GET Journal of Biosecurity and One Health

Global Emerging Pathogens Treatment Consortium

GET JOURNAL OF BIOSECURITY & ONE HEALTH

PUBLISHED BY:

Global Emerging Pathogens Treatment Consortium

JOURNAL WEBSITE

www.getjournal.org

Date Received: August 9, 2022
Date Reviewed: March 16, 2023
Date Accepted: May 16, 2023

Gamma Generalized Extended Inverse Exponential Distribution: A Novel Distribution in Modeling COVID-19 Cases in Nigeria

Ogunde AA[1*]; Chukwu AU[2]; Oseghale OI[3] and Nwanyibuife OB[4]

[1,2] Department of Statistics, University of Ibadan, Ibadan, Oyo State, Nigeria.

[3]Department of Mathematics and Statistics, Joseph Babalola University, Arakeji, Osun State, Nigeria.

[4]Department of Statistics, the Federal University of Technology Owerri, Imo State, Nigeria.

ORCID ID: 0000-0001-8708-8612

ABSTRACT

In this study, we developed a novel distribution called Gamma Inverse Exponential (GIE) distribution, which has proved to be a more flexible distribution in modeling COVID-19 case fatality in Nigeria. We studied some statistical properties of the new distribution, which include: moments, incomplete moments, quantile function, Renyi entropy, and mean deviation. A real-life data application to a number of reported cases of COVID-19 infection between March 2019 to 2021 shows that the GIE distribution has a better fit than some competing distributions in fitting the data. Time series analysis of the COVID-19 data is also considered.

Keywords: COVID-19; Characteristics Function; Mean Residual Function; Case Fatality; Gamma Inverse Exponential Distribution.

INTRODUCTION

In probability theory, the extreme value (EV) distribution is a family of continuous distributions developed within the framework of extreme value theory, which includes the popular Gumbel, Frechet and Weibull distributions which are respectively known as type I, II and III extreme value distributions. The EV distributions sometimes exist as a limiting distribution for the minimum or maximum of a sample of independent, identically distributed random variables. Extreme value is the theory of modeling and measuring events which occur with a very small probability. These distributions and their generalized forms can be applied in finance, material sciences, telecommunications, economics and many others. Frechet distribution, because of its heavy tail, has been applied in the modeling of market returns which is related to finance [1]. The cumulative distribution function (CDF) of the standard Frechet distribution is

$$F(x) = e^{\left(-\frac{\lambda}{x}\right)^{\rho}}, \qquad x > 0, \lambda, \rho > 0 \tag{1.1}$$

Where λ is a scale parameter, and ρ is a shape parameter

By letting $\rho = 1$, we obtain another distribution known as Inverse Exponential (IE) distribution which CDF is given by

Corresponding Author's Email: aa.ogunde@ui.edu.ng

 GET Journal of Biosecurity and One Health (2023) 2(1), 38-45.
DOI:10.36108/GJOBOH/3202.20.0120

$$F(x) = e^{-\frac{\lambda}{x}}, \qquad x > 0, \lambda > 0 \qquad (1.2)$$

The associated probability density function (PDF) to (1.2) is given by

$$f(x) = \lambda x^{-2} e^{-\frac{\lambda}{x}}, \qquad x > 0, \lambda > 0 \qquad (1.3)$$

Here, the λ is a scalar parameter.

In many real-life situations, the classical distribution does not give a sufficient fit to lifetime data, most especially when the data exhibit different shapes of the hazard function. Therefore, various generators are proposed to produce a new model with improved modeling potential [2-13].

THE GAMMA INVERSE EXPONENTIAL (GIE) DISTRIBUTION

Zografos and Balakrishnan [12] and Ristic and Balakrishnan [13] proposed a family of univariate distributions generated using gamma random variables. Given any baseline CDF $F(x)$, and $x \in R$, they defined the gamma-G distribution with an extra shape parameter σ > 0 and PDF $g(x)$ and CDF $G(x)$ given by

$$g(x) = \frac{1}{\Gamma(\sigma)} [-log\{1 - F(x)\}]^{\sigma-1} f(x) \qquad (2.1)$$

And

$$G(x) = \frac{\gamma(\sigma, -log[1 - F(x)])}{\Gamma(\sigma)} = \frac{1}{\Gamma(\sigma)} \int_0^{-log\{1-F(x)\}} t^{\sigma-1} e^{-t} dt, \quad (2.2)$$

respectively, where $f(x) = d\,F(x)/dx$, $\Gamma(\sigma) = \int_0^\infty t^{\sigma-1} e^{-t} dt$ and $\gamma(\sigma, y) = \int_0^y t^{\sigma-1} e^{-t} dt$ are the gamma and the incomplete gamma functions. The shape parameter σ controls skewness and kurtosis through the tail weight

Putting (1.1) and (1.2) in (2.1), we obtain the PDF of the GIE distribution given by

$$g(x) = \frac{\lambda\sigma}{\Gamma(\sigma)} x^{-2} e^{-\frac{\lambda}{x}} \left[-log\left\{1 - e^{-\frac{\lambda}{x}}\right\}\right]^{\sigma-1} \qquad (2.3)$$

And the corresponding CDF to (2.3) is given by

$$G(x) = \frac{\gamma\left(\sigma, -log\left[1 - e^{-\frac{\lambda}{x}})\right]\right)}{\Gamma(\sigma)} \qquad (2.4)$$

An expression for the survival and the hazard function is, respectively as

$$S(x) = 1 - \frac{\gamma\left(\sigma, -log\left[1 - e^{-\frac{\lambda}{x}})\right]\right)}{\Gamma(\sigma)}, \qquad (2.5)$$

And

$$h(x) = \frac{\lambda\sigma x^{-2} e^{-\frac{\lambda}{x}} \left[-log\left\{1 - e^{-\frac{\lambda}{x}}\right\}\right]}{\Gamma(\sigma) - \gamma\left(\sigma, -log\left[1 - e^{-\frac{\lambda}{x}})\right]\right)}. \qquad (2.6)$$

The graphs of the density and the hazard function are given in Figures 1 and 2 for various values of the parameters.

QUANTILE FUNCTION OF GIE DISTRIBUTION

From Nadarajah et al. [9], we can generate GIE random variables from the quantile function given by

$$F^{-1}(u) = \lambda \left[-log\left(1 - e^{\left(-Q^{-1}(\sigma, 1-u)\right)}\right)\right] \qquad (2.7)$$

Where $Q^{-1}(\sigma, u)$ denote the inverse function of $Q(\sigma, x) = 1 - {\gamma(\sigma, x)}/{\Gamma(\sigma, x)}$; setting $u = 0.5$

in (2.7), we obtain the median (M) of GIE distribution as

$$M = \lambda \left[-log\left(1 - e^{\left(-Q^{-1}(\sigma, 0.5)\right)}\right)\right] \qquad (2.8)$$

PROPERTIES OF GIE DISTRIBUTION

Here, we obtain an expression for the moment, incomplete moment, mean deviation and the entropy of the GIE model.

MOMENTS AND INCOMPLETE MOMENTS OF GIE DISTRIBUTION

We now obtain the r^{th} moment about zero of X, say $E(X^r)$. The r^{th} moment of $X \sim GIE\ (\lambda, \sigma)$ is found by

$$E(X^r) = \int_{-\infty}^{\infty} x^r f(x)dx \qquad (3.0)$$

Substitute from Equation (2.3) into Equation (3.0), and we will get the r^{th} moment as follows.

$$E(X^r) = \lambda^r \Gamma(1-r) \sum_{i,l=0}^{\infty} \frac{(-1)^{l+1}(i+1)V_i}{(1+l)^{(1-r)}} \binom{i}{l} \qquad (3.1)$$

Where

$$= \frac{(-1)^i}{(i+1)!} \sum_{k=0}^{\infty} \frac{(\sigma+k)}{\Gamma(\sigma+k-i)} \Gamma(\sigma+k)$$

Next, we derive a simple formula for the r^{th} incomplete moment of X, say $m_r(t) = E(X_r|X < t)$. From equation (2.3), we obtain

$$m_r(t) = \lambda^r \Gamma(1-r) \sum_{i,l=0}^{\infty} \frac{(-1)^{l+1}(i+1)V_i}{(1+l)^{(1-r)}} \binom{i}{l} \left\{ [(1+l)\lambda]^{-(1-r)} - \gamma\left[1-r,(1+l)\frac{\lambda}{t}\right] \right\} \qquad (3.2)$$

MEAN DEVIATIONS

The mean deviations about the mean $\left(\delta_1(X) = E(|X-\mu_1'|)\right)$ and about the median $\left(\delta_2(X) = E(|X-M|)\right)$ of X can be represented as:

$$\delta_1(X) = 2\mu_1' F(\mu_1') - 2m_1(\mu_1') \text{ and } \delta_2(X) = \mu_1' - 2m_1(M) \qquad (3.3)$$

Respectively, where $\mu_1' = E(X)$ can be obtained from (3.1), $M = median\ (X)$ is the median given in (2.8).$F(\mu_1')$ can be easily obtained from the CDF (2.2), and $m_1(t) = \int_0^t xf(x)dx$ is the first incomplete moment obtained from (3.2) by setting $r = 1$.

Applications of these equations can be used to obtain an explicit expression for the Bonferroni and Lorenz curves defined for a given probability π by $B(\pi) = \frac{m_1(h)}{\pi\mu_1'}$ and $L(\pi) = \frac{m_1(h)}{\mu_1'}$, respectively, where

$h = F^{-1}(\pi)$ is the GIE function at π defined from (2.7).

RENYI ENTROPY

The entropy of a random variable X with density function f (x) is a measure of the variation of the uncertainty. For any real parameter $\rho > 0$ and $\rho = 1$, the Renyi entropy is given by

$$I_R(\rho) = \frac{1}{1-\rho} log \int_0^{\infty} f^{\rho}(x)dx \qquad (3.4)$$

Putting equation (2.3) in (2.4), we have an expression for the moments of GIE distribution as

$$\frac{1}{1-\rho} log \left\{ \begin{array}{c} \frac{(\sigma-1)\rho\lambda^{\rho}}{[\Gamma(\sigma)]^{\rho}} \sum_{k,l,m}^{\infty} (-1)^{k+l+m+1} \frac{\binom{k-(\sigma-1)\rho}{k}\binom{k}{m} p_{m,k}}{(\sigma-1)\rho - m} \frac{\Gamma[(\sigma-1)\rho+1]\Gamma(w)}{l!\,[(k+l+\sigma w)]^{w}} \\ \times \Gamma[(\sigma-1)\rho+1-l] \end{array} \right\} \qquad (3.5)$$

Where,

$$w = \rho[(\sigma-1)\rho+1-l] \text{ and } p_{m,k} = k^{-1} \sum_{j=0}^{k} \frac{(-1)^j [j(m+1)-k]}{(j+1)} p_{m,k-j}$$

For $k = 1,2,\ldots, and\ p_{m,0} = 1$

ESTIMATION

Here, we determine the maximum likelihood estimates (MLEs) of the parameters of the GIE distribution from complete samples only. Let $x_1, \ldots, x_n$ be a random sample of size n from the GIE (λ, σ) distribution. The log-likelihood function can be expressed as

$$l(\lambda, \sigma) = -nlog\Gamma(\sigma) + nlog(\lambda\sigma) - \lambda \sum_{i=1}^{n} x_i^{-1} - 2\sum_{i=1}^{n} log(x_i)$$

$$+(\sigma-1)\sum_{i=1}^{n} \left[log\left(-log\left[1-e^{-\frac{\lambda}{x_i}}\right]\right)\right]$$

The element of the score vector for GIE model is given as

GET Journal of Biosecurity and One Health (2023) 2(1), 38-45.
DOI:10.36108/GJOBOH/3202.20.0120

$$\frac{\partial l}{\partial \lambda} = \frac{\left[n-\lambda \sum_{i=1}^{n}(x_i^{-1})\right]}{\lambda} + (\sigma-1)\sum_{i=1}^{n}\left[\frac{\frac{\sigma}{x_i}e^{\frac{\sigma}{x_i}}}{\lambda\left(1-e^{\frac{\sigma}{x_i}}\right)log\left(1-e^{\frac{\sigma}{x_i}}\right)}\right] \qquad (4.1)$$

$$\frac{\partial l}{\partial \sigma} = n\psi(\sigma) + \sum_{i=1}^{n} log\left(-log\left[1-e^{\frac{\sigma}{x_i}}\right]\right) \qquad (4.2)$$

APPLICATION

In this section, we present the usefulness of the GIE distribution by applying it to the case fatality COVID-19 data set, Nigeria's experience between the months of March 2020 to December 2021 and compare the performance of the model with its sub-model. Now, we apply the Akaike information criterion (AIC), consistent Akaike information criterion (CAIC), Hannan-Quinn criterion (HQIC), and Kolmogorov-Smirnov(K) goodness of fit test to verify which distribution fits better to these data. In general, the smaller the values of the statistics AIC, CAIC, HQIC and KS, the better the fit to the data. Table I contains the exploratory data analysis for the COVID-19 data. Table 2 shows the measure of goodness of fit test. Table 3 presents the Auto-Regressive Integrated Moving Average (ARIMA) model for the COVID-19 data. The graph of the daily number of reported cases is drawn in Figure 4.

CONCLUSION

The Gamma Inverse Exponential model is considered a better model when compared to its sub-model in modeling the COVID-19 data because it returns the minimum values of information criteria. ARIMA (1, 0, 1) is the best model under the ARIMA scheme, and the point forecast obtained indicates that if a drastic step is not taken by January 2023, the monthly reported cases of COVID-19 patients will be 23008.

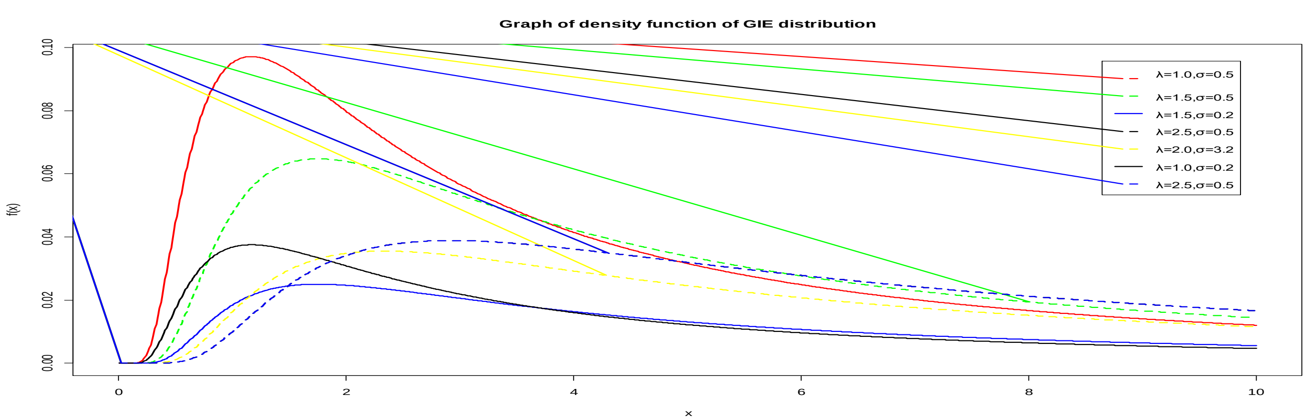

Figure 1: Graph of the density function of GIE distribution

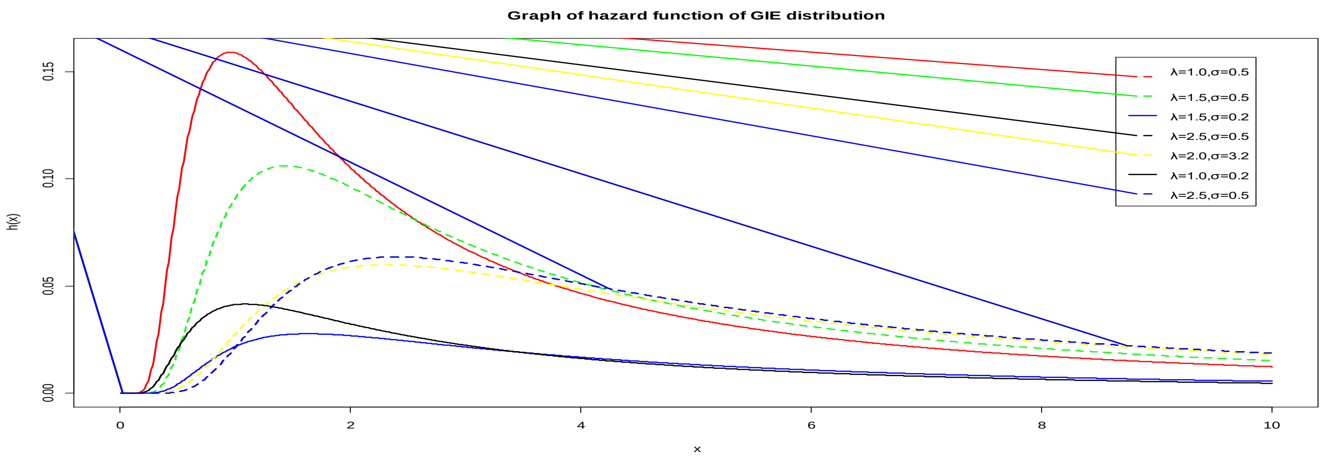

Figure 2: Graph of the hazard function of GIE distribution

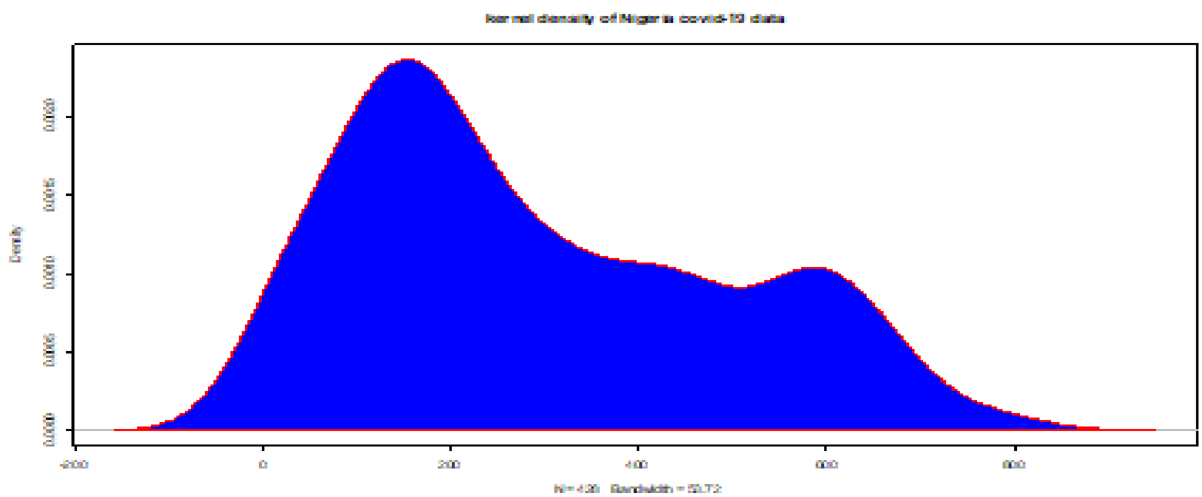

Figure 3: Kernel density curve for COVID-19 data

 GET Journal of Biosecurity and One Health (2023) 2(1), 38-45.
DOI:10.36108/GJOBOH/3202.20.0120

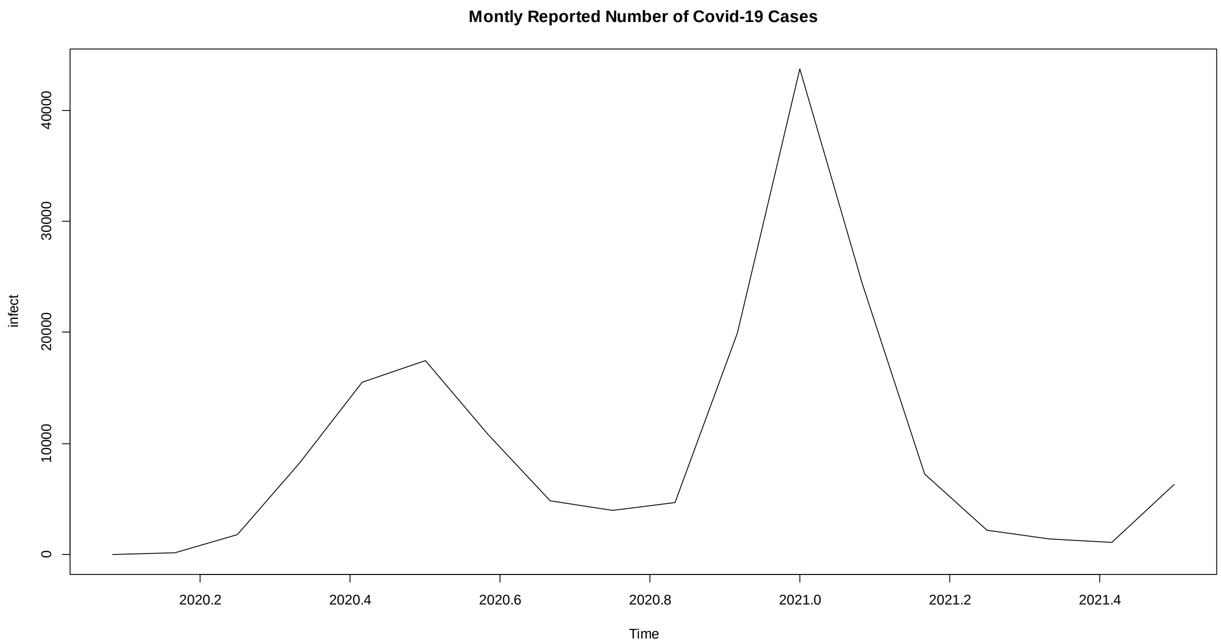

Figure 4: Graph of daily reported cases of COVID-19 in Nigeria.

Table 1: Exploratory data Analysis of COVID-19 data

Min.	q_1	q_2	Med.	Mean	Max.	Var.	Skew.	Kurt.
1.0	86.0	499.5	203.0	345.4	2314.0	139613	2.0	7.3

Table 2: Measures of goodness of fit of GIE model

$Model$	λ	σ	$-l$	AIC	$CAIC$	$HQIC$	KS
GIE	20.05 (0.93)	1.46 (0.04)	4403.42	8810.85	8810.87	8814.29	0.2347
IE	5.72 (2.23)	− (−)	4585.35	9172.72	9172.71	9174.42	0.3090

GET Journal of Biosecurity and One Health (2023) 2(1), 38-45.
DOI:10.36108/GJOBOH/3202.20.0120

Table 3: Models with information criterion

ARIMA models	**Value of mean**	**AIC**
$ARIMA(2,0,2)$	$With\ nozero\ mean$	376.3729
$ARIMA(0,0,0$	$With\ non-zero\ mean$	506.9031
$ARIMA(1,0,0)$	$With\ non-zero\ mean$	391.5391
$ARIMA(0,0,1)$	$With\ non-zero\ mean$	398.7286
$ARIMA(0,0,0)$	$With\ zero\ mean$	504.9151
$ARIMA(1,0,2)$	$With\ non-zero\ mean$	374.1229
$ARIMA(0,0,2)$	$with\ non-zero\ mean$	394.9124
$ARIMA(1,0,1)$	$with\ non-zero\ mean$	384.8265
$ARIMA(1,0,3)$	$with\ non-zero\ mean$	374.6944
$ARIMA(0,0,2)$	$with\ non-zero\ mean$	396.022
$ARIMA(2,0,1)$	$with\ non-zero\ mean$	386.9209
$ARIMA(2,0,3)$	$with\ non-zero\ mean$	377.9495
$ARIMA(1,0,2)$	$with\ zero\ mean$	372.4052
$ARIMA(0,0,2)$	$with\ zero\ mean$	393.029
$ARIMA(1,0,1)$	$with\ zero\ mean$	383.1155
$ARIMA(2,0,2)$	$with\ zero\ mean$	374.5912
$ARIMA(1,0,3)$	$with\ zero\ mean$	372.9165
$ARIMA(0,0,1)$	$with\ zero\ mean$	396.9129
$ARIMA(0,0,3)$	$with\ zero\ mean$	394.151
$ARIMA(2,0,1)$	$with\ zero\ mean$	385.072
$ARIMA(2,0,3)$	$with\ zero\ mean$	376.1708

ARIMA (1,0,2) with zero mean : 392.2688

GET Journal of Biosecurity and One Health (2023) 2(1), 38-45.
DOI:10.36108/GJOBOH/3202.20.0120

REFERENCES

[1] Alves IF, Neves C. Extreme Value Distributions. 2010. http://docentes.deio.fc.ul.pt/fragaalves/fraga_alves_lexicon.pdf. (Accessed 5th May 2023).

[2] Alexander C, Cordeiro GM, Ortega EMM, Sarabia JM. Generalized Beta-Generated Distributions. Comput. Stat. Data Anal. 2012; 56: 1880–1897.

[3] Alzaghal A, Famoye F, Lee C. Exponentiated T-X Family of Distributions with Some Applications. Int J probab Stat. 2013; 2: 1-31.

[4] Amini M, Mir Mostafaee SMTK, Ahmadi J. Log-Gamma-Generated Families of Distributions. Statistics. 2012; 1: 1-20.

[5] Bourguignon M, Silva RB, Cordeiro GM. The Weibull-G Family of Probability Distributions. Data Sci J. 2014; 12: 53–68.

[6] Cordeiro GM, Alizadeh M, Diniz Marinho PR. The Type I Half-Logistic Family of Distributions. J. Stat. Comput. Simul. 2016; 86: 707–728.

[7] Ghosh I, Alizadeh M, Cordeiro GM, Pinho LG. The Gompertz-G Family of Distributions. J. Stat. Theory Pract 2016; 11(1): 179–207.

[8] Hosseini B, Afshari M, Alizadeh M. The Generalized Odd Gamma-G Family of Distributions: Properties and Applications. Austrian J. Stat. 2018; 47: 69–89.

[9] Nadarajah S, Korkmaz MC, Cordeiro GM, Yousof HM, Pescim RR, Afify AZ. The Weibull Marshall–Olkin Family: Regression Model and Application to Censored Data. Commun Stat. 2019; 48: 4171–4194.

[10] Nadarajah S, Kotz S. The Exponentiated Fr'Echet Distribution. InterSat. 2003. Available from: http://interstat.statjournals.net/YEAR/2003/abstracts/0312001.php (Accessed 5th May 2023).

[11] Reis LDR, Cordeiro GM, Lima MCS. The Gamma-Chen Distribution: A New Family of Distributions with Applications. Span J Stat.. 2020; 2: 23–40.

[12] Zografos K, Balakrishnan N. On Families of Beta- and Generalized Gamma-Generated Distributions and Associated Inference. Stat Methodol. 2009; 6: 344–362.

[13] Ristic MM, Balakrishnan N. The Gamma Exponentiated Exponential Distribution. J Stat Comput Simul. 2003; 82(8); 1191-1206.