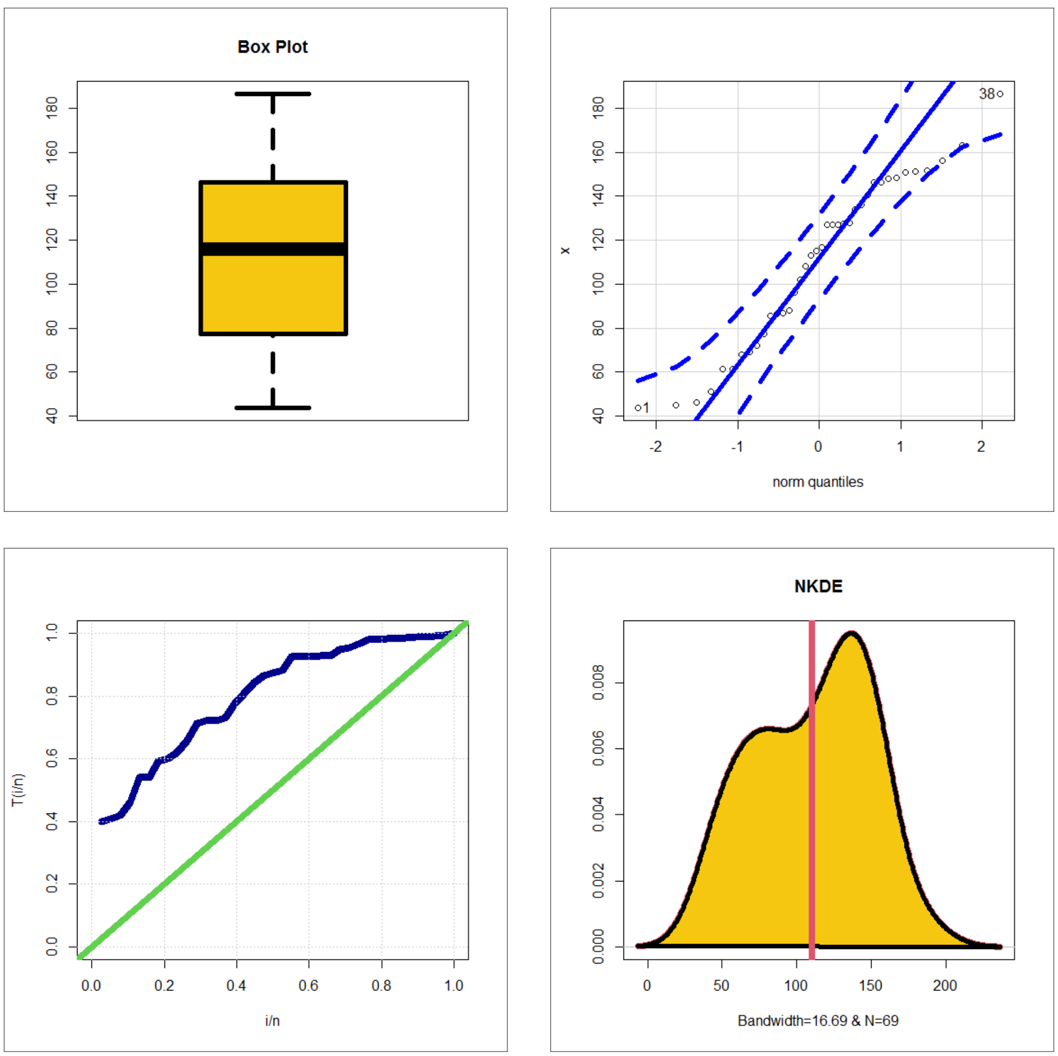

Figure 6: Box plot, Q-Q plot, TTT plot and NKDE plot for the minimum flow data.

Table 4: Results for comparing competing model under the minimum flow data.

Model	AICR	BICR	CVMS	AD	K.S	P-Value
BXC	**389.44**	**394.35**	**0.051**	**0.444**	**0.1101**	**0.7621**
MOC	391.57	395.49	0.098	0.615	0.1277	0.6890
Chen	397.63	400.91	0.102	0.646	0.1686	0.2624
KC	392.27	396.78	0.115	0.667	0.1441	0.4090
TC	388.51	393.46	0.198	0.661	0.1553	0.3839
TEC	392.70	399.28	0.114	0.702	0.1578	0.3771
EC	389.99	395.80	0.133	0.729	0.1533	0.3486
BC	394.25	397.77	0.125	0.753	0.1519	0.3436
GC	392.49	399.04	0.288	1.715	0.5791	<0.010

Table 5: The MXLEs (SEs) for the minimum flow data.

Model	MXLE (Standard errors)			
BXC(ϑ,a,b)	0.5443	0.0017	0.3596	
	(0.094)	(0.0013)	(0.0041)	
TEC(ϑ,λ,a,b)	2.7372	-0.2488	0.0132	0.3488
	(1.216)	(0.478)	(0.013)	(0.029)
BC(ϑ,λ,a,b)	3.0144	0.7734	0.0144	0.3544
	(1.909)	(1.249)	(0.011)	(0.055)
GC(ϑ,λ,a,b)	3.1352	4.3649	0.0965	0.3455
	(1.143)	(4.437)	(0.025)	(0.022)
KC(ϑ,λ,a,b)	4.5144	21.111	0.0222	0.2735
	(2.022)	(42.855)	(0.024)	(0.053)
MOC(ϑ,a,b)	13.0010	0.0232	0.3455	
	(18.667)	(0.026)	(0.048)	
EC(ϑ,a,b)	2.8593	0.01447	0.3554	
	(0.987)	(0.004)	(0.028)	
TC(ϑ,a,b)	-1.0044	0.0039	0.3685	
	(0.707)	(0.0024)	(0.017)	
Chen(a,b)	0.00324	0.3655		
	(0.0019)	(0.019)		

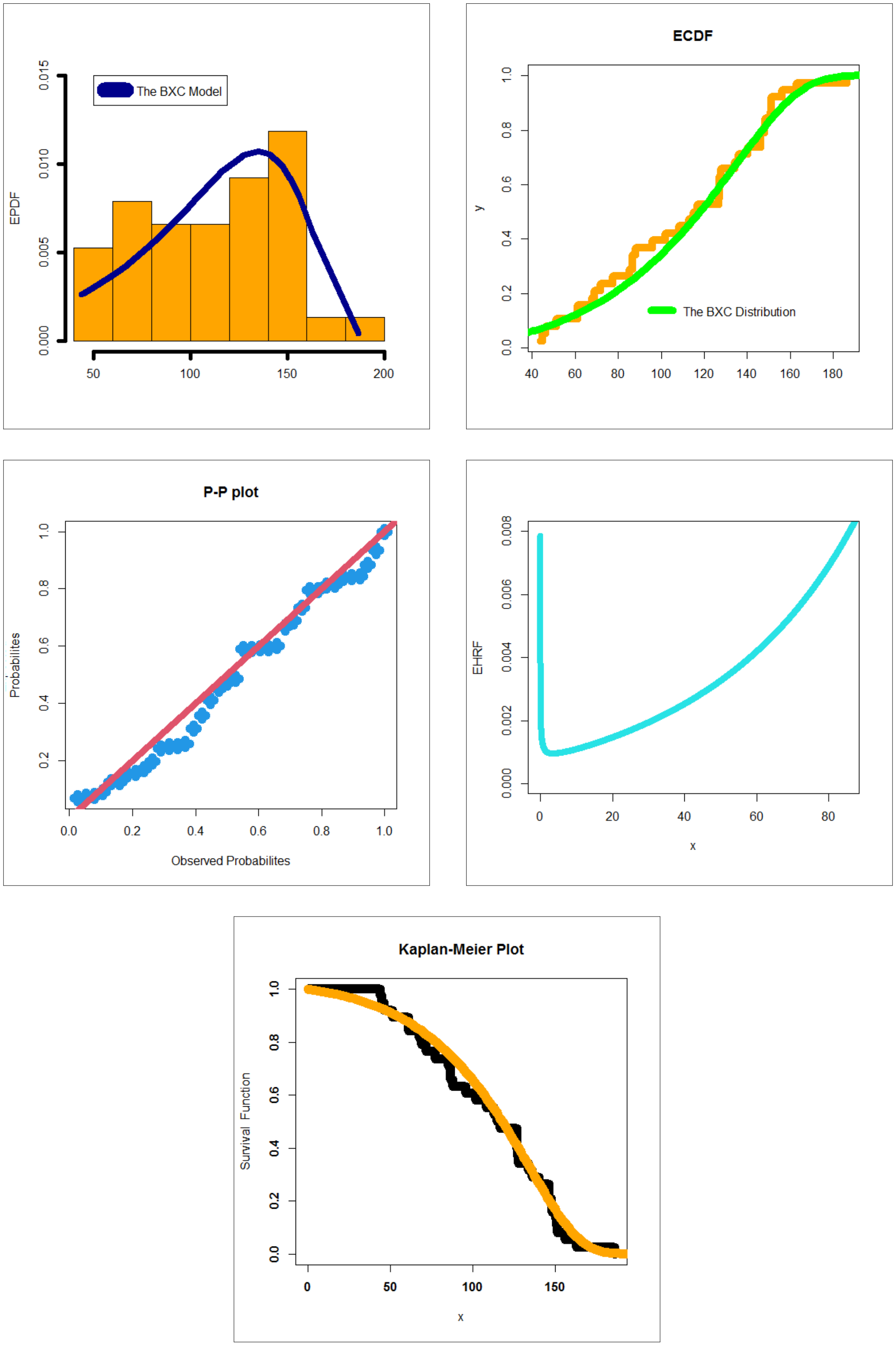

Figure 7: EPDF, ECDF, P-P, EHRF and Kaplan-Meier plot for the minimum flow data.

Table 6 gives the mean, variance, skewness and kurtosis using the estimated parameters. For the relief times data $\mathrm{S}(X) = 0.431397 > 0$, $\mathrm{K}(X) = 3.282876 > 3$ and $1 > \mathrm{ID}(X) > 0$. For the minimum flow data $\mathrm{S}(X) = 8.22105 > 0$, $\mathrm{K}(X) = 72.80223 > 3$ and $\mathrm{ID}(X) > 1$.

Data availability: The data set is available in the paper.
Funding: This research received no external funding.
Conflicts of Interest: The authors declare no conflict of interest.

8. Conclusions

A new flexible Chen extension which accommodates many useful hazard rate shapes such as the "**J**", "monotonically increasing", "**U** (bathtub)" and "upside down (reversed bathtub)" hazard rate shapes. The new model is derived based on Burr type X G family and called the Burr type X Chen (BXC) model. Moments, conditional moments, moment generating function, mean residual life, and mean past lifetime are among the pertinent statistical features of the BXC model that are obtained. Many common copulas, such the Farlie-Gumbel-Morgenstern, modified FGM copula, Renyi's entropy, and Clayton copula, are used to create some bivariate BXC type distributions. For calculating the BXC model parameters, we took into account the maximum likelihood estimation approach. Maximal likelihood estimators are evaluated via graphical simulations. To demonstrate the BXC model's applicability, two actual data sets are taken into account and examined.

The Beta-Chen distribution, Gamma-Chen distribution, Marshall-Olkin Chen distribution, Transmuted Chen, transmuted exponentiated Chen, Extended Chen, Kumaraswamy Chen distribution, and Standard Chen distribution are just a few examples of common competitive models that the BXC model is compared to. The Akaike Information criterion, Bayesian Information criterion, Cramer-von Mises criterion, Anderson-Darling criterion, Kolmogorov-Smirnov test statistic, and its associated p-value are all used to do the comparison.

The BXC distribution can be used to model the distribution of financial risk, which is important for estimating the probability of financial losses for individuals and organizations. Its flexibility in modeling non-normal and skewed data is useful for analyzing complex financial data. The BXC distribution can be used to model the distribution of investment returns, which is important for estimating the risk and return of investment portfolios. Its ability to model non-normal and skewed data is useful for analyzing investment return data that may not follow a normal distribution. For the purpose of calculating the expected value of claims and the risk of losses for insurance firms, it is critical to model the distribution of insurance claims using the BXC distribution. It is helpful for analyzing complicated insurance claim data because of its adaptability in modelling non-normal and skewed data.

References

1. Ali, M. M., Ibrahim, M. and Yousof, H. M. (2022). A New Flexible Three-Parameter Compound Chen Distribution: Properties, Copula and Modeling Relief Times and Minimum Flow Data. Bulletin of the Malaysian Mathematical Sciences Society, 45(1), 130-160.
2. Alzaatreh, A., Famoye, F. and Lee, C. (2014). The gamma-normal distribution: Properties and applications. Computational Statistics and Data Analysis, 69, 67-80.
3. Balakrishnan, N., Cohen, A. C. (2014). Order Statistics & Inference: Estimation Methods. Academic Press, London: Elsevier.
4. Chaubey, Y. P., Zhang, R. (2015). An extension of Chen's family of survival distributions with bathtub shape or increasing hazard rate function. Communications in Statistics-Theory and Methods 44(19), 4049--4064.
5. Chen, Z. (2000). A new two-parameter lifetime distribution with bathtub shape or increasing failure rate function. Statistics & Probability Letters 49(2), 155--161.
6. Cordeiro, G. M. and de Castro, M. (2011). A new family of generalized distributions. Journal of Statistical Computation and Simulation, 81, 883-893.
7. Dey, S., Kumar, D., Ramos, P. L. and Louzada, F. (2017). Exponentiated Chen distribution: Properties and estimation. Communications in Statistics-Simulation and Computation, 46(10), 8118-8139.
8. Elgohari, H. and Yousof, H. M. (2020a). A Generalization of Lomax Distribution with Properties, Copula and Real Data Applications. Pakistan Journal of Statistics and Operation Research, 16(4), 697-711. https://doi.org/10.18187/pjsor.v16i4.3260
9. Elgohari, H. and Yousof, H. M. (2021b). A New Extreme Value Model with Different Copula, Statistical

Properties and Applications. Pakistan Journal of Statistics and Operation Research, 17(4), 1015-1035. https://doi.org/10.18187/pjsor.v17i4.3471

10. Elgohari, H. and Yousof, H. M. (2020c). New Extension of Weibull Distribution: Copula, Mathematical Properties and Data Modeling. Statistics, Optimization & Information Computing, 8(4), 972-993. https://doi.org/10.19139/soic-2310-5070-1036
11. Elgohari, H., Ibrahim, M. and Yousof, H. M. (2021). A New Probability Distribution for Modeling Failure and Service Times: Properties, Copulas and Various Estimation Methods. Statistics, Optimization & Information Computing, 8(3), 555-586.
12. Emam, W.; Tashkandy, Y.; Goual, H.; Hamida, T.; Hiba, A.; Ali, M.M.; Yousof, H.M.; Ibrahim, M. A New One-Parameter Distribution for Right Censored Bayesian and Non-Bayesian Distributional Validation under Various Estimation Methods. Mathematics 2023a, 11, 897. https://doi.org/10.3390/math11040897
13. Emam, W.; Tashkandy, Y.; Hamedani, G.G.; Shehab, M.A.; Ibrahim, M.; Yousof, H.M. A Novel Discrete Generator with Modeling Engineering, Agricultural and Medical Count and Zero-Inflated Real Data with Bayesian, and Non-Bayesian Inference. Mathematics 2023b, 11, 1125. https://doi.org/10.3390/math11051125
14. Eugene, N., Lee, C., and Famoye, F. (2002). Beta-Normal distribution and its Applications. Communications in Statistics -Theory and Methods, 31, 497-512.
15. Farlie, D. J. G. (1960) The performance of some correlation coefficients for a general bivariate distribution. Biometrika, 47, 307-323.
16. Gumbel, E. J. (1961). Bivariate logistic distributions. Journal of the American Statistical Association, 56(294), 335-349.
17. Gumbel, E. J. (1960) Bivariate exponential distributions. Journ. Amer. Statist. Assoc., 55, 698-707.
18. Ibrahim, M., Aidi, K., Ali, M. M. and Yousof, H. M. (2022). A Novel Test Statistic for Right Censored Validity under a new Chen extension with Applications in Reliability and Medicine. Annals of Data Science, forthcoming. doi.org/10.1007/s40745-022-00416-6
19. Ibrahim, M.; Emam, W.; Tashkandy, Y.; Ali, M.M.; Yousof, H.M. Bayesian and Non-Bayesian Risk Analysis and Assessment under Left-Skewed Insurance Data and a Novel Compound Reciprocal Rayleigh Extension. Mathematics 2023, 11, 1593. https://doi.org/10.3390/ math11071593
20. Johnson, N. L. and Kotz, S. (1975) On some generalized Farlie-Gumbel-Morgenstern distributions. Commun. Stat. Theory, 4, 415-427.
21. Johnson, N. L. and Kotz, S. (1977) On some generalized Farlie-Gumbel-Morgenstern distributions- II: Regression, correlation and further generalizations. Commun. Stat.Theory, 6, 485-496.
22. Johnson, N. L., Kemp, A. W. and Kotz, S. (2005). Univariate discrete distributions, 3rd edn. Wiley, Hoboken.
23. Jose, K. K. (2011). Marshall-Olkin family of distributions and their applications in reliability theory, time series modeling and stress-strength analysis. Proc. ISI 58th World Statist. Congr Int Stat Inst, 21st-26th August, 3918-3923.
24. Khan, M. S., King, R. and Hudson, I. L. (2013). A new three parameter transmuted Chen lifetime distribution with application. Journal of Applied Statistical Science, 21, 239-259.
25. Khan, M. S., King, R. and Hudson, I. L. (2016). Transmuted exponentiated Chen distribution with application to survival data. ANZIAM Journal, 57, 268-290.
26. Korkmaz, M. Ç., Altun, E., Chesneau, C. and Yousof, H. M. (2022). On the unit-Chen distribution with associated quantile regression and applications. Mathematica Slovaca, 72 (2022), No. 3, 765-786.
27. Mohamed, H. S., Ali, M. M. and Yousof, H. M. (2022a). The Lindley Gompertz Model for Estimating the Survival Rates: Properties and Applications in Insurance, Annals of Data Science, 10.1007/s40745-022-00451-3
28. Mohamed, H. S., Cordeiro, G. M. and Yousof, H. M. (2022b). The synthetic autoregressive model for the insurance claims payment data: modeling and future prediction. Statistics, Optimization & Information Computing, forthcoming.
29. Mohamed, H. S., Cordeiro, G. M., Minkah, R., Yousof, H. M. and Ibrahim, M. (2022c). A size-of-loss model for the negatively skewed insurance claims data: applications, risk analysis using different methods and statistical forecasting. Journal of Applied Statistics, forthcoming.
30. Morgenstern, D. (1956). Einfache beispiele zweidimensionaler verteilungen. Mitteilingsblatt fur Mathematische Statistik, 8, 234-235.
31. Pougaza, D. B. and Djafari, M. A. (2011). Maximum entropies copulas. Proceedings of the 30th international workshop on Bayesian inference and maximum Entropy methods in Science and Engineering, 329-336.
32. Rodriguez-Lallena, J. A. and Ubeda-Flores, M. (2004). A new class of bivariate copulas. Statistics and Probability Letters, 66, 315--25.
33. Shehata, W. A. M. and Yousof, H. M. (2022). A novel two-parameter Nadarajah-Haghighi extension: properties,

copulas, modeling real data and different estimation methods. Statistics, Optimization & Information Computing, 10(3), 725-749.

34. Shehata, W. A. M. and Yousof, H. M. (2021). The four-parameter exponentiated Weibull model with Copula, properties and real data modeling. Pakistan Journal of Statistics and Operation Research, 17(3), 649-667.
35. Shehata, W. A. M., Yousof, H. M., & Aboraya, M. (2021). A Novel Generator of Continuous Probability Distributions for the Asymmetric Left-skewed Bimodal Real-life Data with Properties and Copulas . Pakistan Journal of Statistics and Operation Research, 17(4), 943-961. https://doi.org/10.18187/pjsor.v17i4.3903
36. Shehata, W. A. M., Butt, N. S., Yousof, H., & Aboraya, M. (2022). A New Lifetime Parametric Model for the Survival and Relief Times with Copulas and Properties. Pakistan Journal of Statistics and Operation Research, 18(1), 249-272.
37. Yousof, H., Afify, A., Hamedani, G. G., and Aryal, G. (2017). The Burr X Generator of Distributions for Lifetime Data. Journal of Statistical Theory and Applications, 16: 288-305.
38. Yousof, H. M., Emam, W., Tashkandy, Y., Ali, M. M., Minkah, R. and Ibrahim, M. (2023). A Novel Model for Quantitative Risk Assessment under Claim-Size Data with Bimodal and Symmetric Data Modeling. Mathematics, 11, 1284. https://doi.org/10.3390/math11061284
39. Yousof, H. M., Korkmaz, M. Ç., K., Hamedani, G. G and Ibrahim, M. (2022). A novel Chen extension: theory, characterizations and different estimation methods. Eur. J. Stat, 2(2022), 1-20.