Nepal Journal of Mathematical Sciences (NJMS)
ISSN: **2738-9928 (*online*), 2738-9812 (*print*)**
Vol. 3, No. 1, 2022 (February): 11-30
DOI: https://doi.org/10.3126/njmathsci.v3i1.44100

Research Article
Received Date: October 5,2021
Accepted Date: February 18, 2022
Published Date: February 28, 2022

A New Flexible Extension of Xgamma Distribution and its Application to COVID-19 Data

Harsh Tripathi[1], Abhimanyu Singh Yadav[2], Mahendra Saha[3] and Shivanshi Shukla[4].
[1]*Department of Mathematics, Lovely Professional University, Punjab, India*
[2]*Department of Statistics, Banaras Hindu University, Varansi, India*
[3]*Department of Statistics, Central University of Rajasthan, Rajasthan, India*

Email: [1]*mahendrasaha@curaj.ac.in*, [2]*asybhu10@gmail.com*, [3]*rsearchstat21@gmail.com*, [4]*speak2shivanshi@gmail.com*

Corresponding author's email: [4]*speak2shivanshi@gmail.com*

Abstract: *In this article, a new flexible extension of xgamma probability distribution has been proposed. Several well known distributional properties viz., raw moments, generating functions, conditional moments, mean deviation, quantile functions etc., of this flexible extension model have derived and studied in detail. Further, the estimation of the unknown model parameters along with the survival function and hazard function are estimated using maximum likelihood estimation technique. The Monte Carlo simulation has been performed to check the consistency of the proposed estimators for the different variation of sample size and model parameters. Finally, the superiority of proposed extension over several well known lifetime models has been illustrated using four data sets pertaining to COVID-19 cases in different country of the world.*

Keywords: *Xgamma distribution, Moments, Generating function, Conditional moments, Maximum likelihood method of estimation*.

1. Introduction

The development of new probability model plays an important role, as it is equipped with more flexibility that provides for explaining much wider range of real life situation. Foremost and vital requirement to analyse the considered (or given) data is the information about the probability distribution. For analysing the survival and reliability characteristics of the considered (or given) data, we must have the information of probability distribution which suited best to the considered (or given) data set before hand. In literature there exist several univariate, bivarite and multivariate probability distributions. For the complete analysis, data set must follow a specific pattern of the particular probability density function (PDF) and should have more or less similar shape of hazard rate function (HRF). HRF behaves in different ways like constant, monotone increasing, monotone decreasing, bathtub and inverted bathtub in real life scenario. Bathtub shape of HRF consists of two change points and a constant part enclosed within the change points. Hence depending upon the shape of hazard rate of survival data, we decide the plausible corresponding distribution model to state the interpretation and make the inferences about the considered data set. From the probabilistic point of view, we have large number of choices of models to analyse the data and pass the statement about the nature of a data and further obtain the various statistical properties of that data on the basis of chosen model.

In current fashion, the new probability distribution are developed by adding the additional parameter(s). There are many ways to add a new parameter to the distribution and expand the family of the distribution. Marshall-Olkin also suggested a approach to introduce a new parameter in a particular probability distribution and generate the family of considered distribution. For the more detailing of the Marshall-Olkin family readers may refer to see the Marshall and Olkin[15]. Marshall-Olkin[15] added a new parameter and developed the family of the exponential distribution (ED) and Weibull distribution (WD). The survival function (SF) of new extended distribution by Marshall-Olkin approach is given as;

$$\bar{G}(x,\alpha)=\frac{\alpha\bar{F}(x)}{1-\bar{\alpha}\bar{F}(x)};-\infty<x<\infty,0<\alpha<\infty \tag{1}$$

where, $\bar{\alpha}=1-\alpha$ and when $\alpha=1$ then $\bar{G}=\bar{F}$. $\bar{F}(x)$ is the SF of baseline distribution which is used to generate a new family of distribution. All the commonly-used methods of introducing an additional parameter have a stability property: if the method of adding a new parameter is applied twice on a particular distribution model then nothing new is obtained. For example, power of an exponential random variable has a Weibull distribution, but the power of a Weibull random variable is just another Weibull random variable. Similarly, if, in (1), SF of form $\bar{G}$ is introduced for $\bar{F}$, then (1) yields nothing new. PDF of Marshall-Olkin family corresponding to SF [see Equation 1] is:

$$g(x,\alpha)=\frac{\alpha f(x)}{\{1-\bar{\alpha}\bar{F}(x)\}^2} \tag{2}$$

and HRF of Marshall-Olkin family is obtained by using PDF and SF of the same family. It is given as below;

$$r(x;\alpha)=\frac{1}{\{1-\bar{\alpha}\bar{F}(x)\}}r_F(x) \tag{3}$$

In literature many researchers have developed the Marshall-Olkin family for different probability models. For the details of Marshall-Olkin family, researchers are suggested to follow the articles : Afify et al. [1], Alizadeh et al[2], Bdair et al [3] , Cordeiro and Lemonte [4], Eghwerido et al. [5], Ghitany et al. [6], Ghitany et al. [7], González-Hernández et al. [8], Jayakumar and Mathew [9], Jose et al.[10], Jose and Paul [11], Krishna et al. [12], Korkmaz et al. [13], Klakattawi et al. [14],Marshall-Olkin[15], Maxwell et al.[16], Nadarajah et al.[17], , Ristic and Kundu[18], Santos-Neto et al. [19], Saboor and Pogany[20], and Yousof et al. [22].

The main purpose of this contribution is to propose a new two-parameter extension of xgamma distribution, by applying the approach suggested by Marshall–Olkin. From now, we call it as flexible extension of xgamma distribution (FEXg). The generalization obtained by this method provides better flexibility to analyze reliability/survival data with monotone hazard rate as compared to some well known lifetime distributions. The different distributional properties such as moments, reliability and hazard functions, conditional moments, generating function, Quantile function, aging intensity, entropy etc., have been derived. Further, the maximum likelihood estimation technique has been applied to estimate the unknown parameters, survival function and hazard function of the proposed extension. The performances of the proposed estimators are studied in terms of mean square error using Monte Carlo simulations. Lastly, four

COVID-19 survival data sets of different countries are taken for the illustration of FEXg extension in real life.

Rest of article is organized as follows: FEXg is introduced and survival characteristics of FEXg are derived in section 2. Important statistical properties such as moments, conditional moments, mean deviation, Bonferroni and Lorenz curves and procedure of random number generation of FEXg are discussed in section 3. In section 4, we have discussed about the classical method of estimation of parameters of FEXg through MLE. In section 5, Monte Carlo simulation study is carried out to assess the performance of the above cited classical method for SF and HRF in terms of mean square error (MSE). For illustrative purposes, four real data sets are analyzed in section 6. Finally, concluding remarks are given in section 7.

2. The Model and It's Generalization

Xgamma distribution (XGD) [see, Sen et al.[21]) is a mixture of expenential and gamma distribution with specific proportion.Sen et al. [21] showed the superiorty of XGD over existing distributions by using real life examples. FEXg is the extended version of XGD (see, Sen et al.[21]) by adding a new parameter through the Marshall-Olkin approach (see, Marshall and Olkin [15]). Marshall-Olkin have discussed the stability property over the adding the parameter by using all common methods and the property is: if the method is applied twice, nothing new is obtained the second time around. They have also discussed the behaviour of HRF of Marshall-Olkin family. Adding a new parameter in the XGD by using Marshall-Olkin family makes XGD more flexible thereby enhancing its beauty and practical significance. Thus, FEXg become more realistic and useful in real life situation. Let X be a random variable follows XGD and then the SF and PDF of XGD are given below:

$$S(x;\theta)=\frac{\left(1+\theta+\theta x+\frac{\theta^2x^2}{2}\right)}{(1+\theta)}e^{-\theta x}\ ;x>0,\theta>0 \tag{4}$$

and

$$f(x;\theta)=\frac{\theta^2}{(1+\theta)}\left(1+\frac{\theta}{2}x^2\right)e^{-\theta x}\ ;x>0,\theta>0 \tag{5}$$

To obtained the SF of FEXg, we use SF of XGD [see, Equations 4]. The SF of FEXg $\bar{G}(x;\theta,\alpha)$ when X follows the XGD is as follows;

$$\bar{G}(x;\theta,\alpha)=\frac{\alpha\frac{\left(1+\theta+\theta x+\frac{\theta^2x^2}{2}\right)}{(1+\theta)}e^{-\theta x}}{\left[1-\bar{\alpha}\frac{\left(1+\theta+\theta x+\frac{\theta^2x^2}{2}\right)}{(1+\theta)}e^{-\theta x}\right]} \tag{6}$$

Graphical representation of SF of FEXg is shown in Figure 2. PDF $g(x;\theta,\alpha)$ and CDF $G(x;\theta,\alpha)$ of FEXg with the scale parameter θ and newly introduced parameter α is given as:

$$g(x;\theta,\alpha)=\frac{\alpha\theta^2}{(1+\theta)}\frac{\left(1+\frac{\theta}{2}x^2\right)e^{-\theta x}}{\left[1-\bar{\alpha}\frac{\left(1+\theta+\theta x+\frac{\theta^2x^2}{2}\right)}{(1+\theta)}e^{-\theta x}\right]^2};\quad x>0,\theta>0,\alpha>0 \tag{7}$$

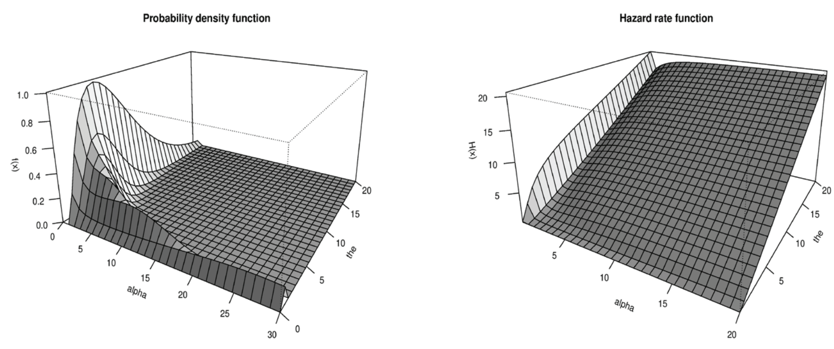


Figure 1: Density and hazard function of FEXg.

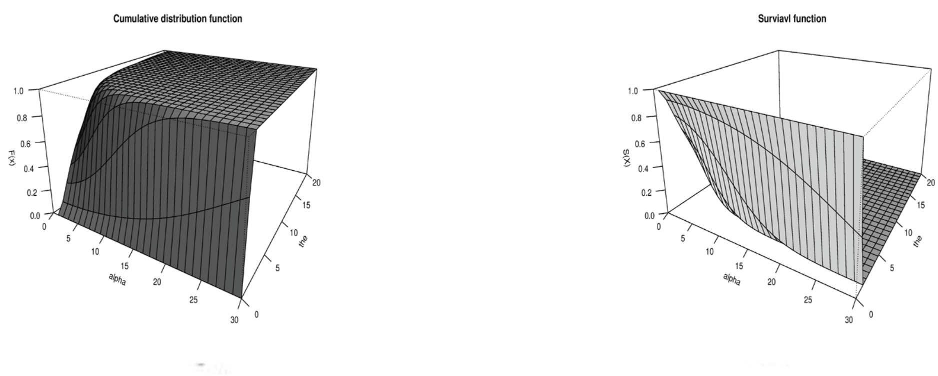


Figure 2: Distribution function and survival function of FEXg.

and

$$G(x;\theta,\alpha)=1-\left\{\frac{\alpha\frac{\left(1+\theta+\theta x+\frac{\theta^2x^2}{2}\right)}{(1+\theta)}e^{-\theta x}}{\left[1-\bar{\alpha}\frac{\left(1+\theta+\theta x+\frac{\theta^2x^2}{2}\right)}{(1+\theta)}e^{-\theta x}\right]}\right\};\quad x>0,\theta>0,\alpha>0 \tag{8}$$

if $\alpha=1$, then FEXg coincide with the XGD with the scale parameter θ. HRF is the most important property to know the survival behaviour of any life time model. Hazard rate better known as instantaneous failure rate is the dynamic speed with which an system or component fails, expressed in failures per unit of time. HRF $H(x;\theta,\alpha)$ of FEXg is:

$$H(x;\theta,\alpha)=\frac{\frac{\theta^2}{(1+\theta)}\left(1+\frac{\theta}{2}x^2\right)e^{-\theta x}}{\left\{\frac{\left(1+\theta+\theta x+\frac{\theta^2x^2}{2}\right)}{(1+\theta)}e^{-\theta x}\right\}\left\{1-\bar{\alpha}\frac{\left(1+\theta+\theta x+\frac{\theta^2x^2}{2}\right)}{(1+\theta)}e^{-\theta x}\right\}} \tag{9}$$

Also graphical representation of PDF and CDF of newly proposed model has been shown by the Figure 1 and 2, respectively. Figures of CDF, PDF, SF and HRF plot for various choices of α and θ. From the Figure 1, it is observed that the HRF of FEXg can take every possible shape, i.e, increasing, decreasing and bathtub shape for the chosen value of the scale parameter (θ) and the newly introduced parameter α. As compared to xgamma distribution, proposed distribution HRF is more flexible than xgamma distribution because Sen et al.[21] have been shown that the HRF of xgamma distribution can take only bathtub shape.

3. Statistical Properties of MOEXg

In this section, mathematical expressions of statistical properties viz., raw moments, generating functions, conditional moments, order statistics etc. of the proposed model have been obtained. These statistical properties play key role to analyze the data and we conclude about the behaviour of the data on the basis of these statistical properties.

3.1 Raw moments

Raw moments are the extremely important aspect of any distributional form because they help to determine the skewness and kurtosis of a model. The r-th raw moments about origin is given as;

$$E(X^r) = \int_0^\infty x^r g(x;\theta,\alpha)dx$$

$$E(X^r) = \int_0^\infty x^r \frac{\alpha\theta^2}{1+\theta}\left(1+\frac{\theta}{2}x^2\right)e^{-\theta x}\left[1-\bar{\alpha}\frac{\left(1+\theta+\theta x+\frac{\theta^2x^2}{2}\right)}{(1+\theta)}e^{-\theta x}\right]^{-2} dx$$

$$E(X^r) = \int_0^\infty x^r \frac{\alpha\theta^2}{1+\theta}e^{-\theta x}\left[1-\bar{\alpha}\frac{\left(1+\theta+\theta x+\frac{\theta^2x^2}{2}\right)}{(1+\theta)}e^{-\theta x}\right]^{-2} dx$$

$$+\frac{\theta}{2}\int_0^\infty x^2x^r \frac{\alpha\theta^2}{1+\theta}e^{-\theta x}\left[1-\bar{\alpha}\frac{\left(1+\theta+\theta x+\frac{\theta^2x^2}{2}\right)}{(1+\theta)}e^{-\theta x}\right]^{-2} dx \qquad (10)$$

We use the following Lemma to obtain the closed form solution of the above equation.

Lemma 1: Let X be random variable having FEXg then

$$I_1(a,r,c) = \int_0^\infty x^r e^{-ax}\left[1-\bar{\alpha}\frac{\left(1+c+cx+\frac{c^2x^2}{2}\right)}{(1+c)}e^{-cx}\right]^{-2} dx$$

$$= \sum_{i=0}^{\infty}\sum_{j=0}^{i}\sum_{k=0}^{j}\sum_{l=0}^{k} i_j\ j_k\ k_l\ (i+1)\left(\frac{\bar{\alpha}}{1+c}\right)^i \frac{c^{j+l}}{2^l}\frac{\Gamma(r+k+l+1)}{(a+ic)^{r+k+l+1}} \qquad (11)$$

and

$$I_2(a,r,c) = \int_0^\infty x^2x^r e^{-ax}\left[1-\bar{\alpha}\frac{\left(1+c+cx+\frac{c^2x^2}{2}\right)}{(1+c)}e^{-cx}\right]^{-2} dx$$

$$= \sum_{i=0}^{\infty}\sum_{j=0}^{i}\sum_{k=0}^{j}\sum_{l=0}^{k} i_j\ j_k\ k_l\ (i+1)\left(\frac{\bar{\alpha}}{1+c}\right)^i \frac{c^{j+l}}{2^l}\frac{\Gamma(r+k+l+3)}{(a+ic)^{r+k+l+3}} \qquad (12)$$

Proof:

$$I_1(a,r,c) = \sum_{i=0}^{\infty}(i+1)\left(\frac{\bar{\alpha}}{1+c}\right)^i\int_0^\infty x^r e^{-ax}e^{-icx}\left(1+c+cx+\frac{c^2x^2}{2}\right)^i dx$$

$$= \sum_{i=0}^{\infty}(i+1)\left(\frac{\bar{\alpha}}{1+c}\right)^i\sum_{j=0}^{i} i_j\ c^j\int_0^\infty x^r e^{-ax-icx}\left(1+x+\frac{cx^2}{2}\right)^j dx$$

$$= \sum_{i=0}^{\infty} (i+1)\left(\frac{\bar{\alpha}}{1+c}\right)^{i} \sum_{j=0}^{i} \binom{i}{j} c^{j} \sum_{k=0}^{j} \binom{j}{k} \int_0^{\infty} x^{r} e^{-ax-icx}\left(x+\frac{cx^2}{2}\right)^{k} dx$$
$$= \sum_{i=0}^{\infty} (i+1)\left(\frac{\bar{\alpha}}{1+c}\right)^{i} \sum_{j=0}^{i} \binom{i}{j} c^{j} \sum_{k=0}^{j} \binom{j}{k} \sum_{l=0}^{k} \binom{k}{l} x^{k} \int_0^{\infty} x^{r} e^{-ax-icx}\left(\frac{cx}{2}\right)^{l} dx$$
$$= \sum_{i=0}^{\infty} \sum_{j=0}^{i} \sum_{k=0}^{j} \sum_{l=0}^{k} \binom{i}{j}\binom{j}{k}\binom{k}{l} (i+1)\left(\frac{\bar{\alpha}}{1+c}\right)^{i} \frac{c^{j+l}}{2^{l}} \int_0^{\infty} x^{r+k+l} e^{-x(a+ic)} dx$$

Integral of above equation can be easily solved by use of gamma function.

$$I_1(a,r,c) = \sum_{i=0}^{\infty} \sum_{j=0}^{i} \sum_{k=0}^{j} \sum_{l=0}^{k} \binom{i}{j}\binom{j}{k}\binom{k}{l} (i+1)\left(\frac{\bar{\alpha}}{1+c}\right)^{i} \frac{c^{j+l}}{2^{l}} \frac{\Gamma(r+k+l+1)}{(a+ic)^{r+k+l+1}}$$

To get the solution of I_2 we have to proceed in similar manner as $I_1(a,r,c)$. So solution of $I_2(a,r,c)$ is:

$$I_2(a,r,c) = \sum_{i=0}^{\infty} \sum_{j=0}^{i} \sum_{k=0}^{j} \sum_{l=0}^{k} \binom{i}{j}\binom{j}{k}\binom{k}{l} (i+1)\left(\frac{\bar{\alpha}}{1+c}\right)^{i} \frac{c^{j+l}}{2^{l}} \frac{\Gamma(r+k+l+3)}{(a+ic)^{r+k+l+3}}$$

Put the $a=\theta$ and $c=\theta$ in Lemma 1, then the r-th raw moment is:

$$E(X^r) = \frac{\alpha\theta^2}{\theta+1}\left[I_1(\theta,r,\theta) + \frac{\theta}{2} I_2(\theta,r,\theta)\right] \tag{13}$$

Expression of first four raw moments are obtained by putting r =1,2,3,4 respectively in the Equation 13.

$$E(X) = \frac{\alpha\theta^2}{\theta+1}\left[I_1(\theta,1,\theta) + \frac{\theta}{2} I_2(\theta,1,\theta)\right]$$
$$E(X^2) = \frac{\alpha\theta^2}{\theta+1}\left[I_1(\theta,2,\theta) + \frac{\theta}{2} I_2(\theta,2,\theta)\right]$$
$$E(X^3) = \frac{\alpha\theta^2}{\theta+1}\left[I_1(\theta,3,\theta) + \frac{\theta}{2} I_2(\theta,3,\theta)\right]$$
$$E(X^4) = \frac{\alpha\theta^2}{\theta+1}\left[I_1(\theta,4,\theta) + \frac{\theta}{2} I_2(\theta,4,\theta)\right]$$

Formulas of central moments rely on the raw moments thus, central moments can be calculated with the help of above expressions of raw moments. Further, the skewness and kurtosis of the proposed model can be determined from the central moments. The first two moments along with the cofficients of skewness (SK) and kurtosis (KR) helps us in getting a general glimpse of the data. Mean (computed using first raw moment) gives the idea of measure of central tendency. Variance (the second central moment) narrates about the spread of data whereas SK and KR comments upon the measure of assymetricity and of peakedness respectively. Pearson formulated the skewness and kurtosis by following formulas:

$$SK = \frac{\mu_3^2}{\mu_2^3} \ \ and \ \ KR = \frac{\mu_4}{\mu_2^2}$$

3.2 Generating functions

This subsection consists the brief theory and expressions of the generating functions viz., moment generating function $M_x(t)$, characteristic function $\Phi_x(t)$ and cumulant generating function $k_x(t)$. Moment generating function can be calculated in following manner:

$$M_x(t) = E(e^{tx}) = \int_0^{\infty} e^{tx} g(x;\theta,\alpha) dx$$

$$M_x(t) = \int_0^{\infty} e^{tx} \frac{\alpha\frac{\theta^2}{(1+\theta)}\left(1+\frac{\theta}{2}x^2\right)e^{-\theta x}}{\left[1-\bar{\alpha}\frac{\left(1+\theta+\theta x+\frac{\theta^2 x^2}{2}\right)}{(1+\theta)}e^{-\theta x}\right]^2} dx \tag{14}$$

To get the solution of the integral of $M_x(t)$ [see, Equation 14], we have to used Lemma 1. Mathematical formulation of the $M_x(t)$ for the proposed plan is given below:

$$M_x(t) = \frac{\alpha\theta^2}{\theta+1}[I_1(\theta - t, 0, \theta) + I_2(\theta - t, 0, \theta)] \tag{15}$$

MGF suffers from the drawback that it is defined for a specified range i.e; $-h < t < h$ where h is a small positive number thus, we resort to characteristic function which is given for the real axis. Further, characteristic function always exist unlike MGF. For proposed distribution it can obtained with the help of Equation (15) and we replace the dummy parameter t by it. Therefore, the expression of the characteristic function $\Phi_x(t)$.

$$\Phi_x(t) = \frac{\alpha\theta^2}{\theta+1}[I_1(\theta - it, 0, \theta) + I_2(\theta - it, 0, \theta)] \tag{16}$$

Cumulants generating function is defined as the logarithmic of the characteristic function and is given in following equation [see, Equation 17].

$$k_x(t) = \log(\Phi_x(t))$$

$$k_x(t) = \log\left(\frac{\alpha\theta^2}{\theta+1}\right) + \log\{[I_1(\theta - it, 0, \theta) + I_2(\theta - it, 0, \theta)]\} \tag{17}$$

3.3 Conditional moments

Given that the life of the unit under observation exceeds a specified value say 'x' then one may be interested for the expression for moments under such condition. Such moments are called as conditional moments. Thus, about origin the expression for n-th conditional moment is;

$$E(X^n | X > x) = \int_x^\infty x^n \frac{g(x;\theta,\alpha)}{1 - G(x;\theta,\alpha)} dx$$

where, $g(x;\theta,\alpha)$ and $G(x;\theta,\alpha)$ is PDF and CDF of FEXg, given in Equations (7, 8).

$$E(X^n | X > x) = \frac{1}{1 - G(x;\theta,\alpha)} \frac{\alpha\theta^2}{1+\theta}\left[\int_x^\infty x^n \, a \, dx + \int_x^\infty x^{n+2} \, \frac{\theta}{2} \, a \, dx\right] \tag{18}$$

where $a = e^{-\theta x}\left(1 - \bar{\alpha}\frac{\left(1+\theta+\theta x+\frac{\theta^2 x^2}{2}\right)}{(1+\theta)} e^{-\theta x}\right)^{-2}$.

Two complicated integrals are involve in above equation. For the solution of these integrals [see, Equation (18)], we will apply the following lemma:

Lemma 2: Let X be random variable having FEXg then

$$L_1(a, r, c, \delta) = \sum_{j=0}^{i} \sum_{k=0}^{j} \sum_{l=0}^{k} \binom{i}{j} \binom{j}{k} \binom{k}{l} (i+1) \left(\frac{\bar{\alpha}}{1+c}\right)^i \frac{c^{j+l}}{2^l} \frac{\Gamma(r+k+l+1), \delta(a+ic)}{(a+ic)^{(r+k+l+1)}}$$

and

$$L_2(a, b, c, \delta) = \sum_{j=0}^{i} \sum_{k=0}^{j} \sum_{l=0}^{k} \binom{i}{j} \binom{j}{k} \binom{k}{l} (i+1) \left(\frac{\bar{\alpha}}{1+c}\right)^i \frac{c^{j+l}}{2^l} \frac{\Gamma(r+k+l+3), \delta(a+ic)}{(a+ic)^{(r+k+l+3)}}$$

Proof: Proof of lemma 2 is similar as the lemma 1. Hence the expression of conditional moment is:

$$E(X^r | X > x) = \frac{1}{1 - G(x;\theta,\alpha)} \frac{\alpha\theta^2}{1+\theta}\left[L_1(a, r, c, \delta) + \frac{\theta}{2} L_2(a, r, c, \delta)\right] \tag{19}$$

Put a= θ, r=n, c= θ and $\delta = x$ in Equation (19) and get the expression of n-th conditional moment of FEXg.

$$E(X^n | X > x) = \frac{1}{1 - G(x;\theta,\alpha)} \frac{\alpha\theta^2}{1+\theta}\left[L_1(\theta, n, \theta, x) + \frac{\theta}{2} L_2(\theta, n, \theta, x)\right] \tag{20}$$

Using Lemma 2, the first four conditional moments are given as:

$$E(X|X>x)=\frac{1}{1-G(x)}\frac{\alpha\theta^2}{1+\theta}\left[L_1(\theta,1,\theta,x)+\frac{\theta}{2}L_2(\theta,1,\theta,x)\right]$$

$$E(X^2|X>x)=\frac{1}{1-G(x;\theta,\alpha)}\frac{\alpha\theta^2}{1+\theta}\left[L_1(\theta,2,\theta,x)+\frac{\theta}{2}L_2(\theta,2,\theta,x)\right]$$

$$E(X^3|X>x)=\frac{1}{1-G(x;\theta,\alpha)}\frac{\alpha\theta^2}{1+\theta}\left[L_1(\theta,3,\theta,x)+\frac{\theta}{2}L_2(\theta,3,\theta,x)\right]$$

$$E(X^4|X>x)=\frac{1}{1-G(x;\theta,\alpha)}\frac{\alpha\theta^2}{1+\theta}\left[L_1(\theta,4,\theta,x)+\frac{\theta}{2}L_2(\theta,4,\theta,x)\right]$$

3.4 Mean deviation

Mean deviation of the FEXg about the mean has been defined in following equation:

$$MD=\int_0^{\infty}|x-\mu|g(x;\theta,\alpha)dx$$

where, μ mean of the FEXg.

$$MD=\int_0^{\mu}(\mu-x)g(x;\theta,\alpha)dx+\int_{\mu}^{\infty}(x-\mu)g(x;\theta,\alpha)dx$$
$$=2\mu G(\mu)-2\mu+2\int_{\mu}^{\infty}xg(x;\theta,\alpha)dx$$

$$MD=2\mu G(\mu)-2\mu+2\frac{\alpha\theta^2}{1+\theta}\left(L_1(\theta,1,\theta,\mu)+\frac{\theta}{2}L_2(\theta,1,\theta,\mu)\right) \tag{21}$$

3.5 Order statistics

Let $X_1,X_2,X_3,\ldots,X_n$ is a random sample of size from FEXg. Then, the ordered observations $X_{(1)}<X_{(2)}<X_{(3)}<\ldots..<X_{(n)}$ constitute the order statistic. Let $Y=X_{(k;n)}$ denotes the k-th order statistic, then the PDF and CDF of k-th order statistic are computed as follows:

$$g(y;\theta,\alpha)=n\genfrac{}{}{0pt}{}{}{k}\,G^{(k-1)}(y;\theta,\alpha)[1-G(y;\theta,\alpha)]^{n-k}g(y;\alpha,\theta)$$
$$=n\genfrac{}{}{0pt}{}{}{k}\,G^{k-1}(y;\theta,\alpha)\sum_{i=0}^{n-k}(-1)^i n-k\genfrac{}{}{0pt}{}{}{i}\,[G(y;\alpha,\theta)]^i g(y;\alpha,\theta)$$

$$g(y;\theta,\alpha)=n\genfrac{}{}{0pt}{}{}{k}\,\sum_{i=0}^{n-k}(-1)^i n-k\genfrac{}{}{0pt}{}{}{i}\,[G(y;\theta,\alpha)]^{(k+i-1)}g(y;\theta,\alpha) \tag{22}$$

Equation (22) represents the PDF of k-th order statistics. Now, the CDF of k-th order statistics is:

$$G(y;\theta,\alpha)=\sum_{j=k}^{n}n\genfrac{}{}{0pt}{}{}{j}\,G^j(y;\theta,\alpha)[1-G(y;\theta,\alpha)]^{n-j}$$

$$G(y;\theta,\alpha)=\sum_{j=k}^{n}\sum_{i=0}^{n-j}n\genfrac{}{}{0pt}{}{}{j}\,n-j\genfrac{}{}{0pt}{}{}{i}\,(-1)^i[G(y;\theta,\alpha)]^{j+i} \tag{23}$$

By putting the value of PDF and CDF [see, Equations (7) and (8) respectively] of FEXg in Equations (22) and (23) then we get the PDF and CDF of k-th oreder staistics of FEXg. Also, the distribution of $X_{(1)}=\min(X_{(1)}<X_{(2)}<X_{(3)}<\ldots..<X_{(n)})$ and $X_{(n)}=\max(X_{(1)}<X_{(2)}<X_{(3)}<\ldots..<X_{(n)})$ can be computed with help of above Equations (8) by putting $k=1$ and $k=n$ respectively.

3.6 Bonferroni and Lorenz curves

Bonferroni and Lorenz curves are very important tools in actuarial and population science to study the income and poverty level. Let X be a random variable with PDF $g(x;\alpha,\theta)$, defined in Equation (7) then

Bonferroni curve $B(p)$ and Lorenz curve $L(p)$ are defined by the following Equations (24) and (25) respectively.

$$B(p) = \frac{1}{p\mu}\left[\mu - \int_q^\infty xg(x;\theta,\alpha)dx\right]$$
$$L(p) = \frac{1}{\mu}\left[\mu - \int_q^\infty xg(x;\theta,\alpha)dx\right]$$

After simplification, the final expression of $B(p)$ and $L(p)$ are obtained as:

$$B(p) = \frac{1}{p\mu}\left[\mu - \frac{\alpha\theta^2}{1+\theta}\left(L_1(\theta,1,\theta,q) + \frac{\theta}{2}L_2(\theta,1,\theta,q)\right)\right] \tag{24}$$

and

$$L(p) = \frac{1}{\mu}\left[\mu - \frac{\alpha\theta^2}{1+\theta}\left(L_1(\theta,1,\theta,q) + \frac{\theta}{2}L_2(\theta,1,\theta,q)\right)\right] \tag{25}$$

where, $\mu = E(x)$. Bonferroni and Gini indices are helpful in several fields such as income, wealth, reliability, insurance, demography and medicine. Mathematical expressions thes indices based on the above two curves are given as;

$$B = 1 - \int_0^1 B(p)dp, \qquad G = 1 - 2\int_0^1 L(p)dp$$

3.7 Entropy measurement

Entropy is used to measure the randomness of systems and it is widely used in areas like physics, molecular imaging of tumors and sparse kernel density estimation. General expression of generalized entropy is given below:

$$G_E = \frac{\nu_\lambda \mu^{-\lambda} - 1}{\lambda(\lambda-1)}; \qquad \lambda \neq 0,1$$

where, $\nu_\lambda = \int_0^\infty x^\lambda g(x;\theta,\alpha)dx$. ν_λ is determined by the λ-th raw moments. Now the expression of generalized entropy in case of FEXg model is:

$$G_E = \frac{\frac{\alpha\theta^2}{\theta+1}\left[I_1(\theta,\lambda,\theta) + \frac{\theta}{2}I_2(\theta,\lambda,\theta)\right]\mu^{-\lambda} - 1}{\lambda(\lambda-1)} \tag{26}$$

3.8 Ageing intensity

Ageing is a important aspect in study of the survival and reliability analysis and ageing is a basic characteristic of the any system or product. Ageing characteristic of the system can be calculated mathematically using formula given in Equation (27). Ageing intensity is a function of x and is defined as the ratio of hazard rate to baseline hazard rate. Ageing intensity is denoted by AI, is given below:

$$L_x(t) = \frac{1}{\bar{G}_x(t)} \frac{t\, g_x(t)}{\log(\bar{G}_x(t))} \tag{27}$$

Using the expression of $g_x(t)$, $\bar{G}_x(t)$ for the proposed probability distribution, we get

$$L_x(t) = -\frac{t\,\theta^2\left(1+\frac{\theta}{2}t^2\right)}{\log\left(\frac{\alpha\frac{\left(1+\theta+\theta t+\frac{\theta^2t^2}{2}\right)}{(1+\theta)}e^{-\theta t}}{\left[1-\bar{\alpha}\frac{\left(1+\theta+\theta t+\frac{\theta^2t^2}{2}\right)}{(1+\theta)}e^{-\theta t}\right]}\right)\left(1+\theta+\theta t+\frac{\theta^2t^2}{2}\right)\left[1-\bar{\alpha}\frac{\left(1+\theta+\theta t+\frac{\theta^2t^2}{2}\right)}{(1+\theta)}e^{-\theta t}\right]} \tag{28}$$

Pattern of AI depends on the hazard rate. If hazard rate is increasing, decreasing and constant then ageing is positive, negative and non-ageing respectively. When X is a non negative random variable then $L_x(t)$ can take three value namely, $= 1$, < 1 and > 1 for all $t > 0$. Value of $L_x(t)$ is 1 if and only if hazard rate is constant. $L_x(t)$ is > 1 if hazard rate is increasing in t and $L_x(t)$ is < 1 if hazard rate is decreasing function in t.

3.9 Residual lifetime functions

Residual lifetime function is used to determine the remaining lifetime associated with any particular system. In this section, we have derived the residual lifetime functions for the FEXg. It is defined by the R_t and reversed residual lifetime function is defined by $\overline{R_t}$ which denotes the time elapsed from the failure of a component given that it has life less or equal to t. Expressions of residual lifetime function and reversed lifetime function are:

$$R_t = P(x - t|x > t); \quad t \geq 0$$

$$\overline{R_t} = P(t - x|x \leq t); \quad t \geq 0$$

The survival function of residual life function and reverse residual life function are given by;

$$\bar{G}_{R_T}(x) = \frac{\bar{G}(x+t)}{\bar{G}(t)}, \qquad \bar{G}_{\bar{R}_T}(x) = \frac{G(t-x)}{G(t)}$$

respectively. Hence, using the survival function of the proposed model the required expressions of $\bar{G}_{R_T}(x)$, $\bar{G}_{\bar{R}_T}(x)$ are obtained as;

$$\bar{G}_{R_T}(x)|_{x>t} = \frac{\dfrac{\alpha\dfrac{\left(1+\theta+\theta(x+t)+\frac{\theta^2(x+t)^2}{2}\right)}{(1+\theta)}e^{-\theta(x+t)}}{\left[1-\bar{\alpha}\dfrac{\left(1+\theta+\theta(x+t)+\frac{\theta^2(x+t)^2}{2}\right)}{(1+\theta)}e^{-\theta(x+t)}\right]}}{\dfrac{\alpha\dfrac{\left(1+\theta+\theta t+\frac{\theta^2t^2}{2}\right)}{(1+\theta)}e^{-\theta t}}{\left[1-\bar{\alpha}\dfrac{\left(1+\theta+\theta t+\frac{\theta^2t^2}{2}\right)}{(1+\theta)}e^{-\theta t}\right]}}, \quad \bar{G}_{\bar{R}_T}(x)|_{x<t} = \frac{1-\left\{\dfrac{\alpha\dfrac{\left(1+\theta+\theta(t-x)+\frac{\theta^2(t-x)^2}{2}\right)}{(1+\theta)}e^{-\theta(t-x)}}{\left[1-\bar{\alpha}\dfrac{\left(1+\theta+\theta(t-x)+\frac{\theta^2(t-x)^2}{2}\right)}{(1+\theta)}e^{-\theta(t-x)}\right]}\right\}}{1-\left\{\dfrac{\alpha\dfrac{\left(1+\theta+\theta t+\frac{\theta^2t^2}{2}\right)}{(1+\theta)}e^{-\theta t}}{\left[1-\bar{\alpha}\dfrac{\left(1+\theta+\theta t+\frac{\theta^2t^2}{2}\right)}{(1+\theta)}e^{-\theta t}\right]}\right\}} \tag{29}$$

Corresponding PDFs and hazard rate function can be obtained with help of above defined survival functions for both the lifetimes.

4. Maximum Likelihood Estimation

In this section, we have considered maximum likelihood estimation (MLE) procedure for the estimation of the unknown model parameters and the survival characteristics (SF, HRF). Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from Equation (7). Then, the log-likelihood function for the observed random sample $x_1, x_2, \ldots, x_n$ is given as;

$$\ell(\theta,\alpha) = n\log\alpha + 2n\log\theta - n\log(1+\theta) + \sum_{i=1}^{n}\log(1+\frac{\theta}{2}x_i^2) - \theta\sum_{i=1}^{n}x_i - 2\sum_{i=1}^{n}\log(U_i)$$

where,

$$U(x_i;\theta,\alpha) = \left[1 - \bar{\alpha}\left(1+\theta+\theta x_i + \frac{\theta^2 x_i^2}{2}\right)\frac{e^{-\theta x_i}}{1+\theta}\right].$$

The resulting partial derivatives of the log-likelihood function are

$$\frac{\partial\,\ell(\theta,\alpha)}{\partial\alpha} = \frac{n}{\alpha} - 2\sum_{i=1}^{n}\frac{\left(1+\theta+\theta x_i+\frac{\theta^2x_i^2}{2}\right)\frac{e^{-\theta x_i}}{1+\theta}}{U(x_i)} \tag{31}$$

$$\frac{\partial\,\ell(\theta,\alpha)}{\partial\theta}=\frac{2n}{\theta}-\frac{n}{\theta+1}+\sum_{i=0}^{n}\frac{x_i^2/2}{(1+\frac{\theta}{2}x_i^2)}-\sum_{i=0}^{n}x_i-2\sum_{i=0}^{n}\frac{1}{U_i}\times A_i \tag{32}$$

$$A_i=\bar{\alpha}\left[\frac{\theta x_i e^{-\theta x_i\{(1+\theta+1)\}}+\frac{\theta^2}{2}x_i^2 e^{-\theta x_i\{x_i(1+\theta)+1\}}}{(1+\theta)^2}\right]$$

Equating these partial derivatives to zero does not yield closed-form solutions for the MLEs and thus a numerical method is used for solving these equations simultaneously. So for the determination of MLEs from above equations [see, Equation 31 and 32], we have used nlm() function in R. Substituting the MLEs $(\hat{\alpha}_{mle},\hat{\theta}_{mle})$ of (α,θ) and using the invariance properties of MLEs, we can get the estimators of $\bar{G}(t)$ and $H(t)$ as;

$$\hat{\bar{G}}(t)_{mle}=\frac{\hat{\alpha}_{mle}^{\frac{\left(1+\hat{\theta}_{mle}+\hat{\theta}_{mle}t+\frac{\hat{\theta}_{mle}^2t^2}{2}\right)}{(1+\hat{\theta}_{mle})}e^{-\hat{\theta}_{mle}t}}}{\left[1-\hat{\bar{\alpha}}_{mle}^{\frac{\left(1+\hat{\theta}_{mle}+\hat{\theta}_{mle}t+\frac{\hat{\theta}_{mle}^2t^2}{2}\right)}{(1+\hat{\theta}_{mle})}e^{-\hat{\theta}_{mle}t}}\right]} \tag{33}$$

and

$$\hat{H}(t)_{mle}=\frac{\frac{\hat{\theta}_{mle}^2}{(1+\hat{\theta}_{mle})}\left(1+\frac{\hat{\theta}_{mle}}{2}t^2\right)e^{-\hat{\theta}_{mle}t}}{\left\{\frac{\left(1+\hat{\theta}_{mle}+\hat{\theta}_{mle}t+\frac{\hat{\theta}_{mle}^2t^2}{2}\right)}{(1+\hat{\theta}_{mle})}e^{-\hat{\theta}_{mle}t}\right\}\left\{1-\hat{\bar{\alpha}}_{mle}^{\frac{\left(1+\hat{\theta}_{mle}+\hat{\theta}_{mle}t+\frac{\hat{\theta}^2{}_{mle}t^2}{2}\right)}{(1+\hat{\theta}_{mle})}e^{-\hat{\theta}_{mle}t}}\right\}} \tag{34}$$

respectively for the given value of t.

5. Simulation Study

In this section, Monte Carlo simulation study is compiled to check the performances of maximum likelihood estimators for the proposed distribution. Monte carlo simulation study provides an enviourment under which we repeat the same program for prefixed number of times under similar conditions, and this enables us to estimate the average estimate of parameter and average mean square error for considered model. The performances of the estimators i.e. survival and hazard functions $[\bar{G}(t)\text{and}H(t)]$ are studied in terms of average mean square errors. The study is carried out for the different variation of sample size (n) and parametric values (θ,α). In particular, we have taken, $n=10,20,30,40,50,100,150,200$ and (α,θ)=(1.60,0.32), (0.60,0.32), (1.75,0.50), (2.25,0.45), (1.20,0.15), (1.05,0.35). The estimates of the survival function and hazard function are computed using invariance property of the MLE for arbitrary chosen specified mission time $t=2,3\&\,4$. For each considered combination (α,θ), we generate the sample of size n from the FEXg using accept reject method of sample generation. Next, we calculated the average value (AV) and mean square error (MSE) the proposed estimators based on $N=5000$ replications using the following expressions;

$$AV(\theta)=\frac{1}{N}\sum_{j=1}^{N}\theta_j,\;AV(\alpha)=\frac{1}{N}\sum_{j=1}^{N}\alpha_j$$

$$MSE\,(\bar{G}(t))=\frac{1}{N}\sum_{j=1}^{N}\{\bar{G}(t))_j-\bar{G}(t)\}^2,\;MSE\,(H(t))=\frac{1}{N}\sum_{j=1}^{N}\{H(t)_j-H(t)\}^2$$

From Table (2), it is observed that the average MSEs of the $\bar{G}(t)$ and $H(t)$ decrease as sample sizes increases for all the considered combination of (α,θ). The decreasing trend of MSEs proves that the estimators of $\bar{G}(t)$ and $H(t)$ are consistent. All simulations were performed using programs written in the open source statistical package R.

6 Real life Examples and Discussion

Coronavirus disease (COVID-19) is a zoonatic disease caused by coronavirus. The wide spread of this disease has shaped into a global pandemic. It has worstly affected the different spheres of life. The study of data pertaining to coronavirus has become inevitable so that we can study it's effect on social and economic front. Thus, in this section we attempt to illustrate the practical applicability of our proposed model on the data releated to Coronavirus disease (COVID-19). Description of the considered examples and associated data are given below. Also, the descriptive summary, viz., Minimum, Q_1, median, mean, Q_3, maximum, coefficient of skewness (CS) and coefficient of kurtosis (CK) are displayed in Table 7. For the applications part, first we checked whether the considered data sets comes from FEXg or not by using Kolmogrov-Smirnov (K-S) goodness-of-fit test. Thes test based on the K-S statistic compares an empirical and a theoretical model by computing the maximum absolute difference between the empirical and theoretical CDFs and is defined as $D_n = Sup_x|F_n(x) - F(x;\Theta)|$, where, $\Theta = (\alpha, \theta)$ and Sup_x is the supremum of the set of the distances, $F_n(x)$ is the empirical distribution function and $F(x;\Theta)$ is the CDF. Note that, K-S statistic to be used only to verify the goodness-of-fit not as a discrimination criteria. Therefore, we consider two discrimination criteria based on the likelihood-function evaluated at the MLEs. The criterion are: Akaike's Information Criteria (AIC) and Bayesian Information Criteria (BIC). These statistics are given by $AIC = -2l(\widehat{\Theta}) + 2k$, $BIC = -2l(\widehat{\Theta}) + 2\ln(n)$, where, $l(\widehat{\Theta})$ denotes the log-likelihood function evaluated at the MLEs, k is the number of model parameters and n is the sample size. The model with lowest values for these two statistics could be chosen as the best model to fit the data. Tables 3, 4, 5 and 6 are all about to show the flexibility of proposed model over other life time models.

Data I: Following observations represents the new cases of Covid-19 in Italy during 31^{st} May 2020 to 30^{th} June 2020 [see, https://www.worldometers.info/coronavirus/country/italy/] and the observations are:

334,200,319,322,177,519,270,197,280,283,202,380,163,347,337,301

210,329,332,251,264,224,221,113,190,296,255,175,174,126,142

Model fitting summary of considered data set has been given in Table 3. The values of MLEs of the parameters, $l(\widehat{\Theta})$, AIC, BIC, K-S Statistic with corresponding p values are reported. From Table 3, it is observed that the proposed model is best fit as compared to generalized exponential distribution (GED), Weibull distribution (WD), Transmuted Rayleigh (TR) distribution, Frechet distribution (FD), Lindley distribution (LD), inverse Weibull distribution (IWD), Akash distribution (AKD) and xgamma distribution (XGD) in terms of p value.

Data II: Data represent the percentage of death rate in India due to Covid-19 pandemic from 12^{th} March 2020 to 12^{th} April, 2020, for more detail one may visit to the website and URL of the website is www.worldometers.info/coronavirus/country/india/.

20,16.67,16.67,13.33,13.33,17.65,17.65,16.67,17.86,17.86,22.58

22.73,20,21.82,30.77,21.52,22.22,22.13,23.88,22.15,28.16,27.38

30.94,30.18,26.46,26.61,25.48,26.02,26.33,24.34,22.91,23.46

Model fitting summary of considered data set II has been given in Table 4. The values of MLEs of the parameters, $l(\widehat{\Theta})$, AIC, BIC, K-S Statistic with corresponding p values are displayed in Table 4. From Table 4, it is observed that the proposed model is best fit as compared to generalized exponential distribution (GED), Frechet distribution (FD), Lindley distribution (LD), Akash distribution (AKD), xgamma distribution (XGD), inverted exponential distribution (IED), , exponential distribution (ED) and inverse Weibull distribution (IWD) in terms of p value.

Data III: Data represents the death rate due to Coronaviruses (CoVID-19). These are large family of viruses that cause illness ranging from the common cold to more severe diseases such as Middle East Respiratory Syndrome (MERS-CoV) and Severe Acute Respiratory Syndrome (SARS-CoV). Following data represents the death rate due to CoVID-19 for a particular day data in China. URL is towardsdatascience.com/an-r-package-to-explore-the-novel-coronavirus-590055738ad6

2.4,2.4,4.9,4.4,3.2,1.5,1.5,2.1,2,3,3,3.2,2.8,2.9,2.3.

Model fitting summary of considered data set III has been given in Table 5. The values of MLEs of the parameters, $l(\widehat{\Theta})$, AIC, BIC, K-S Statistic with corresponding p values of data III are displayed in Table 5. From Table 5, it is observed that the proposed model is best fit as compared to exponential power distribution (EPD), Frechet distribution (FD), , Weibull distribution (WD), xgamma distribution (XGD), Akash distribution (AKD), inverted exponential distribution (IED), exponential distribution (ED) and Lindley distribution (LD) in terms of p value.

Data IV: Here, we consider the corona-virus cases distribution among the fifteen countries viz.,France, Italy, Spain, US, Germony, UK, Turkey, Iran, Russia, China, Brazil, Canada, Belgium, Netherlands and Switzerland. Data has taken from a website and URL is https://www.worldometers.info/coronavirus/coronavirus-cases/. Data is given in percentage and the observations are:

5.37,6.56,7.61,32.83,5.24,5.06,3.65,3.03,2.89,2.74,2.10,1.57,1.55,1.27,0.97

Model fitting summary of considered data set IV has given in Table 6. The MLEs of the parameters, $l(\widehat{\Theta})$, AIC, BIC, K-S Statistic with corresponding p values of data IV are displayed in Table 6. From Table 6, it has been observed that the proposed model is best fit as compared to xgamma distribution (XGD), Lindley distribution (LD), Akash distribution (AKD), inverted exponential distribution (IED), inverse xgamma distribution (IXGD), inverse Lindley distribution (ILD), Pareto type-2 Lomax distribution (Pt2LD), inverse Pareto (IP) and exponential power distribution (EPD) in terms of p value. Descriptive summary of the data IV has given in Table 7.

The descriptive summary of all the considered data sets are shown via box plots in Figure 3. Further, the empirical cumulative distribution function (ECDF) plots for the proposed model with all the considered competitive models are given in Figures 4. From ECDF plots, it has been noticed that FEXg distribution provides better fit as compared to the other two parametric distributions. Also, the estimated value of survival function and hazard rate function for all the considered data sets for the differently chosen value of mission time t are reported in Table 8.

We found that proposed probability distribution suited well to real life scenarios as compared to some popular existing probability distributions and to supposrt this statement we have provided numerical results in Tables 3-6.

7. Conclusions

In this article, we have proposed a new lifetime probability distribution, named as FEXg distribution. Different distribution properties such as moments, conditional moments, mean deviation, generating function, entropy, order statistics, ageing intensity, residual and reverse residual functions etc. are derived. The MLE method is used to estimate the survival function and hazard function for specified time. Further, the trend of the estimators are studied through Monte Carlo simulation, and noticed that the MSEs of the SF and HRF are decreases as the magnitude of sample size is increases which insure the trend of consistency of the estimators. Lastly, the applicability of the introduced extension has been shown using noble Coronavirus data of different countries. It is worthless to mention that the proposed FEXg model is a good competitor of several one and two parametric family of lifetime distributions. Hence, the proposed extension might be chosen by researcher to model the real life data with monotone increasing, monotone decreasing and bathtub hazard rate behaviour.

References

[1] Afify, A. Z., Cordeiro, G. M., Yousof, H. M., Saboor, A., and Ortega, E. M. **(2018)**. The Marshall-Olkin additive Weibull distribution with variable shapes for the hazard rate. *Hacettepe Journal of Mathematics and Statistics*, 47(2): 365-381.

[2] Alizadeh, M., Cordeiro, G. M., De Brito, E., and Demétrio, C. G. B. **(2015)**. The beta Marshall-Olkin family of distributions. *Journal of Statistical Distributions and Applications*, 2(1): 4.

[3] Bdair, O., & Haj Ahmad, H. **(2019)**. Estimation of the marshall-olkin Pareto distribution parameters: comparative study. *Revista Investigacion Operacional*, 41(2).

[4] Cordeiro, G. M., and Lemonte, A. J. **(2013)**. On the Marshall-Olkin extended weibull distribution. *Statistical papers*, 54(2): 333-353.

[5] Eghwerido, J. T., Ogbo, J. O., & Omotoye, A. E. **(2021)**. The Marshall-Olkin Gompertz Distribution: Properties and Applications. *Statistica*, *81*(2) :183-215.

[6] Ghitany, M. E., Al-Awadhi, F. A., and Alkhalfan, L. **(2007)**. Marshall-Olkin extended Lomax distribution and its application to censored data. *Communications in Statistics-Theory and Methods*, 36(10) : 1855-1866.

[7] Ghitany, M. E., Al-Mutairi, D. K., Al-Awadhi, F. A., and Al-Burais, M. M. **(2012)**. Marshall-Olkin extended Lindley distribution and its application. *International Journal of Applied Mathematics*, 25(5): 709-721.

[8] González-Hernández, I. J., Granillo-Macías, R., Rondero-Guerrero, C., & Simón-Marmolejo, I. **(2021)**. Marshall-Olkin distributions: a bibliometric study. *Scientometrics*, *126*(11): 9005-9029.

[9] Jayakumar, K., and Mathew, T. **(2008)**. On a generalization to Marshall-Olkin scheme and its application to Burr type XII distribution. *Statistical Papers*, 49(3): 421-439.

[10] Jose, K. K., Naik, S. R., and Ristić, M. M. **(2010)**. Marshall-Olkin q-Weibull distribution and max–min processes. *Statistical papers*, 51(4): 837-851.

[11] Jose, K. K. and Paul, A. **(2018)**. Marshall olkin exponential power distribution and its generalization: Theory and application. *IAPQR Transaction*, 43(1).

[12] Krishna, E., Jose, K. K., Alice, T., and Ristić, M. M. **(2013)**. The Marshall-Olkin Fréchet distribution. *Communications in Statistics-Theory and Methods*, 42(22) : 4091-4107.

[13] Korkmaz, M. Ç., Yousof, H. M., Hamedani, G. G., and Ali, M. M. **(2018)**. The Marshall-Olkin Generalized G poisson family of distribution. *Pakistan Journal of Statistics*, 34(3): 251-267.

[14] Klakattawi, H., Alsulami, D., Elaal, M. A., Dey, S., & Baharith, L. **(2022)**. A new generalized family of distributions based on combining Marshal-Olkin transformation with TX family. *PloS one*, *17*(2), e0263673.

[15] Marshall, A. W., and Olkin, I. **(1997)**. A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families. *Biometrika*, 84(3) : 641-652.

[16] Maxwell, O., Chukwu, A. U., Oyamakin, O. S., and Khaleel, M. A. **(2019)**. The Marshall-Olkin Inverse Lomax Distribution (MO-ILD) with Application on Cancer Stem Cell. *Journal of Advances in Mathematics and Computer Science*, 1-12.

[17] Nadarajah, S., Jayakumar, K., and Ristić, M. M. **(2013)**. A new family of lifetime models. *Journal of Statistical Computation and Simulation*, 83(8) : 1389-1404.

[18] Ristic, M. M., and Kundu, D. **(2015)**. Marshall-Olkin generalized exponential distribution. *Metron*, 73(3): 317-333.

[19] Santos-Neto, M., Bourguignon, M., Zea, L. M., Nascimento, A. D., and Cordeiro, G. M. **(2014)**. The Marshall-Olkin extended Weibull family of distributions. *Journal of Statistical Distributions and Applications*, 1(1): 9.

[20] Saboor, A., and Pogány, T. K. **(2016)**. Marshall–Olkin gamma–Weibull distribution with applications. *Communications in Statistics-Theory and Methods*, 45(5): 1550-1563.

[21] Sen, S., Maiti, S. S., Chandra, N. **(2016)**. The xgamma distribution: statistical properties and application. *Journal of Modern Applied Statistical Methods*, 15(1): 774-788.

[22] Yousof, H. M., Rasekhi, M., Alizadeh, M., and Hamedani, G. G. **(2018)**. The Marshall-Olkin exponentiated generalized G family of distributions: properties, applications and characterizations. *Journal of Nonlinear Sciences and Applications* forthcoming.

Table 1: Mean, Variance, Skewness, Kurtosis and Median for the various values of α and θ.

α	θ	Mean	Median	Variance	Skewness	Kurtosis
0.30	0.50	0.92869	0.94905	0.27157	-0.07013	1.82952
	0.75	0.46093	0.45349	0.07515	0.05575	1.79105
	1.00	0.30432	0.30274	0.02990	-0.02234	1.83991
	1.50	0.16922	0.16624	0.00975	0.04956	1.82119
0.80	0.50	2.40994	2.45545	1.95419	-0.03050	1.80273
	0.75	1.26778	1.28084	0.52100	-0.01948	1.82356
	1.00	0.80748	0.81068	0.20391	0.03872	1.87685
	1.50	0.45196	0.45149	0.07033	0.00713	1.74957
1.75	0.50	5.23050	5.18624	7.77678	0.07377	1.79422
	0.75	3.11015	3.05415	4.41424	0.02345	1.79923

	1.00	2.24733	2.29250	1.72972	-0.07194	1.74407
	1.50	1.26756	1.36504	1.56621	-0.11267	1.77688
2.00	0.50	5.49671	5.39148	8.52423	0.07095	1.75142
	0.75	3.42854	3.36312	4.14281	0.00947	1.84185
	1.00	2.28991	2.31023	1.91979	0.05839	1.78504
	1.50	1.31076	1.28566	0.65196	0.04810	1.78938
2.5	0.50	6.21731	6.08132	9.59394	0.03831	1.79644
	0.75	3.80539	3.76465	3.89006	0.00885	1.87036
	1.00	2.72659	2.71487	2.72596	0.04900	1.80169
	1.5	1.52906	1.48998	0.83842	0.04219	1.77426
3.0	0.50	7.07237	7.13003	11.3914	-0.05354	1.81212
	0.75	3.99226	3.91462	3.83259	0.07685	1.84265
	1.00	2.86224	2.80709	2.40283	0.08433	1.82191
	1.5	1.79435	1.84541	1.05947	-0.07654	1.82514

Table 2: MSE, Average estimated value of $\bar{G}(t)$ and $H(t)$ for different values of t.

n	α,θ	t	$\bar{G}(t)$	$H(t)$	$\hat{\bar{G}}(t)_{mle}$	$\hat{H}(t)_{mle}$	MSE of $\bar{G}(t)$	MSE of $H(t)$
10	1.60,0.32	2	0.91095	0.05107	0.99182	0.00507	0.002875	0.000925
20					0.96300	0.02258	0.001747	0.000531
30					0.95337	0.028377	0.001351	0.000396
40					0.94908	0.030902	0.001194	0.000343
50					0.94648	0.032473	0.001102	0.000310
100					0.94240	0.034859	0.000913	0.000247
150					0.94051	0.035962	0.000830	0.000220
200					0.94019	0.036155	0.000822	0.000217
10	0.60,0.32	2	0.79324	0.11859	0.94828	0.03633	0.016152	0.005944
20					0.89035	0.07524	0.008476	0.002721
30					0.86824	0.08977	0.005443	0.001593
40					0.86332	0.09284	0.004671	0.001232
50					0.85352	0.09923	0.003889	0.000987
100					0.84401	0.10534	0.002779	0.000524
150					0.84118	0.10716	0.002416	0.000354
200					0.83804	0.10918	0.002120	0.000267
10	1.75,0.50	3	0.73538	0.13489	0.93482	0.04327	0.015222	0.005593
20					0.85156	0.09519	0.009101	0.002410
30					0.82436	0.11131	0.007148	0.001694
40					0.81282	0.11800	0.005819	0.001161
50					0.80610	0.12192	0.005086	0.000922
100					0.79268	0.12940	0.003511	0.000447
150					0.78985	0.13087	0.003124	0.000293
200					0.78766	0.13193	0.002929	0.000241

10	2.25,0.45	3	0.81607	0.09022	0.96156	0.02227	0.009287	0.002290
20					0.90783	0.05207	0.005703	0.001206
30					0.89173	0.06052	0.004496	0.000868
40					0.88168	0.06577	0.003806	0.000663
50					0.87726	0.06799	0.003353	0.000557
100					0.86744	0.07301	0.002583	0.000362
150					0.86365	0.07502	0.002261	0.000280
200					0.86268	0.07554	0.002182	0.000255
10	1.20,0.15	4	0.93333	0.02165	0.96004	0.00612	0.001438	0.000151
20					0.98095	0.00612	0.001438	0.000149
30					0.97691	0.00741	0.001353	0.000143
40					0.97384	0.00837	0.001256	0.000135
50					0.97303	0.00863	0.001227	0.000132
100					0.97016	0.00953	0.001197	0.000129
150					0.96904	0.00988	0.001170	0.000127
200					0.96884	0.00995	0.001179	0.000126
10	1.05,0.35	4	0.69183	0.12068	0.91869	0.04286	0.019907	0.005154
20					0.81544	0.09348	0.011091	0.002119
30					0.78163	0.10899	0.007858	0.001337
40					0.76920	0.11519	0.006249	0.001017
50					0.76008	0.11936	0.005367	0.000885
100					0.74579	0.12550	0.003520	0.000499
150					0.74019	0.12789	0.002873	0.000397
200					0.73670	0.12940	0.002432	0.000321

Table 3: The model fitting summary for the considered data set I.

Model	MLE	L-L	AIC	BIC	KS	p value
FEXg	$\alpha = 21.8159, \theta = 0.0254$	-181.6155	367.2311	370.099	0.0882	0.9516
GED	$\alpha = 17.61869, \lambda = 74.04075$	-180.8895	365.779	368.647	0.0949	0.9178
WD	$\alpha = 3.1472, \lambda = 285.6358$	-181.827	367.6539	370.5219	0.0935	0.9259
TR	$\sigma = 211.4643, \lambda = 0.4999$	-187.7672	379.5345	382.4025	0.2605	0.0241
FD	$\theta = 2.907283, \sigma = 203.2133$	-183.7472	371.4944	374.3624	0.1304	0.6205
LD	$\theta = 0.0077$	-192.7683	387.5366	388.9706	0.2681	0.0185
IWD	$\alpha = 2.0226, \lambda = 47401.04$	-187.0328	378.0656	380.9336	0.2271	0.0690
AKD	$\theta = 0.0117$	-187.7058	377.4116	378.8455	0.2051	0.1276
XGD	$\theta = 0.0116$	-187.9498	377.8996	379.3336	0.20848	0.1166

Table 4: The model fitting summary for the considered data set II.

Model	MLE	L-L	AIC	BIC	KS	p value
FEXg	$\alpha = 483.4629, \theta = 0.4420$	-95.74962	195.4992	198.4307	0.11774	0.7667
GED	$\alpha = 86.2955, \lambda = 4.4804$	-96.73035	197.4607	200.3922	0.14751	0.4893
FD	$\theta = 4.3195, \sigma = 19.4864$	-99.71265	203.4253	206.3568	0.17762	0.2649
LD	$\theta = 0.0858$	-121.0507	244.1014	245.5671	0.3835	0.0001
AKD	$\theta = 0.1333$	-113.8608	229.7216	231.1873	0.32534	0.0022
XGD	$\theta = 0.1275$	-116.4718	234.9435	236.4093	0.35396	0.0006
IED	$\theta = 0.0469$	-131.497	264.9941	266.4598	0.49784	2.583e-07
ED	$\theta = 0.0447$	-131.4435	264.887	266.3527	0.46289	2.215e-06
IWD	$\alpha = 4.3195, \lambda = 372436.7$	-99.71265	203.4253	206.3568	0.17762	0.2649

Table 5: The model fitting summary for the considered data set III.

Model	MLE	L-L	AIC	BIC	KS	p value
FEXg	$\alpha = 95.3409, \theta = 2.5237$	-19.91846	43.83693	45.25303	0.13851	0.9358
EPD	$\beta = 2.1505, \eta = 3.8339$	-21.05967	46.11933	47.53543	0.24581	0.3250
FD	$\theta = 3.2302, \sigma = 2.2390$	-19.77290	43.5458	44.9619	0.1489	0.8952
WD	$\alpha = 3.1958, \lambda = 3.0951$	-19.85424	43.70848	45.12458	0.19547	0.6153
XGD	$\theta = 0.8192$	-28.26686	58.53372	59.24177	0.38825	0.0217
AKD	$\theta = 0.8803$	-26.25298	54.50595	55.21400	0.31111	0.1096
IED	$\theta = 0.4001$	-30.29933	62.59866	63.30671	0.40868	0.0133
ED	$\theta = 0.3605$	-30.30075	62.6015	63.30955	0.41776	0.0106
LD	$\theta = 0.5876$	-27.82462	57.64924	58.35729	0.3559	0.0447

Table 6: The model fitting summary for the considered data set IV.

Model	MLE	L-L	AIC	BIC	KS	p value
FEXg	$\alpha = 0.0851, \theta = 0.1608$	-40.4889	84.9779	86.394	0.1973	0.5385
XGD	$\theta = 0.4263$	-43.5558	89.11175	89.8198	0.2501	0.2585
LD	$\theta = 0.3198$	-41.42952	-84.8590	85.5671	0.2131	0.4424
AKD	$\theta = 0.5047$	-44.5656	91.1313	91.8394	0.2687	0.1905
IED	$\theta = 0.3853$	-38.0421	78.0843	78.7923	0.2222	0.3912
IXGD	$\theta = 3.7157$	-38.3578	78.7156	79.4237	0.2405,	0.2998

ILD	$\theta = 3.2110$	-37.84871	77.6974	78.4054	0.2118	0.4500
Pty2Lomax	$\theta = 20.6565, \alpha = 4.8074$	-39.9882	83.9765	85.3926	0.1979	0.5348
IP	$\theta = 0.0052, \alpha = 488.1109$	-38.0610	80.1221	81.5382	0.2209	0.3986
EPD	$\beta = 0.6095, \eta = 11.6948$	-42.9405	89.8811	91.2972	0.2471	0.2709

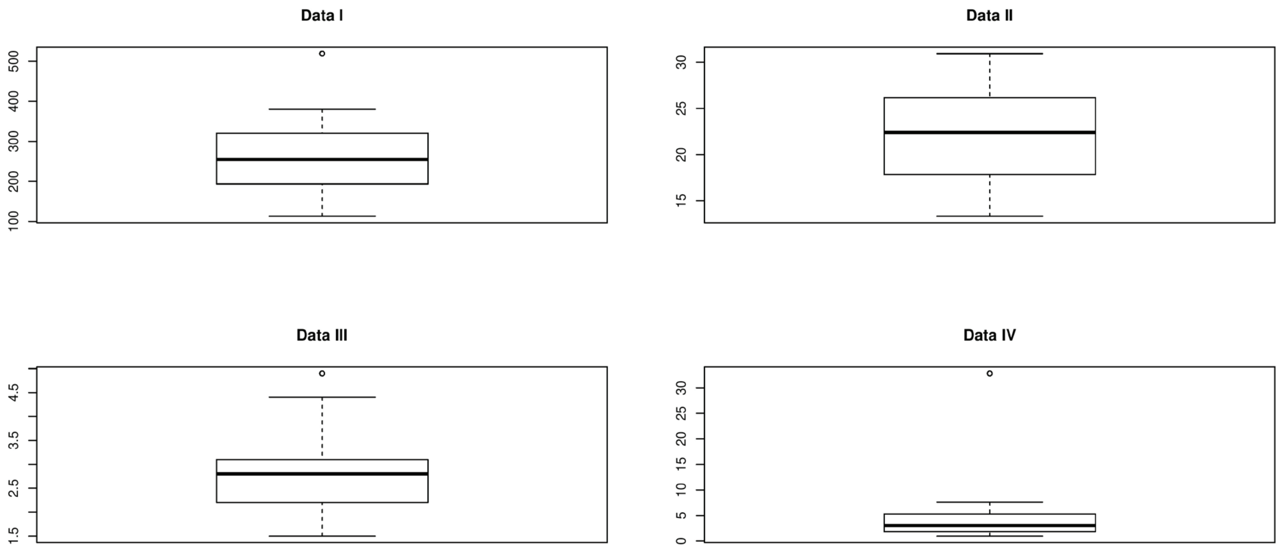
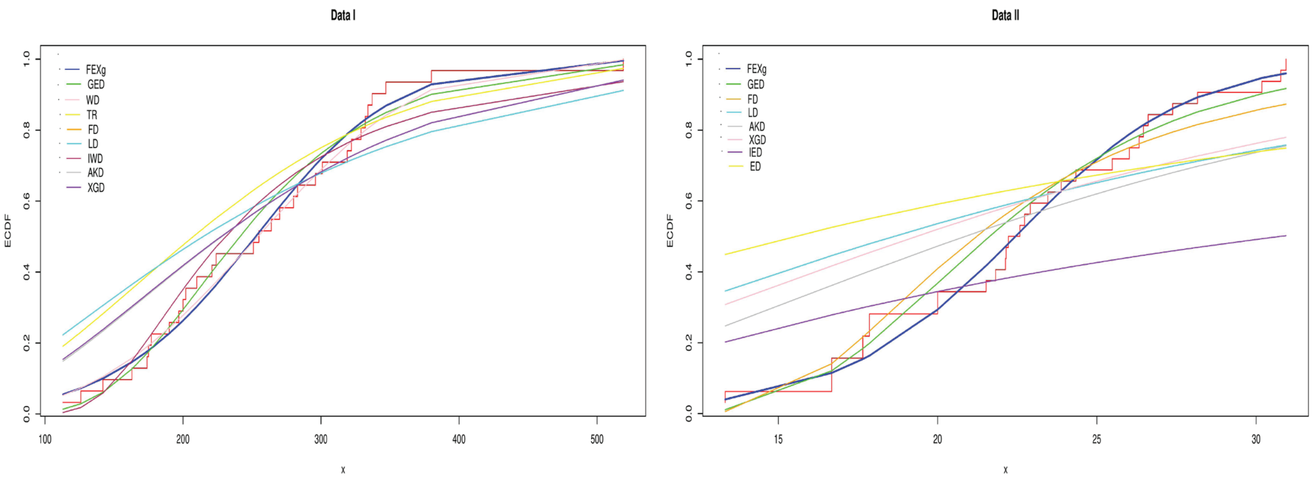


Figure 3: Boxplots of considered data set I, II, III, IV, V.



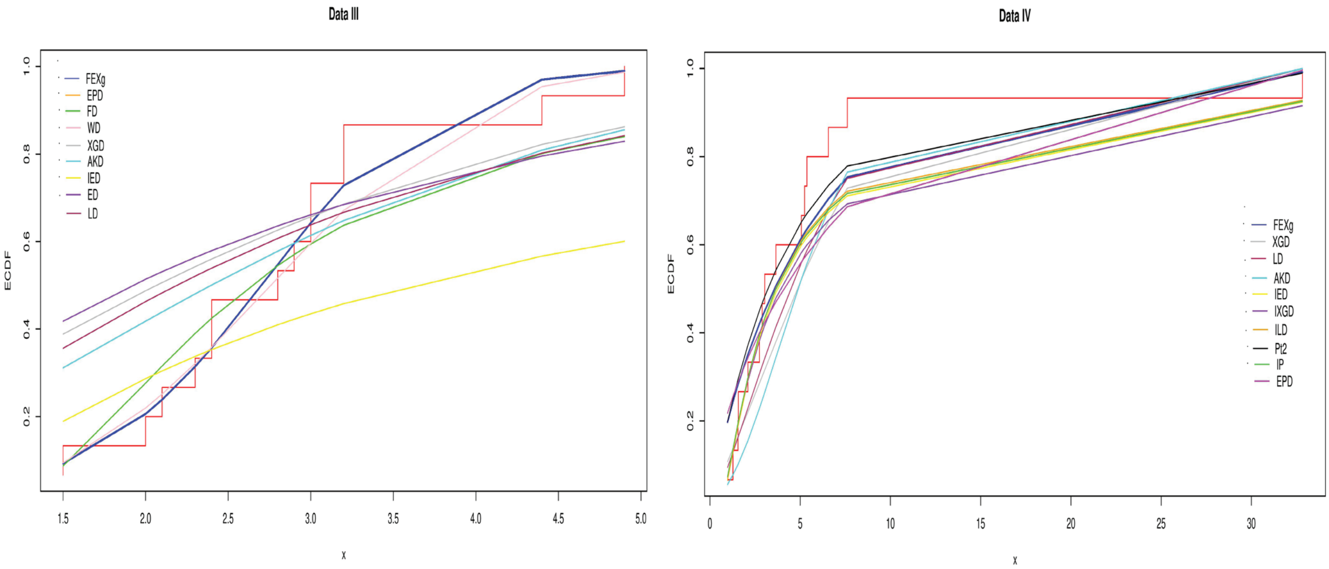


Table 4: ECDF plots of the considered data set I, II, III, IV, V.

Table 7: Descriptive summary for the considered data sets.

Data	Minimum	Q_1	Q_2	Mean	Q_3	Maximum	CS	CK
I	113	193.5	255	255.9	320	519.0	0.6954	3.8598
II	13.33	17.86	22.40	22.37	26.10	30.94	-0.0402	2.2894
III	1.500	2.200	2.800	2.773	3.100	4.90	0.7932	3.2362
IV	0.970	1.835	3.030	5.496	5.305	32.83	3.0901	11.4119

Table 8: Estimated survival function and hazard rate function for the considered data sets.

Data	$\hat{\alpha}_{mle}$	$\hat{\theta}_{mle}$	t	$\hat{\bar{G}}(t)_{mle}$	$\hat{H}(t)_{mle}$
I	21.8159	0.0254	50	0.99194	0.00036
			255.9	0.48893	0.01000
			100	0.95976	0.00102
II	483.2629	0.4420	50	2.265712e-05	0.40385
			22.37	0.50684	0.17893
			100	2.183881e-14	0.42246
III	95.34097	2.52378	50	3.443906e-50	2.48411
			2.773	4.659548e-01	1.04944
			100	2.149399e-104	2.50386
IV	0.08515374	0.1608966	50	0.00098	0.12722
			5.596	0.35637	0.17599
			100	1.107039e-06	0.14215