Nepal Journal of Mathematical Sciences (NJMS)
ISSN: 2738-9928 (*online*), 2738-9812 (*print*)
Vol. 2, No. 2, 2021 (August): 43-56
DOI: https://doi.org/10.3126/njmathsci.v2i2.40078

Research Article
Received Date: May 18, 2021
Accepted Date: June 25, 2021
Published Date: August 30, 2021

Exponentiated Marshall-Olkin Exponential Distribution: Application of COVID-19 Second Wave in Nepal

Ramesh Prasad Tharu[1], Sakar Pahari[2], Gopal Prasad Sedhai[3] & Govinda Prasad Dhungana[4*]

[1]*Department of Statistics, Tribhuvan University, Mahendra Multiple Campus, Nepalgunj, Nepal*
[2,] *Graduate Research Assistant, Dept of Civil and Environmental Engineering ,Michigan State University; USA*

[3]*Department of Statistics, Tribhuvan University, Prithwi Narayan Campus, Pokhara, Nepal*
[4]*Department of Statistics Tribhuvan University; Birendra Multiple Campus, Bharatpur Chitwan, Nepal*

Email:[1]*rameshpt* 02*@gmail.com*,[2]*psakar*007*@gmail.com*,[3]*gsedhai*45*@gmail.com*,

[4]*dhunganagovindana*2012*@gmail.com*

Corresponding Author: ***Govinda Prasad Dhungana***

Abstract: *The present study aims to generate the new flexible probability distribution from classical probability distribution by introducing one or more additional parameter (s). Recently, many methods (or techniques) have been proposed to develop and extend new univariate continuous probability models. The aim of this study is to explore the new univariate continuous Exponentiated Marshall –Olkin Exponential distribution to predict mortality rate of COVID-*19 *second wave in Nepal. To develop the proposed model, the exponential distribution is compounded with Marshall-Olkin family of distribution. Explicit expressions of reliability/survival function, hazard rate function, revised hazard rate function, quintile function, asymptotic behavior, moments, residual life function, R`enyi entropy and q entropy, probability weighted moment and order statistics of proposed model. The data were collected from April* 1 *to May* 14 *from MOHP Nepal regarding the COVID-*19 *second wave. The parameters* $\hat{\alpha}$, $\hat{\beta}$ *and* $\hat{\theta}$ *are estimated by maximum likelihood methods and predict the mortality rate per day by COVID-*19 *second wave. The COVID-*19 *second wave data set is used to validate the proposed model. This model is valid statistical tools like as P-P and Q-Q plots; KS, Anderson darling test and Cramer von miss test; and value of log-likelihood, AIC, BIC and CAIC. Hence, proposed model is a satisfactory model to predict the events. From this model daily predicted mortality rate was quite high (*94.61*%). As a result, if the current situation persists, there is a high risk of up to* 150 *people dying per day in Nepal.*

Keywords: *COVID-*19*, Exponentiated Marshall-Olkin exponential, Maximum likelihood estimation, Order statistics*

1. Introduction

In the last few decades, numerous approaches (or techniques) have been proposed for the development and expansion of new family of distributions or probability distribution in different applied areas. In the literature, some familiar univariate family of distribution like as, muth generated family of distributions [1], Kumaraswamy type I half logistic family of distributions [11], the new Weibull-G family [32], bivariate odd Weibull-G family [9] and Generalized Odd Gamma-G family [15] and many others. After introducing the family of distribution, the new flexible probability model can be derived by adding one or more auxiliary parameters, so that it is better fitted to a huge amount of data in the different applied fields [20]. Likewise, inclusion of one or more additional shape parameter(s) to the baseline distribution, it makes the distribution more flexible, particularly when studying the tail properties [19, 31].

In literature, various probability distributions have been derived from different family of distributions in the last decade. Risti' Miroslav M. et al. [29] introduced a three-parameter distribution of Marshall Olkin Generalized Exponential (MOGE). This distribution is very flexible because the shape of the hazard function increases, decreases, the bathtub or the inverted bathtub. This model has good statistical behavior and is satisfactory in the analysis of real data. Similarly, UL Haq, et al. [33] expanded the distribution of Length-biased exponential (LBE) and proposed a two-parameter distribution of Marshall-Olkin Length-Biased Exponential (MOLBE). This distribution is better fitted than a number of other competitive models and tested empirically using real life data. Likewise, the three-parameter and unimodal, Marshall-Olkin Logistic-Exponential (MOLE) was introduced by Mansoor M, et al. [22]. The characteristic of the hazard rate function is increased, decreased, bathtub and upside down, which indicates that the distribution is more applicable to the real data analysis. Similarly, other different distributions were also suggested by different authors, such as Marshall–Olkin Extended Weibull by Ghitany et al. [12], Marshall-Olkin Exponential Weibull by Pogány et al. [26], Marshall–Olkin Gamma–Weibull by Saboor et al. [30], Marshall-Olkin Fréchet by Krishna et al. [17], a new Poisson Inverted Exponential distribution from the Poisson family of distribution Dhungana [8] and Exponentiated Rayleigh Poisson distribution by Joshi & Dhungana [16]. Hence, the concept of generating the new distribution is introducing the flexible distribution which has applicable in different fields like health, engineering, demography, survival analysis, and reliability analysis.

The COVID-19 pandemic has devastated the world and has accompanied economic, social and behavioral challenges and responses. More than 160 million cases have been confirmed worldwide as of now and the disease has claimed over 3.3 million lives [10]. Likewise, more than 4,500 people have died in Nepal by 14th May, 2021 [13]. The situation seemed to be within control in Nepal in early March, at one point only 47 daily cases being reported in 6th March, 2021 [13]. But by mid-April, 14 districts, including the Kathmandu Valley, were declared as Covid-19 hotspots due to the second wave (The Times of India 2021, April 15). The total reported cases and total deaths due to Covid-19 from 1st April to 14th May account to almost 37% of total reported cases and 35% of total deaths since the first case was reported in Nepal back in January of 2020 [3].

Therefore, the aim of this study is to generate the more flexible probability distribution and data analysis of COVID-19 second wave of Nepal from this distribution.

2. Material and Methods

Initially, Marshall and Olkin (1997) suggested a new family of distribution having cumulative distribution functions (cdf);

$$F(x) = \frac{G(x)}{G(x) + \beta\{1 - G(x)\}}; x > 0, \beta > 0. \tag{1}$$

The family of distribution has been expanded by introducing an auxiliary parameter, and then it becomes exponentiated or generalized distribution [7, 31]. Let, $\theta > 0$, has an auxiliary parameter has been added in proposed model, called Exponentiated Marshall and Olkin family of distribution, which is

$$F(x) = \left[\frac{G(x)}{G(x) + \beta\{1 - G(x)\}}\right]^{\theta}; x > 0, \beta > 0, \theta > 0. \tag{2}$$

Where, $G(x)$ is cdf of baseline distribution and β is a scale and θ is shape parameter. In probability theory and statistics, exponential distribution plays a significant role in the study of survival results. This is the probability distribution of time between events in the Poisson point process, i.e., a process in which events occur continuously and independently at a constant average rate. This is a special case of the distribution of gamma. It's a continuous analog of geometric distribution, which has the main property of being memoryless. In addition to being used for the study of Poisson point systems, it can

be found in a number of other ways. So, the exponential distribution is used as a base line probability distribution having cdf;

$$G(x) = 1 - e^{-\alpha x}; x > 0, \alpha > 0. \tag{3}$$

For proposed model, cdf of the equation (3) is compounded with equation (2), then cdf of Exponentiated Marshall –Olkin Exponential (EMOE) distribution is

$$F(x) = \left[\frac{1 - e^{-\alpha x}}{1 + (\beta - 1)e^{-\alpha x}}\right]^{\theta}; x > 0, \theta > 0, \alpha > 0, \beta > 0. \tag{4}$$

Now, differentiation of equation (4) with respect to x, then pdf of proposed model is

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

The shape of the distribution depends upon the value of parameters. This distribution is positively skewed and unimodal (Fig.1, left panel).

Similarly, the survival function of proposed model is

$$R(x) = 1 - F(x) = \frac{\left[1 + (\beta - 1)e^{-\alpha x}\right]^{\theta} - \left[1 - e^{-\alpha x}\right]^{\theta}}{\left[1 + (\beta - 1)e^{-\alpha x}\right]^{\theta}}. \tag{6}$$

The hazard rate function is the conditional density given that the event has not yet occurred before time x. Mathematically, let x be a survival time of a component or item and we want to calculate the probability that it will not survive for an additional time Δx, then hazard rate function is $h(x) = \lim_{\Delta x \to 0} \frac{prob(x \le X \le \Delta x)}{\Delta x . R(x)} = \frac{f(x)}{1 - F(x)}; x > 0$.

Therefore, the hazard rate function is

$$h(x) = \frac{\alpha\beta\theta e^{-\alpha x}\left[1 - e^{-\alpha x}\right]^{\theta - 1}}{\left[1 + (\beta - 1)e^{-\alpha x}\right]\left[\left\{1 + (\beta - 1)e^{-\alpha x}\right\}^{\theta} - \left\{1 - e^{-\alpha x}\right\}^{\theta}\right]}. \tag{7}$$

The shape of hazard function is increasing, inverted bathtub in nature. This reflects the good statistical behavior in real data modeling (Fig. 1, right panel).

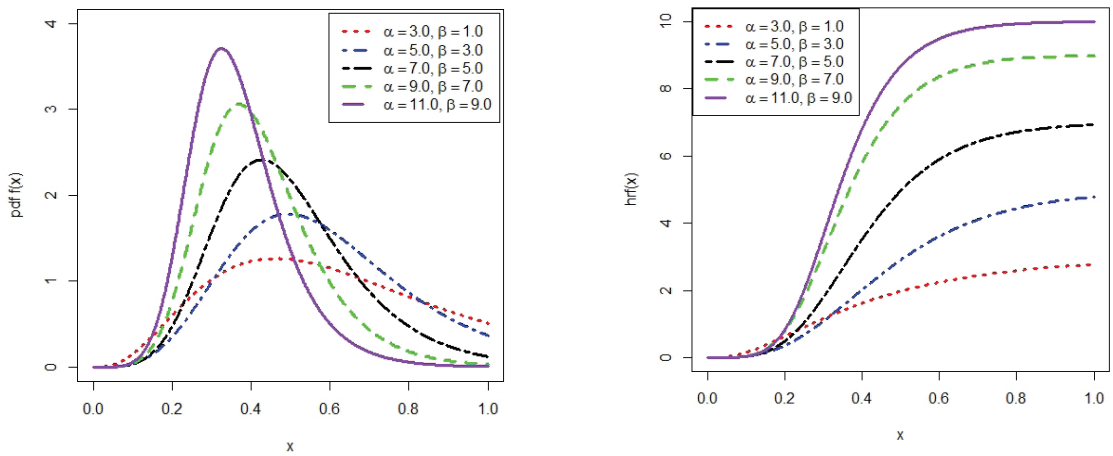

Figure 1: Plot of probability density function (left panel), hazard rate function (right panel) with different value of parameters α and β when $\theta = 4$

The reversed hazard function is;

$$r(x)=\frac{f(x)}{F(x)}=\frac{\alpha\beta\theta e^{-\alpha x}}{\left(1-e^{-\alpha x}\right)\left[1+(\beta-1)e^{-\alpha x}\right]}. \tag{8}$$

Special case:

When θ =1, the proposed model becomes MOEE which was introduced by marshal and Olkin. (1997). The pdf of model is

$$f(x)=\frac{\alpha\lambda e^{-\lambda x}}{\left[1-(1-\lambda)e^{-\lambda x}\right]^2};x>0,\alpha>0,\lambda>0. \tag{9}$$

3. Statistical Properties

In this section major properties of EMOE distribution have been derived.

3.1 Useful expansions

Distribution is derived from the generalized binomial series. For, $|z|<1,\ n>0$, we have;

$$(1+z)^{-n}=\sum_{i=0}^{\infty}(-1)^i\binom{n+i-1}{i}z^i. \tag{10}$$

$$(1-z)^{n}=\sum_{j=0}^{\infty}(-1)^j\binom{n}{j}z^j. \tag{11}$$

Using the binomial theorem (10) and (11) in equation (5), the pdf of proposed model is

$$f(x)=\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\xi_{ij}e^{-(1+i+j)\alpha x} \tag{12}$$

$$\text{Where, } \xi_{ik}=(-1)^{i+j}(\beta-1)^i\alpha\beta\theta\binom{\theta+i}{i}\binom{\theta-1}{j}.$$

The expansion of cumulative function $\left[F(x)\right]^s$, by used equation (10) and (11), then,

$$\left[F(x)\right]^s=\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\phi_{k\ l(s)}e^{-(k+l)\alpha x} \tag{13}$$

$$\text{Where, } \phi_{k\ l(s)}=(-1)^{k+l}\left(\beta-1\right)^l\binom{\theta s+l-1}{l}\binom{\theta s}{k}.$$

Likewise, the expansion of $\left[f(x)\right]^\delta,\delta>0$ is

$$\left[f(x)\right]^\delta=\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\xi_{ij(\delta)}e^{-(\delta+i+j)\alpha x} \tag{14}$$

$$\text{Where, } \xi_{ij(\delta)}=(-1)^{i+j}\left(\alpha\beta\theta\right)^\delta\left(\beta-1\right)^i\binom{\delta(\theta+1)+i-1}{i}\binom{\delta(\theta-1)}{j}.$$

3.2 Asymptotic Behavior

We investigate the behavior of the proposed model as given in equation (5), as $x\to 0$ *and as* $x\to\infty$. Therefore,

$$\lim_{x\to 0}\frac{\alpha\beta\theta e^{-\alpha x}\left[1-e^{-\alpha x}\right]^{\theta-1}}{\left[1+(\beta-1)e^{-\alpha x}\right]^{\theta+1}}=\lim_{x\to\infty}\frac{\alpha\beta\theta e^{-\alpha x}\left[1-e^{-\alpha x}\right]^{\theta-1}}{\left[1+(\beta-1)e^{-\alpha x}\right]^{\theta+1}}=0.$$

These findings confirm that the proposed distribution has a unique mode value. The required necessary and sufficient condition for mode is;

$\frac{df(x)}{dx}=0$ and $\frac{d^2f(x)}{dx^2}<0$.

After applying necessary condition, we get

$$-\alpha+\frac{(\theta-1)\alpha}{\left(e^{\alpha x}-1\right)}+\frac{\alpha(\theta+1)(\beta-1)}{\left[e^{\alpha x}+(\beta-1)\right]}=0. \tag{15}$$

The equation (15) is a non-linear equation. The solution of the equation (15) cannot be determined analytically. The value of the mode can be determined by numerically using the Newton-Raphson method (Table 1).

3.3 Quantile Function

The quantile function is defined as, $Q(u)=F^{-1}(u)$ which as an alternative of pdf and cdf of any probability distribution. It is used to obtain statistical calculations such as mean, median, mode, skewness and kurtosis by generating random numbers. The corresponding quantile function for the proposed model is

$$Q(u)=-\frac{1}{\alpha}\left[\ln\left\{\frac{1-u^{1/\theta}}{1+(\beta-1)u^{1/\theta}}\right\}\right]. \tag{16}$$

Where, $u\sim U$ (0, 1) distribution. The 100 random samples are generated from the equation (16). It is observed that proposed distribution is unimodal, skewed and non-normal (Table 1, Figure 1, left panel). Therefore, the proposed distribution can be used in univariate data analysis either longitudinal or cross-sectional, in different areas such as engineering, actuarial, environmental, medical, biological, demographic, economics and many others applied fields.

Table 1: The mean, median, mode, skewness, Kurtosis of proposed distribution with difference value of parameters

Parameters			Mean	Median	Mode	Skewness	Kurtosis
α	β	θ					
0.1	0.7	1.0	0.12558	0.12990	0.13668	-2.06518	9.62200
0.2	0.8	1.5	0.10537	0.11090	0.08399	-0.74273	2.48239
0.3	0.9	2.0	0.08904	0.08580	0.04675	-0.02866	1.80272
0.4	1.0	2.5	0.07865	0.06435	0.02512	-0.44559	1.99554
0.5	1.1	3.0	0.07256	0.04907	0.01329	-0.80408	2.44920
0.6	1.2	3.5	0.12478	0.03714	0.04371	1.07558	2.94762
0.7	1.3	4.0	0.06841	0.02792	0.00366	1.27667	3.38691
0.8	1.4	4.5	0.06841	0.02042	0.00192	1.42064	3.72947
0.9	1.5	5.0	0.07058	0.01489	0.00101	1.52150	3.97769
1.0	1.6	5.5	0.07310	0.01136	0.00053	1.59184	4.15378

3.4 Moments

Let $X\sim EMOE$ (α,β,θ), the rth raw moment about the origin is defined as;

$$\mu_r'=E(X^r)=\int_0^\infty x^r f(x)dx=\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\xi_{ij}\int_0^\infty x^r e^{-(1+i+j)\alpha x}dx. \tag{17}$$

After integration of (17), the rth raw moment about the origin of proposed model is

$$\mu_r'=\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\xi_{ij}\frac{\Gamma(r+1)}{[\alpha(1+i+j)]^{r+1}}.$$

When $r=1$, mean of proposed model is $\mu_1'=\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\xi_{ij}\frac{1}{[\alpha(1+i+j)]^2}$.

Similarly, second raw moment about the origin is $\mu_2'=\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\xi_{ij}\frac{2}{[\alpha(1+i+j)]^3}$.

Therefore, we can calculate the variance as the relation $V(x) = \mu_2' - (\mu_1')^2$

The incomplete moments, say, $\varphi_s(t)$, is given by; $\varphi_s(t) = \int_0^t x^s f(x)dx.$ (18)

Used the relation (12) in equation (18), and applied lower incomplete gamma function, $\gamma(s,t) = \int_0^t x^{s-1}e^{-x}dx$ and after integration equation (18) then incomplete moment $\varphi_s(t)$ is

$$\varphi_s(t) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\xi_{ij}\frac{\gamma\{(s+1),\alpha(1+i+j)t\}}{[\alpha(1+i+j)]^{(s+1)}}.$$

Similarly, the conditional moment is defined as $\tau_s(t) = \int_t^{\infty} x^s f(x)dx$ (19)

Used the relation (12) in equation (19), and applied upper incomplete gamma function, $\Gamma(s,t) = \int_t^{\infty} x^{s-1}e^{-x}dx$ and after integration equation (19) then conditional moment is

$$\tau_s(t) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\xi_{ij}\frac{\Gamma\{(s+1),\alpha(1+i+j)t\}}{[\alpha(1+i+j)]^{(s+1)}}.$$

Likewise, Moment Generating Function (MGF) is

$$M_X(t) = E[e^{tX}] = \sum_{r=0}^{\infty}\frac{t^r}{r!}E(X^r). \quad (20)$$

Using result of the equation (17) in equation (20), then MGF is

$$M_X(t) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{r=0}^{\infty}\xi_{ij}\frac{t^r\Gamma(r+1)}{r![\alpha(1+i+j)]^{r+1}}.$$

3.5. Residual Life Function

The nth moment of the residual life of X is given by;

$$m_n(t) = \frac{1}{R(t)}\int_t^{\infty}(x-t)^n f(x)dx. \quad (21)$$

Apply the binomial expansion of $(x-t)^n = \sum_{r=0}^{n}(-1)^r\binom{n}{r}x^{n-r}t^r$ into the equation (21) and substitute the value of $f(x)$ from equation (12) then equation (21) can be express as;

$$m_n(t) = \frac{1}{R(t)}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{r=0}^{n}(-t)^r\xi_{ij}\binom{n}{r}\int_t^{\infty}x^{n-r}e^{-(1+i+j)\alpha x}dx. \quad (22)$$

Apply the upper incomplete gamma function $\Gamma(s,t) = \int_t^{\infty} x^{s-1}e^{-x}dx$ in equation (22), n^{th} moment of the residual life of X is

$$m_n(t) = \frac{1}{R(t)}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{r=0}^{n}(-t)^r\xi_{ij}\binom{n}{r}\frac{\Gamma[(n-r+1),\alpha(1+i+j)t]}{[\alpha(1+i+j)]^{(n-r+1)}}.$$

3.6. R`enyi and *q*-entropies

The entropy of a random variable X is a measure of variation of uncertainty and has been used in many fields such as physics, engineering and economics among others. The R`enyi entropy is defined as;

$$I_\delta(X) = \frac{1}{1-\delta}\log\int_{-\infty}^{\infty}\left[f(x)\right]^\delta dx, \quad \delta > 0 \,\&\, \delta \neq 1. \tag{23}$$

Used equation (13) in equation (23) and integration of this equation, R`enyi entropy is

$$I_\delta(\mathrm{X}) = \frac{1}{1-\delta}\log\left[\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\xi_{ij}\frac{1}{\alpha(\delta+i+j)}\right].$$

Similarly, q-entropy is defined as;

$$H_q(x) = \frac{1}{1-q}\log\left[1-\int_{-\infty}^{\infty}\left\{f(x)\right\}^q\right], q > 0 \;\&\; q \neq 1. \tag{24}$$

The q-entropy is obtained by substituting the result (23) into (24), where δ is replaced by q;

$$H_q(x) = \frac{1}{1-q}\log\left[1-\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\xi_{ij}\frac{1}{\alpha(\delta+i+j)}\right].$$

3.7 The Probability Weighted Moments (PWM)

The probability weighted moments can be obtained from the following relation

$$\tau_{r,s} = E(X^r F(x)^s] = \int_{-\infty}^{\infty} x^r f(x)F(x)^s dx. \tag{25}$$

By substituting equations (12) and (13) into (25), then equation (25) becomes

$$\tau_{r,s} = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\xi_{ij}\phi_{k\;l(s)}\int_0^{\infty} x^r e^{-(1+i+j+k+l)\alpha x}dx. \tag{26}$$

After integration the equation (26), Hence, the PWM of proposed model is

$$\tau_{r,s} = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\xi_{ij}\phi_{k\;l(s)}\frac{\Gamma(r+1)}{\left[\alpha(1+i+j+k+l)\right]^{r+1}}.$$

3.8 Order Statistics

Order statistics have been extensively applied in many fields of statistics, such as reliability and life testing. Let $X_1, X_2, ..., X_n$ be independent and identically distributed random variables with their corresponding cumulative distribution function F(x). If these variable are arrange in ascending order of magnitude then can be written as $X_{(1)} \leq X_{(2)} ... \leq X_{(n)}$, we call $X_{(j)}$ as the jth order statistics, $j = 1, 2, ...n$.The pdf of jth order statistic David (1981), is defined as;

$$f\left(X_{(x)}(x_{(j)})\right) = \frac{f(x_{(j)})}{B(j, n-j+1)}\sum_{\upsilon=0}^{n-j}(-1)^\upsilon\binom{n-j}{\upsilon}F(x_{(j)})^{\upsilon+j-1}. \tag{27}$$

Where, $B(.,.)$ is the beta function. By substituting (12) the equation (27), and (13) in (27), where replacing s by υ+j-1. Again let, $\eta^* = (-1)^\upsilon\binom{n-j}{\upsilon}$, then, the pdf of proposed model in term of order statistics is

$$f\left(X_{(x)}(x_{(\mathrm{j})})\right) = \frac{1}{B(\mathrm{j}, \mathrm{n}-j+1)}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\sum_{\upsilon=0}^{n-j}\eta^*\xi_{ij}\phi_{k\;l(\upsilon+j-1)}\int_0^{\infty} e^{-(1+i+j+k+l)\alpha x_{(j)}}dx_{(j)}. \tag{28}$$

The rth raw moment of order statistics is defined as;

$$E\left(X_{(j)}^r\right) = \int_0^{\infty} x_{(j)}^r f\left(X_{(x)}\left(x_{(j)}\right)\right)dx_{(j)}. \tag{29}$$

After integration of equation (29) after substitute the value of equation (28), then finding of rth raw moment is

$$E(X_{(j)}^{r}) = \frac{1}{B(j,n-j+1)} \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\sum_{\upsilon=0}^{n-j} \eta^{*}\xi_{ij}\phi_{kl(\upsilon+j-1)} \frac{\Gamma(r+1)}{\left[\alpha(1+i+j+k+l)\right]^{r+1}}.$$

4. Maximum Likelihood Estimation

The maximum likelihood estimates (MLEs) of the unknown's parameters of the distribution based on $\underset{\sim}{x} = (x_1,\ldots,x_n)$ observed sample value with of set of parameters $\ell(\alpha,\beta,\theta \mid \underset{\sim}{x})$ .The log likelihood function of the parameter $\ell(\alpha,\beta,\theta) = \ell$ is given by;

$$\ell = n\ln(\alpha\beta\theta) - \alpha\sum_{i=1}^{n} x_i + (\theta-1)\left[\sum_{i=1}^{n}\ln\left(1-e^{-\alpha x_i}\right)\right] - (\theta+1)\left[\sum_{i=1}^{n}\ln\left\{1+(\beta\text{-}1)e^{-\alpha x_i}\right\}\right]. \quad (30)$$

MLE is a technique to determine the value of the parameters by differentiating w.r.t.to parameters and equating to zero, Let $\omega_i = \frac{1}{(e^{\alpha x_i}-1)}$; and $\xi_i = \frac{1}{\left[e^{\alpha x_i}+(\beta-1)\right]}$

$$\frac{\partial\ell}{\partial\alpha} = \frac{n}{\alpha} - \sum_{i=1}^{n} x_i + (\theta-1)\sum_{i=1}^{n}\omega_i x_i + (\theta+1)(\beta-1)\sum_{i=1}^{n}\xi_i x_i = 0. \quad (31)$$

$$\frac{\partial\ell}{\partial\beta} = \frac{n}{\beta} - (\theta+1)\sum_{i=1}^{n}\xi_i = 0. \quad (32)$$

$$\frac{\partial\ell}{\partial\theta} = \frac{n}{\theta} + \sum_{i=1}^{n}\ell n(1-e^{-\alpha x_i}) - \sum_{i=1}^{n}\ell n\left[1+(\beta-1)e^{-\alpha x_i}\right] = 0. \quad (33)$$

To estimate the unknown parameters α, β and θ , we need to solve non-linear equation (31), (32) and (33). It is clear that, these equations cannot solve analytically. Therefore, we estimate the value of unknown parameters by applying the Newton-Raphson's iterative technique in log-likelihood function of equation (30) directly, using optim () function in R software [3, 23].

Let us denote the parameter vector by $\underset{\sim}{\delta} = (\alpha,\beta,\theta)$ and corresponding MLE of $\underset{\sim}{\delta}$ as $\hat{\underset{\sim}{\delta}} = (\hat{\alpha},\hat{\beta},\hat{\theta})$, then asymptotic normality result is $\left(\hat{\underset{\sim}{\delta}} - \underset{\sim}{\delta}\right) \to N_3\left(0,\left(I(\underset{\sim}{\delta})\right)^{-1}\right)$. Where, $I(\underset{\sim}{\delta})$ is fisher information matrix. By applying the Newton-Raphson algorithm in equation (30), it produced the observed information matrix $O(\hat{\underset{\sim}{\delta}})$ (elements given in Appendix-A), which is estimated matrix of fisher information matrix. Therefore, observed information matrix $O(\hat{\underset{\sim}{\delta}})$ is used in the common procedure, which can be written as;

$$O(\hat{\underset{\sim}{\delta}}) = -\begin{pmatrix} \frac{\partial^2\ell}{\partial\alpha^2} & \frac{\partial^2\ell}{\partial\alpha\,\partial\beta} & \frac{\partial^2\ell}{\partial\alpha\,\partial\theta} \\ \frac{\partial^2\ell}{\partial\alpha\,\partial\beta} & \frac{\partial^2\ell}{\partial\beta^2} & \frac{\partial^2\ell}{\partial\beta\,\partial\theta} \\ \frac{\partial^2\ell}{\partial\alpha\,\partial\theta} & \frac{\partial^2\ell}{\partial\beta\,\partial\theta} & \frac{\partial^2\ell}{\partial\theta^2} \end{pmatrix}_{|(\hat{\alpha},\hat{\beta},\hat{\theta})} = -H(\underset{\sim}{\delta})_{|\underset{\sim}{\delta}=\hat{\underset{\sim}{\delta}}}. \quad (34)$$

Where, H is the hessian matrix, $\underset{\sim}{\delta} = (\alpha, \beta, \theta)^T$ as a parameter and $\hat{\underset{\sim}{\delta}} = (\hat{\alpha}, \hat{\beta}, \hat{\theta})^T$ as estimated value of corresponding parameters. The inverse of hessian matrix is called variance –covariance matrix which can be expressed as;

$$\left(-H(\delta)_{\big|\delta=\hat{\delta}}\right)^{-1} = \begin{pmatrix} \mathrm{var}(\hat{\alpha}) & \mathrm{cov}(\hat{\alpha},\hat{\beta}) & \mathrm{cov}(\hat{\alpha},\hat{\theta}) \\ \mathrm{cov}(\hat{\alpha},\hat{\beta}) & \mathrm{var}(\hat{\beta}) & \mathrm{cov}(\hat{\beta},\hat{\theta}) \\ \mathrm{cov}(\hat{\alpha},\hat{\theta}) & \mathrm{cov}(\hat{\beta},\hat{\theta}) & \mathrm{var}(\hat{\theta}) \end{pmatrix} \tag{35}$$

Finally, we construct approximate $100(1-\gamma)$ % confidence interval for α, β and θ is

$$\hat{\alpha} \pm z_{\gamma/2}\sqrt{\mathrm{var}(\hat{\alpha})}\ ;\ \hat{\beta} \pm z_{\gamma/2}\sqrt{\mathrm{var}(\hat{\beta})} \text{ and } \hat{\theta} \pm z_{\gamma/2}\sqrt{\mathrm{var}(\hat{\theta})} \tag{36}$$

Where, $Z_{\gamma/2}$ is the upper percentile of standard normal variate.

5. Result and Discussion

The COVID 19 is a worldwide pandemic disease that occurred in Nepal, causing severe acute respiratory syndrome. COVID 19's various waves are on display all over the world. The syndrome of the second wave of the COVID-19 pandemic was commonly observed in Nepal during the first week of April. The death rate in Nepal was high following the second wave of COVID 19. The number of death cases in Nepal on the first day of April was only one; however, this situation has steadily increased, and by the end of the 14th of May, the number of deaths in Nepal was 203 per day. As a result, this study concentrated on the prediction of deaths in Nepal after starting with the second wave (1st April to 14th May). The data (at least one death per day from 1st April to 14th May) were reported in the following (Government of Nepal Ministry of Health and Population, 2021).
(1, 1, 4, 2, 1, 1, 13, 5, 3, 5, 4, 5, 8, 8, 11, 10, 5, 5, 14, 28, 12, 18, 17, 35, 33, 19, 27, 37, 55, 58, 54, 50, 53, 88, 139, 225, 168, 214, 203)

Table 2: The summary of number of death due to COVID 19

Statistical Measure	Minimum	Q_1	Median	Mean	Q_3	Maximum
Number of deaths	1	5	14	43	52	225

From the finding on the average 43 person/day died in Nepal due to COCID 19 second wave. The shape of the data was highly skewed (box plot) and can be predicted by an alternative probability model (TTT plot). Hence, this data has been predicted from our proposed model (EMOE).

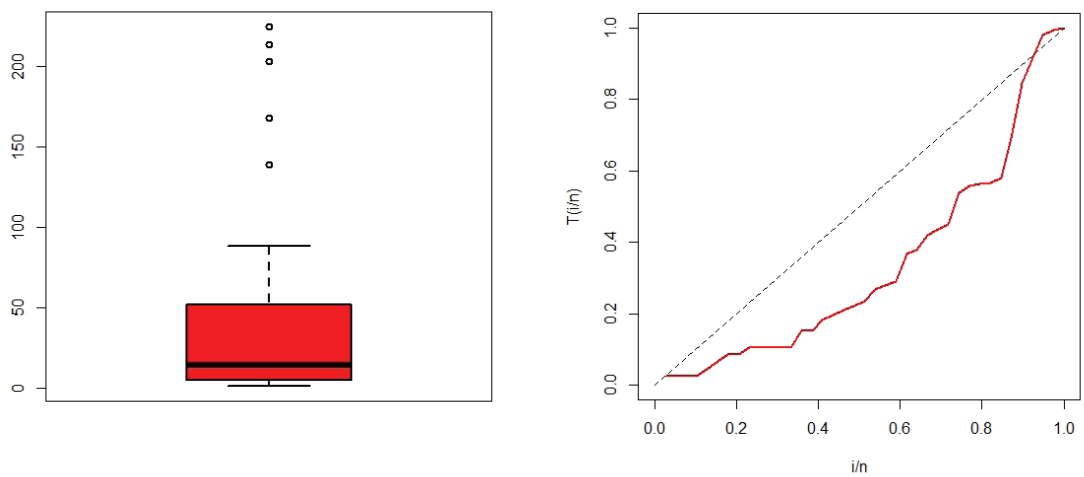

Figure 2: *Box plot (left panel) and TTT plot* (*right panel*)

5.2 Parameter Estimation

We have computed the ML estimates by maximizing the log-likelihood function given in equation (30) directly using optim () function in R software, Braun et al. (2016), R Core Team (2020). Finally, we have to present the value of $\left(\hat{\alpha}, \hat{\beta} \text{ and } \hat{\theta}\right)$ with standard error (SE) (Table 3).

Table 3: Estimated value of parameters

Parameters	MLEs	SE	p-value
$\hat{\alpha}$	0.00728	0.00549	0.185
$\hat{\beta}$	0.09596	0.11430	0.401
$\hat{\theta}$	1.16686	0.47487	0.014

5.3. Model Validation

The probability-probability (*P-P*) plots and quantile -quantile (*Q-Q*) plots can be inspected to verify the model either valid or not. The Q-Q plot may provide information about the lack-of-fit at the tails of the distribution, whereas the *P-P* plot emphasizes the lack-of-fit. Therefore, proposed model has good fit of theoretical distribution versus empirical distribution in both plots (Fig 2).

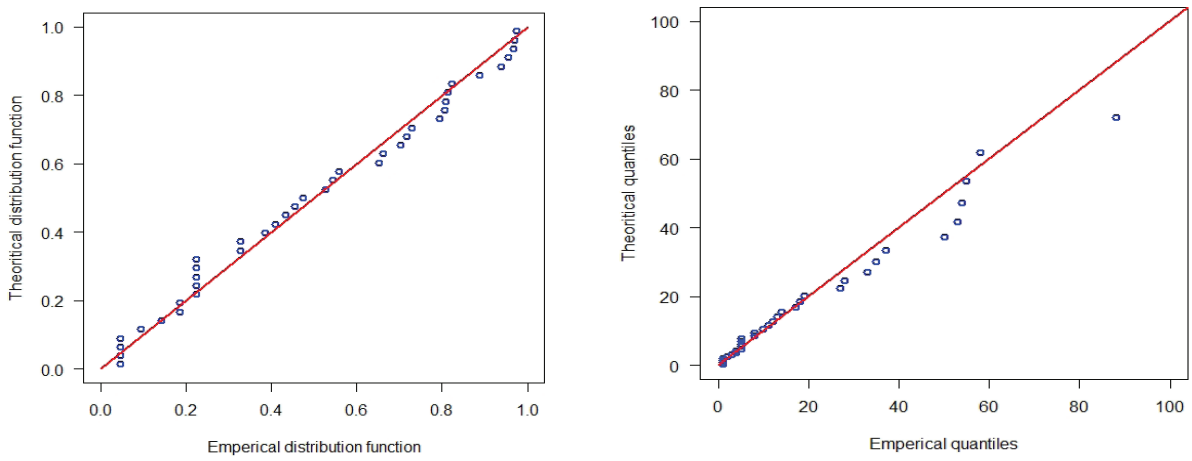

Fig 2: ***P-P plot (left panel) and Q-Q plot (right panel) of proposed model***

Further, validation of the model, we compute the Kolmogorov-Smirnov (KS), Anderson Darling and Cramer-von Mises test. It measured distance between the Empirical Distribution Function (edf) and fitted distribution function. The, *p*-value of KS test, A^2 test and Cramer-von Mises test was 0.752 (D = 0.10811), 0.8273 (D = 0.42053) and 0.8824(D = 0.049439) respectively. Hence, p-value > 0.05 of each statistics indicates that it supports the null hypothesis of goodness of fit. Hence, proposed model satisfactorily fits for further data analysis, Kumar and Ligges (2011).

5.4 Prediction of Death Rate

The proposed model is appropriate for predicting the death rate of Nepal as a result of the COVID 19 second wave. Therefore, we must use the proposed model to predict the various mortality rates. The daily mortality rate was predicted by our proposed model is high (94.61%). As a result, if the current situation persists, there is a high risk of up to 150 people dying per day in Nepal (Table 4).

Table 4: Prediction of death rate due to COVID 19 second wave in Nepal

Number of deaths	0–50	50–100	100–150	150–200	200–250	250–300
Probability	0.7940	0.1106	0.04154	0.02071	0.01177	0.00719

5.5 Model Comparisons

We have considered four alternative models named Generalized Inverted Generalized Exponential (GIGE) by Oguntunde et al. [20], Exponentiated Half Logistic Exponential (EHLE) by Almarashi et al. [2], Exponentiated Generalized Inverted Exponential (EGIE) by Oguntunde et al. [25], Half Logistic Exponential (EHLE) by Almarashi et al. [1] and Exponentiated Inverse Rayleigh (EIR) by [28]. These models are compared with proposed model (MMOE) by Akaike Information Criterion (AIC), Corrected Akaike Information Criterion (CAIC) and Bayesian information criterion (BIC) and value of log-likelihood.

$$\text{AIC}=-2\ell(\hat{\theta})+2k\ ;\ \text{BIC}=-2\ell(\hat{\theta})+k\log(n)\quad\text{and}\quad \text{CAIC}=AIC+\frac{2k(k+1)}{n-k-1}$$

Where k is the number of parameters in the model and n is total sample under consideration.

Here, the parameters are determined by the *optim* ( ) function of each model from R software and determine the values of Log-likelihood, AIC, BIC and CAIC. The least value of each finding shows the better fit as compare to other models. As a result, the intended model is satisfactory model among several models (Table 4).

Table 4: Comparing the value of Log-likelihood, AIC, BIC and CAIC

Models	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\theta}$	$\hat{\lambda}$	$\ell(\hat{\theta})$	AIC	BIC	CAIC
GIGE	0.5785	2.2031^{p}	-	1.6003	-181.4761	368.9522	373.9428	369.6379
EIR	0.21421	-	1.54827^{q}	-	-184.4011	372.8021	376.1293	273.1354
EGIE	0.67672	24.75149	-	0.05817	-180.1067	366.2134	371.2058	366.8991
EHLE	0.01015	0.48769	-	1.87052	-182.6091	371.2182	376.2088	371.9046
EMOE	0.00728	0.09596	1.16686	-	-178.1595	362.3191	367.3096	367.9953

$q=\sigma, p=\gamma$

Likewise, we compared the estimated cdf of the intended model with the estimated cdf of the GIGE, EIR, EGIE and HLE distributions. Again, we plot the estimated pdf of the intended model with the estimated pdf of these well-known distributions. In both cases, our intended model fits perfectly than all other distributions. As a result, the intended model is an alternative model for real data modeling in various fields, which has tremendous upside (Fig 3).

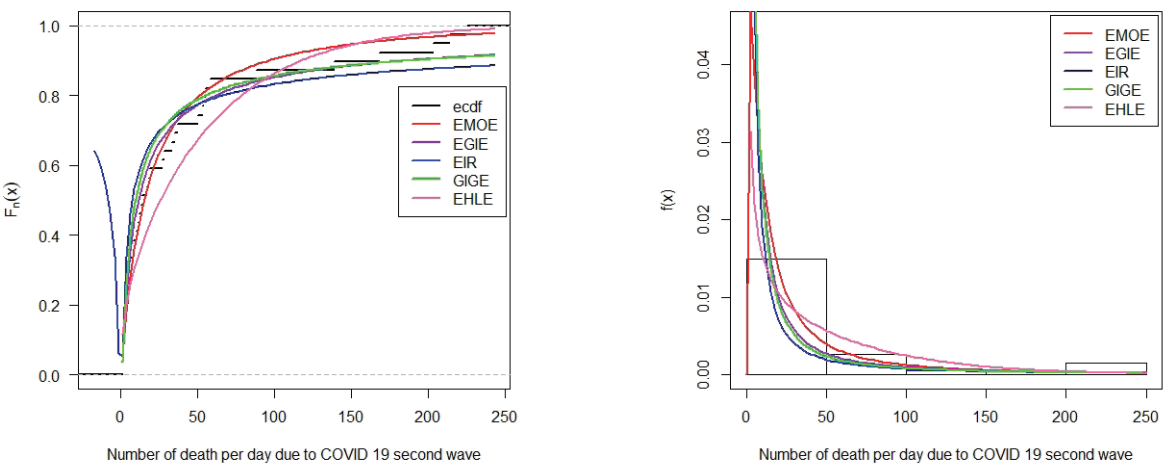

Figure 3: *Estimated fitted cdfs with edf (left panel) and estimated fitted pdf (right panel)*

Conclusion:

In this study, we proposed a new three-parameter probability distribution called the Exponentiated Marshall-Olkin Exponential (EMOE) distribution. Some of the important properties of the distribution, namely quintile; asymptotic behavior, moments, residual life function; R`enyi and q entropy; probability weighted moments and order statistic are investigated of intended model. The value of parameters of proposed model of death rate of COVID 19 in Nepal is estimated by maximum likelihood methods with its confidence interval. We have predicted the mortality rate per day will go as high as up to 150 in Nepal. Therefore, government have to focus on the improvement the health situation of Nepal. From the data analysis (graphical as well as numerical), it is observed that proposed distribution has precisely fitted than others some well-known distribution. Hence, the proposed model is satisfactory model for prediction the mortality rate of COVID 19 in Nepal. Hence it can be applied in skewed and non-normal data modeling and reliability analysis in different areas like as engineering, actuarial, environmental science, medical sciences, biological studies, demography, economics and others.

Appendix –A (Element of Observed Information Matrix)

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

$$\frac{\partial^2 \ell}{\partial \beta^2} = -\frac{1}{\beta^2} - (\theta+1)\sum_{i=1}^{n}\xi_i \tag{2}$$

$$\frac{\partial^2 \ell}{\partial \theta^2} = -\frac{1}{\theta^2} \tag{3}$$

$$\frac{\partial^2 \ell}{\partial \alpha \partial \beta} = (\theta+1)\sum_{i=1}^{n}\left(x_i e^{\alpha x_i}\, \xi_i^2\right) \tag{4}$$

$$\frac{\partial^2 \ell}{\partial \alpha \partial \theta} = \sum_{i=1}^{n}\left(xi\ \omega_i\right) + (\beta-1)\sum_{i=1}^{n}\left(xi\ \xi_i\right) \tag{5}$$

$$\frac{\partial^2 \ell}{\partial \beta \partial \theta} = \sum_{i=1}^{n}\xi_i \tag{6}$$

Appendix B (Probability density function of Compared Model)

$$f_{EHLE}(x) = \frac{2a\lambda\alpha e^{-\alpha\lambda x}[1-e^{-\alpha\lambda x}]^{a-1}}{[1+e^{-\alpha\lambda x}]^{a+1}}; x>0, \alpha>0, \lambda>0. \tag{1}$$

$$f_{GIGE}(x) = \alpha\lambda\gamma x^{-2} e^{-\gamma(\lambda/x)}(1-e^{-\gamma(\lambda/x)})^{\alpha-1}; x>0, \alpha>0, \gamma>0, \lambda>0 \tag{2}$$

$$f_{EIR}(x) = \frac{2\alpha\sigma^2}{x^3} e^{-(\sigma/x)^2}(1-e^{-(\sigma/x)^2})^{\alpha-1}; x>0, \alpha>0, \sigma>0. \tag{3}$$

$$f_{EGIE}(x) = \alpha\beta\lambda x^{-2} e^{-\lambda/x}(1-e^{-\lambda/x})^{\alpha-1}\left[1-(1-e^{-\lambda/x})^{\alpha}\right]^{\beta}; x>0, \alpha>0, \lambda>0, \gamma>0. \tag{4}$$

References

[1] Almarashi, A. M., & Elgarhy, M. **(2018)**. A new muth generated family of distributions with applications. *J. Nonlinear Sci. Appl*, **11**: 1171-1184.

[2] Almarashi, A. M., Khalil, M. G., Elgarhy, M., & ElSehetry, M. M. **(2018)**. Exponentiated half logistic exponential distribution with statistical properties and applications. *Advances and applications in statistics*, **53**(4): 423-440.

[3] Bastola, A., Sah R., Rodriguez-Morales, A.J., Lal B.K., Jha, R. ,& Ojha H.C. **(2020)**. The first 2019 novel coronavirus case in Nepal. *Lancet Infect Diseases*, **20** (3): 279–280.

[4] Braun, W. J., & Murdoch, D. J. **(2016)**. *A first course in statistical programming with R*. Cambridge University Press.

[5] Covid-19: Second wave strikes Nepal; 14 districts declared hotspots **(2021)**. *The Times of India*, retrieved from https://timesofindia.indiatimes.com/.

[6] David, H. A. **(1981)**. *Order Statistics*, Second edition, Wiley, New York.

[7] De Brito, C. R., Rêgo, L. C., De Oliveira, W. R., & Gomes-Silva, F. **(2019)**. Method for generating distributions and classes of probability distributions: the univariate case. *Hacettepe Journal of Mathematics and Statistics*, **48**(3): 897-930.

[8] Dhungana, G. P. **(2020)**. A New Poisson Inverted Exponential Distribution: Model, Properties and Application. *Prithvi Academic Journal*, 136-146.

[9] Eliwa, M. S., & El-Morshedy, M. **(2020)**. Bivariate odd Weibull-G family of distributions: properties, Bayesian and non-Bayesian estimation with bootstrap confidence intervals and application. *Journal of Taibah University for Science*, **14**(1): 331-345.

[10] Elflein, John. **(2021)**. COVID-19 deaths worldwide as of May 14, 2021, by country. https://www.statista.com/statistics/1093256/novel-coronavirus-2019ncov-deaths-worldwide-by-country/.

[11] Elsehetry, M., & El-Sherpieny, E. S. **(2019)**. Kumaraswamy Type I Half Logistic Family of Distributions with Applications. *Gazi University Journal of Science*, **32**(1): 333-349.

[12] Ghitany, M. E., Al-Hussaini, E. K., & Al-Jarallah, R. A. **(2005)**. Marshall–Olkin extended Weibull distribution and its application to censored data. *Journal of Applied Statistics*, 32(10), 1025-1034.

[13] Government of Nepal Ministry of Health and Population. **(2021)**. Health sector response to COVID-19 SitRep#460, SitRep#391

[14] Gupta, R. D., & Kundu, D. **(2001)**. Generalized exponential distribution: different method of estimations. *Journal of Statistical Computation and Simulation*, **69**(4): 315-337.

[15] Hosseini, B., Afshari, M., & Alizadeh, M. **(2018)**. The generalized odd gamma-G family of distributions: properties and applications. *Austrian Journal of Statistics*, **47**(2): 69-89.

[16] Joshi, R. K., & Dhungana, G. P. **(2020)**. Exponentiated Rayleigh Poisson Distribution: Model, Properties and Applications. *American Journal of Theoretical and Applied Statistics*, **9**(6): 272-282.

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

[18] Kumar, V., & Ligges, U. **(2011)**. *reliaR: A package for some probability distributions*. http://cran.r-project.org/web/packages/reliaR/index.html.

[19] Lee, C., Famoye, F., & Alzaatreh, A. Y. **(2013)**. Methods for generating families of univariate continuous distributions in the recent decades. *Wiley Interdisciplinary Reviews: Computational Statistics*, **5**(3): 219-238.

[20] Mahdavi, A., & Kundu, D. **(2017)**. A new method for generating distributions with an application to exponential distribution. *Communications in Statistics-Theory and Methods*, **46**(13): 6543-6557.

[21] Mansoor, M., Tahir, M. H., Cordeiro, G. M., Provost, S. B., & Alzaatreh, A. **(2019)**. The Marshall-Olkin logistic-exponential distribution. *Communications in Statistics-Theory and Methods*, **48**(2): 220-234.

[22] Marshall, A. W., & 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.

[23] Nichols, M. D., & Padgett, W. J. **(2006)**. A bootstrap control chart for Weibull percentiles. *Quality and reliability engineering international*, **22**(2): 141-151.

[24] Oguntunde, P. E., & Adejumo, A. O. **(2015)**. The generalized inverted generalized exponential distribution with an application to a censored data. *Journal of Statistics Applications & Probability*, **4**(2): 223-230.

[25] Oguntunde, P. E., Adejumo, A., & Balogun, O. S. **(2014)**. Statistical properties of the exponentiated generalized inverted exponential distribution. *Applied Mathematics*, **4**(2): 47-55.

[26] Pogány, T. K., Saboor, A., & Provost, S. **(2015)**. The marshall-olkin exponential weibull distribution. *Hacettepe Journal of Mathematics and Statistics*, **44**(6): 1579.

[27] R Core Team R (2020). *A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

[28] Rao, G. S., & Mbwambo, S. **(2019)**. Exponentiated inverse Rayleigh distribution and an application to coating weights of iron sheets data. *Journal of probability and statistics*, 2019.

[29] Ristić, M. M., & Kundu, D. **(2015)**. Marshall-Olkin generalized exponential distribution. *Metron*, **73**(3): 317-333.

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

[31] Tahir, M. H., & Nadarajha, S. **(2015)**. Parameter induction in continuous univariate distributions: Well-established G families. *Anais da Academia Brasileira de Ciências*, **87**(2): 539-568.

[32] Tahir, M. H., Zubair, M., Mansoor, M., Cordeiro, G. M., ALİZADEHK, M., & Hamedani, G. G. **(2016)**. A new Weibull-G family of distributions. *Hacettepe Journal of Mathematics and Statistics*, **45**(2): 629-647.

[33] Ul Haq, M. A., Usman, R. M., Hashmi, S., & Al-Omeri, A. I. **(2019)**. The Marshall-Olkin length-biased exponential distribution and its applications. *Journal of King Saud University-Science*, **31**(2): 246-251.