Modeling COVID-19 Pandemic Data with Beta-Double Exponential Model

N. I. Badmus[1*], Faweya Olanrewaju[2] and A. T. Adeniran[3]

[1]Department of Mathematics, University of Lagos, Akoka, Nigeria.
[2]Department of Statistics, Ekiti State University, Ado-Ekiti, Ekiti State, Nigeria.
[3]Department of Statistics, University of Ibadan, Ibadan, Nigeria.

Authors' contributions

This work was carried out in collaboration among all authors. Author NIB designed the study, performed the statistical analysis, wrote the protocol, and wrote the first draft of the manuscript. Authors NIB, FO and ATA managed the analyses of the study. Author NIB managed the literature searches. All authors read and approved the final manuscript.

Original Research Article

ABSTRACT

Objective: This paper examines and upgrades a two-parameter double exponential distribution to a four-parameter beta double exponential model by compounding the baseline distribution and beta link function to fits and analyse deaths-cases data set of the recent outbreak of the global pandemic coronavirus disease (COVID-19) for both Africa and Non-Africa countries. The new proposed model, although complex in its mathematical structure, yet flexible to implement and its robustness to accommodate non-normal data is an extra advantage to statistical theory and other fields.
Methodology: The statistical properties: the density function, cumulative distribution function, survival function, hazard function, moments, moments generating function, skewness and kurtosis of the developed model were presented. Maximum likelihood method is used for parameters

_____

*Corresponding author: E-mail: nibadmus@unilag.edu.ng;

estimation procedure. The new model is validated and compared with some frontier similar extant parametric family of beta distributions using graphs, Kolmogorov Smirnov (KS) Statistic, Log-likelihood and model criteria statistics like Akaike Information Criteria (AIC), Bayesian Information Criteria (BIC) and Consistent Akaike Information Criteria (CAIC) as tools for comparison.
Results: The graphs, KS, LogL and model criteria statistics values showed that the proposed model fits the COVID-19 pandemic data better than other competing models since the model has lower values as stated: The values from non-African countries KS = 0.1208, LogL = 278.4168, AIC = 560.8336, BIC = 576.1147 and CAIC = 577.1147. Also, from African countries are: KS = 0.0759, LogL = 144.0245, AIC = 292.0490, BIC = 303.9302 and CAIC = 304.9302.
Conclusion: The proposed model showed its applicability and flexibility over other models considered in this work. Therefore, this implies that the new model can be used for modeling other infectious disease data and real data in many fields.

1. INTRODUCTION

Coronavirus disease acronym as "COVID-19" is a recently discovered, contagious, pandemic, airborne viral disease [1]. Like Ebola, Polio, Lassa-fever; COVID-19 is a deadly disease with no vaccine as at present. It belongs to the Orthocoronavirinae sub-family, distinct from Middle East respiratory syndrome-coronavirus with severe acute respiratory syndrome coronavirus (SARSCoV), [2]. The first case of an unexplained new pandemic origin was detected in December 12, 2019 and was later diagnosed by the Chinese Center for Disease Control and Prevention (CDC) as a nonSARS nCoV. The Coronaviridae family consists of a group of large, single, and plus stranded RNA viruses isolated from multiple species, and it is known to cause the common cold and diarrheal diseases in humans.

World Health Organization declared COVID-19 outbreak as the sixth public health emergency of international concern, following H1N1 (2009), polio (2014), Ebola (2014) [2] and [3]. By February 11, 2020, the WHO announced the name for the epidemic disease as coronavirus disease 2019 (COVID-19) and by February 24, 2020, COVID-19 has affected more than 79,331 patients in 29 countries across the globe and has become a major global health concern. Based on this, there is need to assess the incidence and prevalence of this new virus both in Africa and the Western-world using sampling instruments.

The aim of this study is to develop a flexible, robust parametric model that will perform better than any of its family of distributions. This is achieved by introducing more shape parameter(s) to the baseline distribution to accommodate and captures the excessive skewness and kurtosis of any clinical/survival data like the COVID-19 data set [4].

Numerous researchers have worked on different flexible and robust models in literature. These include but not limited to [5] the beta exponentiated Pareto distribution with application to bladder cancer susceptibility, [6] the generalized Weibull exponential distribution with its properties and applications, [7] statistical properties of exponential weighted Weibull distribution, [8] beta generated Kumaraswamy-G and other new families of distributions.

The COVID-19 data set used for the analysis covered 53 African countries and 124 countries from the rest of the world. It elicits information on four different attributes, namely: added, confirmed, recovered and deaths. In the analysis, we used only data from deaths-cases since the model is univariate, where y in equation (5) below stands for the number of deaths-cases recorded in each affected country as at April 16, 2020; when the information was gathered from [9] latest updates news. Some of the numbers of deaths-cases recorded on COVID-19 outbreak data against some non-African countries are: USA (30985), Spain (18812), Italy (21645), Germany (3804), France (17188), UK (12894), China (3352), Iran (4777), Turkey (1518), Belgium (4440), Brazil (1760). Although, there are countries with zero (0) number of deaths-cases such as Cambodia, Mongolia, Autigua and Barbuda, Meldires, Laos, Dominica and so on. Also, from African countries like South-Africa (34), Egypt (183), Algeria (336), Morocco (127), Cameroon (17), Tunisia (35), Ghana (8), Nigeria (12), Guinea (1), Niger (14) etc., and some of those that recorded zero number of deaths-cases from Africa are: Rwanda, Madagascar, Uganda, Equatoria Guinea, Guinea Bissau to mention but few.

The organization of this paper is as follows: in section 2, the model pdf is defined with comprehensive treatment of its statistical properties such as; moment, generating function, skewness and kurtosis, estimation of model parameters. Section 3 presents analysis of the survey data sets, discussion of results, and the concluding remarks. Tables and charts showing the results are also presented in this section. Finally, the article ends with an appendix that contains the data.

2. METHODOLOGY

2.1 The Beta-Double Exponential Model (BDEM)

If Y has a double exponential distribution $Y \sim DE(\mu, \sigma)$, then the probability density function (pdf) is

$$f(y) = \frac{1}{2\sigma} exp\left(-\left(\frac{y-\mu}{\sigma}\right)\right) \tag{1}$$

where Y is real and $\sigma > 0$. Therefore, the cumulative distribution function (cdf) of Y also is

$$F(y) = \begin{cases} \frac{1}{2}e^{-\left(\frac{y-\mu}{\sigma}\right)}, & y \le \mu \\ 1-\frac{1}{2}e^{-\left(\frac{y-\mu}{\sigma}\right)}, & y \ge \mu \end{cases} \tag{2}$$

See [10] and [11]

Then, by letting Y be a random variable with the pdf based on parametric form from the logit of beta distribution as defined by [12] is

$$f(y) = \frac{k[F(y)]^{\alpha k-1}\left[1-F(y)^k\right]^{\beta-1}f(y)}{B(\alpha,\beta)} \tag{3}$$

Now, by substituting (1) and (2) above into (3) we obtain the pdf of the new model (beta-double exponential model) as:

$$f(y) = \frac{k\left[\frac{1}{2}exp\left(-\left(\frac{\mu-y}{\sigma}\right)\right)\right]^{\alpha k-1}\left[1-\frac{1}{2}exp\left(-\left(\frac{y-\mu}{\sigma}\right)^k\right)\right]^{\beta-1}\frac{1}{2\sigma}exp\left(-\left(\frac{y-\mu}{\sigma}\right)\right)}{B(\alpha,\beta)}, \text{ for } y \le \mu \tag{4}$$

$$f(y) = \frac{k\left[1-\frac{1}{2}exp\left(-\left(\frac{y-\mu}{\sigma}\right)\right)\right]^{\alpha k-1}\left[\frac{1}{2}exp\left(-\left(\frac{y-\mu}{\sigma}\right)^k\right)\right]^{\beta-1}\frac{1}{2\sigma}exp\left(-\left(\frac{y-\mu}{\sigma}\right)\right)}{B(\alpha,\beta)}, \text{ for } y \ge \mu \tag{5}$$

where, $y \ge 0$, (y = deaths cases data), $\alpha, \beta, \mu \text{ and } \sigma > 0$, $(B(\alpha,\beta))^{-k}$ is the incomplete beta function, α and β are two shape parameters added to the baseline distribution: the role of the two additional shape parameters vary tail weights and provide flexibility in the shape of the distribution [13] and [10], μ and σ are location and dispersion parameters, while k is a constant which is one (1). Then, we can say that $Y \sim BDEM(\alpha, \beta, \mu, \sigma)$.

Therefore, since $y \ge 0$ (y is the observations of the deaths cases in the data set) equation (5) is the pdf of the BDEM and it shall be used throughout in this work. The pdf plot of BDEM in (4) and (5) and the pdf of the models in (4) and (5) are given below in Figs. 1 and 2 respectively:

The corresponding distribution function is given by

$$F(y) = \int_0^y \frac{\left[1-\frac{1}{2}exp\left(-\left(\frac{y-\mu}{\sigma}\right)\right)\right]^{\alpha-1}\left[\frac{1}{2}exp\left(-\left(\frac{y-\mu}{\sigma}\right)\right)\right]^{\beta-1}\frac{1}{2}\sigma^{-1}exp\left(-\left(\frac{y-\mu}{\sigma}\right)\right)}{B(\alpha,\beta)}dy$$

Setting, $u = 1 - \frac{1}{2}exp\left(-\left(\frac{y-\mu}{\sigma}\right)\right)$, then we have

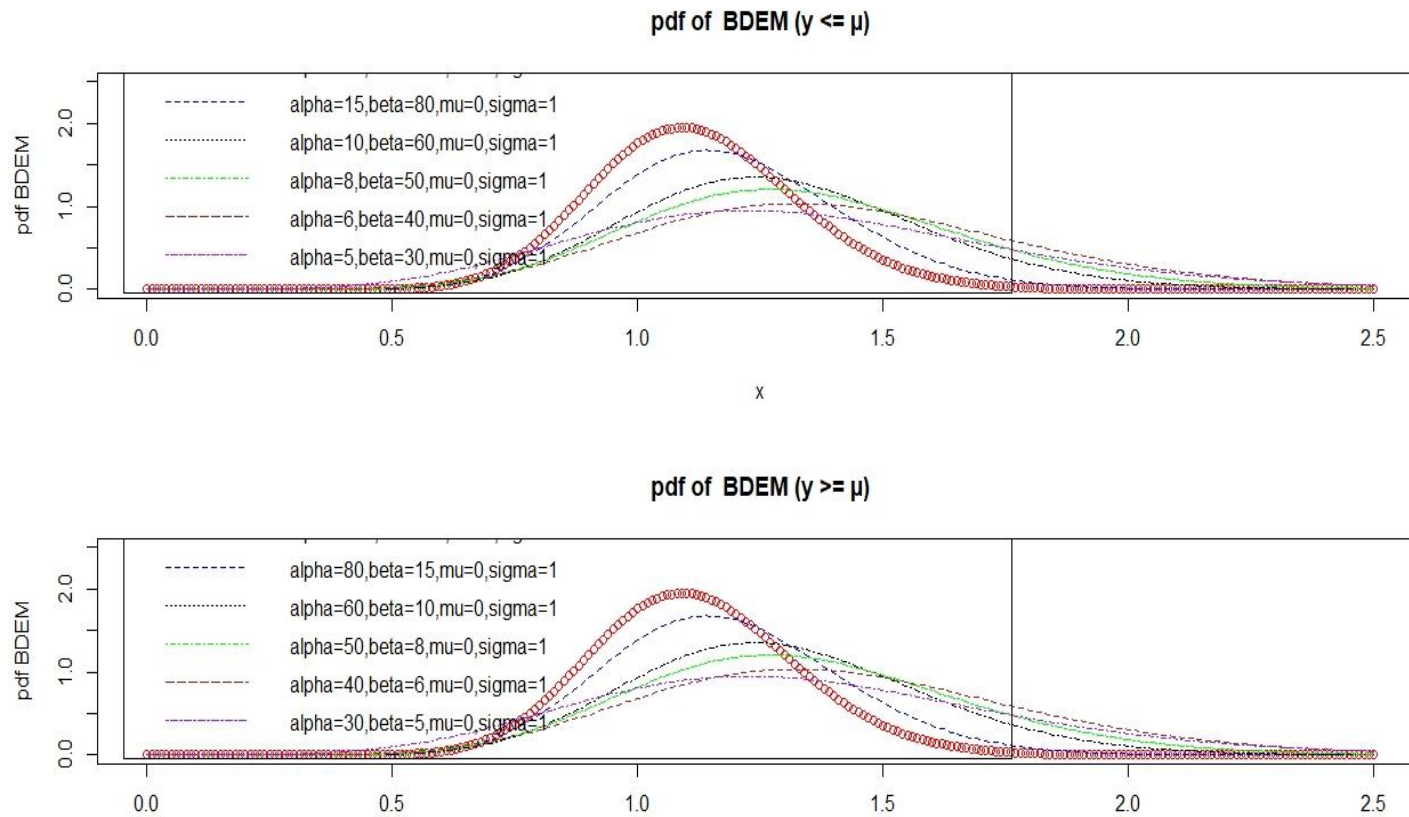$$\frac{1}{B(\alpha,\beta)}\int_0^y [U]^{\alpha-1}[U]^{\beta-1}du$$

pdf of BDEM (y <= μ)

pdf of BDEM (y >= μ)

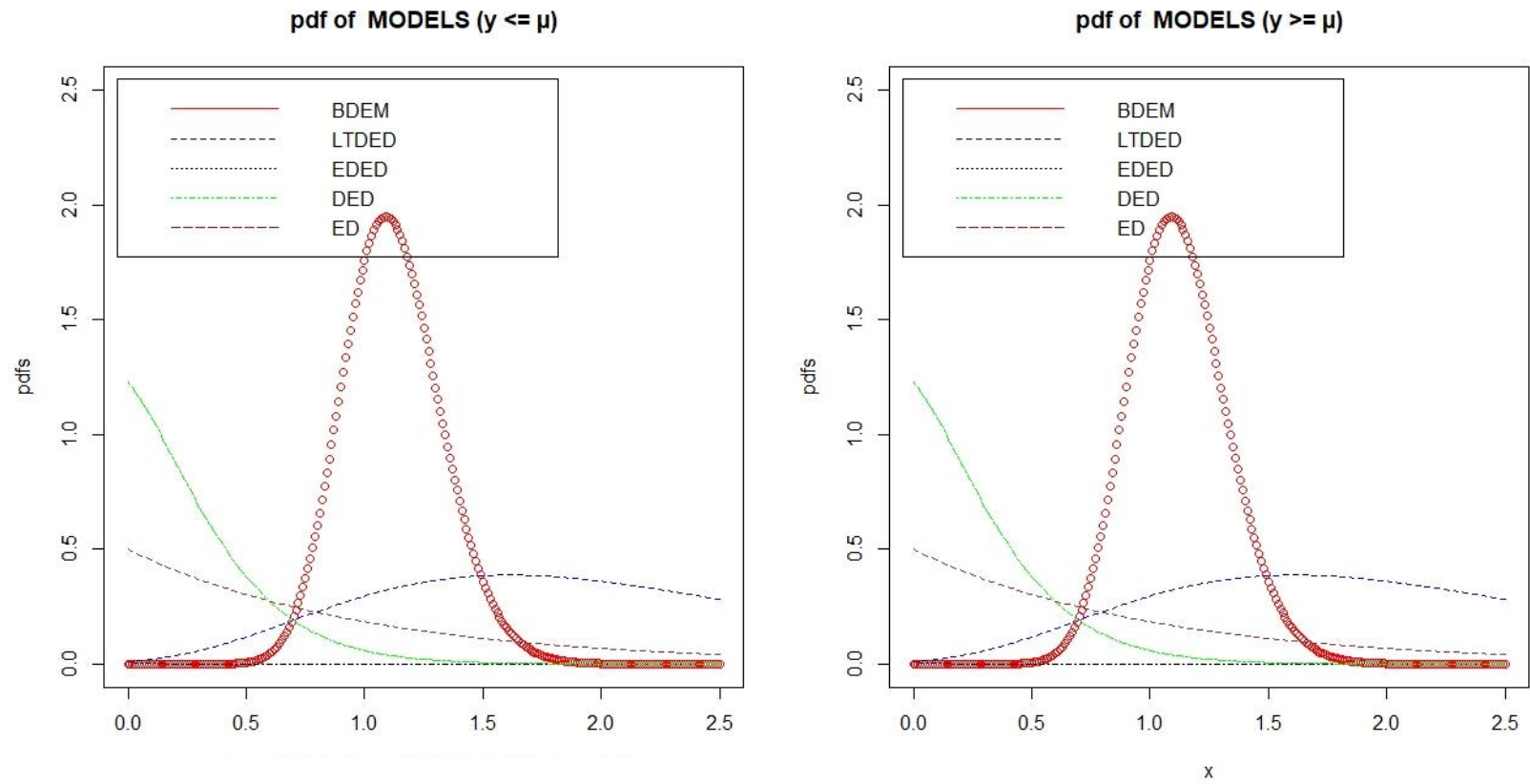**Fig. 1. Plot of the pdf of BDEM in (4) and (5)**

Fig. 2. Plot of the pdf of MODELS in (4) and (5)

where, $\int_0^y [U\ ]^{\alpha-1}[U\ ]^{\beta-1}du = B(u;\ \alpha,\beta) = I_{F(y)}(\alpha,\beta)$

Hence,

$$F(y) = \frac{B(u;\ \alpha,\beta)}{B(\alpha,\beta)} \tag{6}$$

Few sub models arise from the propose model when some parameters equate to 1, such as:

- $\alpha = \beta = 1$; the model becomes double exponential distribution
- $\alpha = 1$; its yields Lehmann Type II double exponential distribution
- $\beta = 1$; it leads to exponentiated double exponential distribution and
- $\alpha = \beta = 1$ and $2\sigma = \sigma$; turns to two-parameter exponential distribution

2.2 The Survival and Hazard Functions BDEM Respectively Obtained as

$$Sur(y) = 1 - F(y) = \frac{B(\alpha,\beta) - B(u;\ \alpha,\beta)}{B(\alpha,\beta)} \tag{7}$$

and hazard rate function is

$$HF(y) = \frac{f(y)}{Sur(y)} = \frac{[U\ ]^{\alpha-1}[U\ ]^{\beta-1}U'}{B(\alpha,\beta) - B(u;\ \alpha,\beta)} \tag{8}$$

where, U' is the pdf of the baseline distribution in (1) above.

3. SOME PROPERTIES OF BDEM

The standardized BDEM random variable defined by $R = \frac{Y-\mu}{\sigma}$ were examined. The density function of R yields

$$\varepsilon(R;\ \alpha,\beta) = \frac{1}{B(\alpha,\beta)}\left[1 - \frac{1}{2}exp\left(-\left(\frac{y-\mu}{\sigma}\right)\right)\right]^{\alpha-1}\left[\frac{1}{2}exp\left(-\left(\frac{y-\mu}{\sigma}\right)\right)\right]^{\beta-1}\frac{1}{2\sigma}exp\left(-\left(\frac{y-\mu}{\sigma}\right)\right)$$

$$= \frac{1}{2B(\alpha,\beta)\sigma}\left[exp(-(r))\ \right]^{\alpha-1}\left[exp(-(r))\ \right]^{\beta-1}exp(-(r)) \tag{9}$$

Also, the associating distribution function is given by

$$F(r) = I_{\left[\frac{1}{2}exp(-(r))\right]}(\alpha,\beta) \tag{10}$$

Meanwhile, when $\alpha = \beta = 1$ corresponds double exponential distribution.

3.1 Moments and Generating Function

Here, we obtained the s-th ordinary moment of BDEM as:

$$\mu'(s) = E(R^s) = \frac{1}{2B(\alpha,\beta)\sigma}\int_{-\infty}^{\infty}R^s\left[exp(-(r))\ \right]^{\alpha-1}\left[exp(-(r))\ \right]^{\beta-1}exp(-(r))dR \tag{11}$$

The expansion of binomial term by [14] and [15] is used as $q = e^R$. We get

$$\mu'(s) = \sum_{i=0}^{\infty}(-1)^i\binom{\beta-1}{i}\int_0^{\infty}log(q)^r\frac{1}{2}e^{-q}\left[\frac{1}{2}e^{-q}\right]^{\alpha(i+1)-1}qdq \tag{12}$$

$$\mu_r^{'} = \frac{1}{B(\alpha,\beta)}\sum_{i=0}^{\infty}(-1)^i\binom{\beta-1}{i}|_{(r,(i+1))} \tag{13}$$

Equation (13) becomes the moment of the BDEM and the tails are controlled by the parameters α and β. This implies that the generating function $M(t) = E(e^{tr})$ according to (9)

$$M(t) = \frac{1}{B(\alpha,\beta)}\sum_{i=0}^{\infty}(-1)^i\binom{\beta-1}{i}\int_0^{\infty}(q)^r\frac{1}{2}e^{-r}\left[\frac{1}{2}e^{-r}\right]^{\alpha(i+1)-1}qdq$$

hence,

$$M(t) = \frac{1}{B(\alpha,\beta)}\sum_{i=0}^{\infty}(-1)^i\binom{\beta-1}{i}[\alpha(i+1)-1]^{-(t+1)} \tag{14}$$

One can derive the moment of (14) from (9) using differentiation method.

Furthermore, the 1st - 4th moments, the skewness and kurtosis of the BDEM using s-th ordinary moment of the BDEM is expressed as:

$$\mu_s^{'} = \int_0^{\infty}logQ\left[\frac{1}{2\sigma B(\alpha,\beta)}[R(q)]^{\alpha-1}[R(q)]^{\beta-1}dr(q)\right]$$

$$\mu_s^{'} = E(q) = exp\left(\frac{\mu(1)}{1-\sigma^2(1^2)}\right)\frac{1}{2\sigma B(\alpha,\beta)}\sum_{i=0}^{\infty}(-1)^i\binom{\beta-1}{i} \tag{15}$$

The mean through the forth moments of the BDEM are given below as follows:

$$\mu = \mu_1^{'}, \ \mu_2 = \mu_2^{'} - \mu^2, \ \mu_3 = \mu_3^{'} - 3\mu\mu_2^{'} + 2\mu^3 \ \text{ and } \mu_4 = \mu_4^{'} - 4\mu\mu_3^{'} + 6\mu^2\mu_2^{'} - 3\mu^4.$$

Both skewness and kurtosis are obtained below:

$$SK = \frac{\mu_3}{\mu_2^{3/5}} = \frac{\mu_3^{'} - 3\mu\mu_2^{'} + 2\mu^3}{\left[\mu_2^{'} - \mu^2\right]^{3/2}} \tag{16}$$

and

$$KUR = \frac{\mu_4}{\mu_2^2} - 3 = \frac{\mu_4^{'} - 4\mu\mu_3^{'} + 6\mu^2\mu_2^{'} - 3\mu^4}{\left[\mu_2^{'} - \mu^2\right]^2} \tag{17}$$

4. PARAMETER ESTIMATION

Here, the estimates of the parameters of the model is obtained and presented through method of Maximum Likelihood Estimation. Let $Y_1, Y_2, \ldots, Y_n$ be a random sample of size n from BDEM distribution with unknown parameter vector $= (\alpha,\beta,\mu,\sigma)$. According to [16] they described log-likelihood function for $\theta = (\alpha,\beta,k,\pi)$, where $\pi = (\mu,\sigma)$, then the likelihood function of BDEM is defined as

$$l(\theta) = nlogk - nlog[B(\alpha,\beta)] + \sum_{i=1}^{n}log[f(Y_i,\pi)] + (\alpha-1)\sum_{i=1}^{n}log[F(Y_i,\pi)]$$

$$+(\beta-1)\sum_{i=1}^{n}log[1-F(Y_i,\pi)] \tag{18}$$

Letting $k = 1$, reduces the generalized beta distribution to the beta generated distribution as we obtain $\theta = (\alpha,\beta,1,\pi)$ as follows:

$$l(\theta|, Y_2, \cdots, Y_n) = -nlog[B(\alpha,\beta)] + \sum_{i=1}^{n} \log[f(Y_i, \pi)] + (\alpha - 1) \sum_{i=1}^{n} \log[F(Y_i, \pi)]$$

$$+(\beta - 1) \sum_{i=1}^{n} \log[1 - F(Y_i, \pi)]$$

where, $f(Y_i, \pi)$ and $F(Y_i, \pi)$, $y \geq \mu$ as in (1) and (2).

$$l(\theta|, Y_2, \cdots, Y_n) = -nlog[B(\alpha,\beta)]$$
$$+ \sum_{i=1}^{n} \log\left[\frac{1}{2\sigma}exp\left(-\left(\frac{y-\mu}{\sigma}\right)\right)\right] + (\alpha - 1) \sum_{i=1}^{n} \log\left[1 - \frac{1}{2}exp\left(-\left(\frac{y-\mu}{\sigma}\right)\right)\right]$$

$$+(\beta - 1) \sum_{i=1}^{n} \log\left[\frac{1}{2}exp\left(-\left(\frac{y-\mu}{\sigma}\right)\right)\right]$$

$$(19)$$

The score function are obtained by using the differential equation in (19) with respect to $(\alpha, \beta, \mu, \sigma)$ and recall that $B(\alpha,\beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$. Hence,

$$\frac{\partial l(\theta)}{\partial \alpha} = -n\frac{\Gamma'(\alpha)}{\Gamma(\alpha)} + \frac{\Gamma'(\alpha + \beta)}{\Gamma(\alpha + \beta)} + \sum_{y=1}^{n} \log\left[1 - \frac{1}{2}exp\left(-\left(\frac{y-\mu}{\sigma}\right)\right)\right]$$

$$(20)$$

$$\frac{\partial l(\theta)}{\partial \beta} = -n\frac{\Gamma'(\beta)}{\Gamma(\beta)} + \frac{\Gamma'(\alpha + \beta)}{\Gamma(\alpha + \beta)} + \sum_{y=1}^{n} \log\left[\frac{1}{2}exp\left(-\left(\frac{y-\mu}{\sigma}\right)\right)\right]$$

$$(21)$$

$$\frac{\partial l(\theta)}{\partial \mu} = \sum_{y=1}^{n} \left[\frac{\frac{\partial}{\partial \mu}\left[\frac{1}{2\sigma}exp\left(-\left(\frac{y-\mu}{\sigma}\right)\right)\right]}{\left[\frac{1}{2\sigma}exp\left(-\left(\frac{y-\mu}{\sigma}\right)\right)\right]}\right] + (\alpha - 1) \sum_{y=1}^{n} \left[\frac{\frac{\partial}{\partial \mu}\left[1 - \frac{1}{2}exp\left(-\left(\frac{y-\mu}{\sigma}\right)\right)\right]}{\left[1 - \frac{1}{2}exp\left(-\left(\frac{y-\mu}{\sigma}\right)\right)\right]}\right]$$

$$+(\beta - 1) \sum_{y=1}^{n} \left[\frac{\frac{\partial}{\partial \mu}\left[\frac{1}{2}exp\left(-\left(\frac{y-\mu}{\sigma}\right)\right)\right]}{\left[\frac{1}{2}exp\left(-\left(\frac{y-\mu}{\sigma}\right)\right)\right]}\right]$$

$$(22)$$

$$\frac{\partial l(\theta)}{\partial \sigma} = \sum_{y=1}^{n} \left[\frac{\frac{\partial}{\partial \sigma}\left[\frac{1}{2\sigma}exp\left(-\left(\frac{y-\mu}{\sigma}\right)\right)\right]}{\left[\frac{1}{2\sigma}exp\left(-\left(\frac{y-\mu}{\sigma}\right)\right)\right]}\right] + (\alpha - 1) \sum_{y=1}^{n} \left[\frac{\frac{\partial}{\partial \sigma}\left[1 - \frac{1}{2}exp\left(-\left(\frac{y-\mu}{\sigma}\right)\right)\right]}{\left[1 - \frac{1}{2}exp\left(-\left(\frac{y-\mu}{\sigma}\right)\right)\right]}\right]$$

$$+(\beta - 1) \sum_{y=1}^{n} \left[\frac{\frac{\partial}{\partial \sigma}\left[\frac{1}{2}exp\left(-\left(\frac{y-\mu}{\sigma}\right)\right)\right]}{\left[\frac{1}{2}exp\left(-\left(\frac{y-\mu}{\sigma}\right)\right)\right]}\right]$$

$$(23)$$

Equating the score functions (20 – 23) to zero can be solved numerically by the system of non-linear equations to obtain the $\hat{\alpha}, \hat{\beta}, \hat{\mu}, \hat{\sigma}$ the MLE of $(\alpha, \beta, \mu, \sigma)$, respectively.

4.1 Application to Death Cases Data Set from COVID-19 Pandemic

The data used for the analysis is from 53 African and 124 global countries (since COVID-19 outbreak cuts across African and global countries, our analysis must also cover both) from the rest of the world consist of four different cases as: added, confirmed, recovered and deaths but only the data for deaths cases were used in the study for analysis. The reason for using data for deaths cases is due to the nature of the proposed BDEM (is a univariate distribution) and the number of deaths cases during the pandemic outbreak are on the high side daily in all affected countries according to the available online data, and there is no way we will not mentioned/talked about other cases. The data was accessed on the 16th of April, 2020 via [17] and [9] news sites. In the analysis, descriptive statistics, normal QQ plot, boxplot, Histogram estimated pdf and cdf, maximum likelihood, The values from non-African countries KS = 0.1208, LogL = 278.4168, AIC = 560.8336, BIC = 576.1147 and CAIC = 577.1147. Also, from African countries are: KS = 0.0759, LogL = 144.0245, AIC = 292.0490, BIC = 303.9302 and CAIC = 304.9302. were all presented in Tables 1 and 2, Figs. 3 and 4; and Tables 3 and 4 respectively.

5. RESULTS

The proposed distribution is validated with COVID-19 Pandemic (death cases only) data set as stated above under section 4.1.

Table 1, contains the descriptive statistics of the deaths cases data set of the global countries. The values of skewness and kurtosis are excessive in nature. The maximum likelihood estimates (MLEs) and their corresponding standard errors (in parentheses) of the model parameters values for deaths cases from global countries are given in Table 3 and we compared the fits of the BDEM, EDED, LTDED, DED and ED distributions. The values in Table 3a now indicate that the BDEM model has smaller values for the AIC, BIC and CAIC statistics (data for

global countries deaths cases) among the fitted distributions. Furthermore, BDEM model proves better fitted model among other competing models. While, Fig. 3 are the Normal QQ plots, boxplot, the Histogram, estimated pdfs and cdfs for the models with global deaths cases from COVID-19 data set. The Normal QQ and boxplot reveal that the data set does not follow normal distribution but follows exponential distribution due to its nature. Some of the observations in the data set are outlier values which requires flexible-robust model like BDEM (as it's graphically display in Fig. 1 to capture its excessive skewness and kurtosis.

Table 2, contains the descriptive statistics of the deaths cases data set of the African countries and the skewness and kurtosis are excessive in nature (their values). The maximum likelihood estimates (MLEs) and their corresponding standard errors (in parentheses) of the model parameters values for deaths cases from global countries are given in Table 3 and we compared the fits of the BDEM, EDED, LTDED, DED and ED distributions. The values in Table 4 show that the Beta Double Exponential (BDEM) model has lowest values for the AIC, BIC and CAIC statistics (data for African countries deaths cases) among the fitted distributions. Furthermore, BDEM model proves better fitted model among other competing models. While, Fig. 4 are the Normal QQ plots, boxplot, the Histogram, estimated pdfs and cdfs for the models with African countries deaths cases from COVID-19 data set. The Normal QQ and boxplot shows that the data set is not normally distributed rather it follows exponentially distributed due to the increase in numbers of deaths cases for COVID-19 worldwide as at the time the data was collected. Some of the observations in the data set are outrageous values which requires robust-flexible model like BDEM to capture its excessive skewness and kurtosis.

Tables 3 and 4 present the values from the Analysis of MLEs of the model parameters, the associating SEs (starred) and the statistics - 2LogL, AIC, BIC and CAIC using Non-African and African death cases

Table 1. Descriptive statistics for global countries COVID-19 deaths Cases

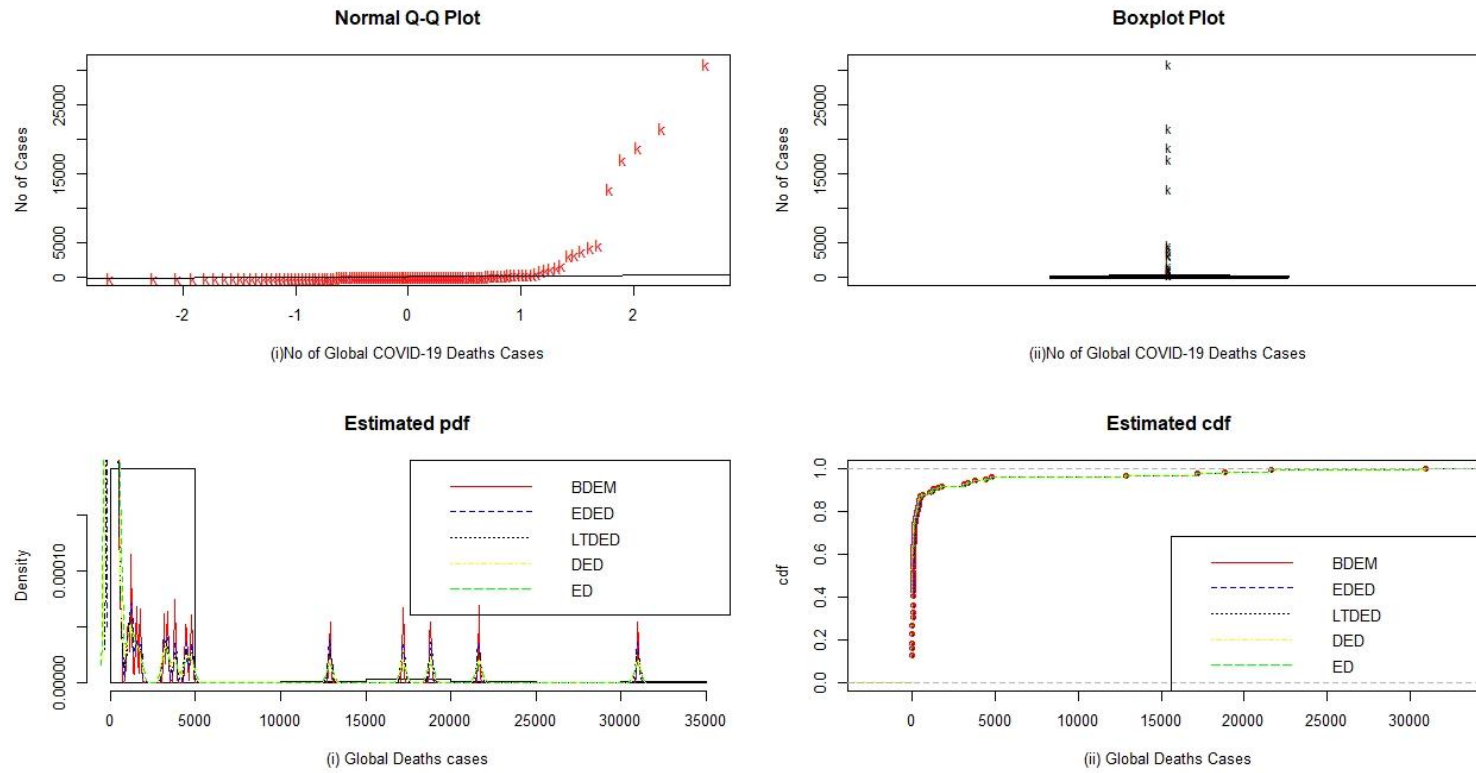| Min	Q1	Median	Mean	Q3	Max	Skewness	Kurtosis
0.00	4	27	1095	154.00	30985	5.097355	30.17454

Fig. 3. Contains the normal QQ plots, boxplots, estimated pdf and cdf of global deaths cases for COVID-19 pandemic
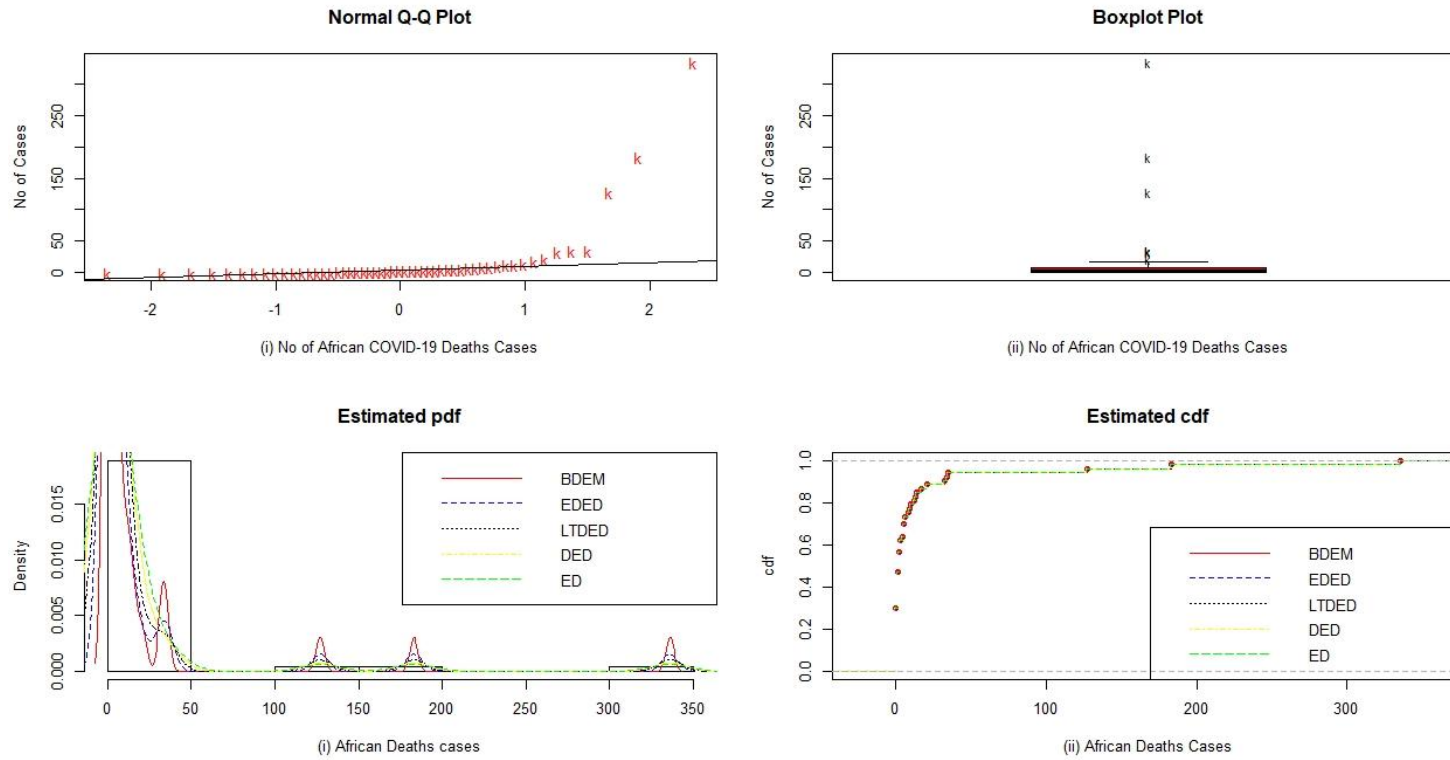
Fig. 4. Contains the normal QQ plots, boxplots, estimated pdf and cdf of African deaths cases for COVID-19 pandemic

Table 2. Descriptive statistics for African countries COVID-19 deaths cases

Min	Q1	Median	Mean	Q3	Max	Skewness	Kurtosis
0.00	0.00	2.00	17.17	8.00	336.00	4.686986	25.78003

Table 3. Non-African deaths-cases

Model	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\mu}$	$\hat{\sigma}$	KS	$-2LogL$	AIC	BIC	CAIC
BDEM	0.5060	1.8970	1.4040	6.1100	0.1208	278.4168	560.8336	576.1147	577.1147
	0.0001*	0.0047*	0.0510*	0.0612*					
EDED	2.5000		2.0000	5.5000	0.1214	366.8265	736.6530	748.1138	749.1138
	0.0033*		0.0112*	0.0104*					
LTDED		1.2000	1.3000	6.0000	0.1975	378.6427	760.2854	771.7462	772.7462
		0.0188*	0.0189*	0.0477*					
DED			5.5000	7.0000	0.2003	421.1143	844.2286	851.6892	852.6892
			0.0084*	0.0151*					
ED			1.3800	4,1500	0.2554	421.8456	845.6912	853.3318	854.3318
			0.0012*	0.0003*					

Table 4. African deaths-cases

Model	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\mu}$	$\hat{\sigma}$	KS	$-2LogL$	AIC	BIC	CAIC
BDEM	0.7162	1.6150	1.9650	1.0320	0.0759	144.0245	292.049	303.9302	304.9302
	0.0001*	0.0089*	0.0072*	0.0063*					
EDED	1.5020		1.6080	1.5000	0.1434	242.8147	488.6294	497.5403	498.5403
	0.0105*		0.0046*	0.0237*					
LTDED		1.5020	1.6070	1.5000	0.1889	253.2551	509.5102	518.4211	519.4211
		0.0129*	0.0056*	0.0284*					
DED			1.5030	4.9420	0.2427	299.0669	600.1338	606.0744	607.0744
			0.0026*	0.0006*					
ED			2.3210	6.7080	0.1260	299.1107	600.2214	606.1620	607.1620
			0.0013*	0.0002*					

6. DISCUSSION

Here, we have improved the characteristics of the double exponential (DE) distribution by adding two shape parameters α and β to the model. Both equations (4) and (5) become the pdfs of the proposed model. The applicability of the new model to fit deaths-cases on COVID-19 pandemic data provides better fits and captured the data due to its flexibility and robustness than parent and generated models.

Fig. 1 shows the plot of the pdf of the proposed model at different values with several shapes; which is clearly shown that the model can accepts increasing, decreasing and bathtub. [14]. Also Fig. 2 depicts the pdfs of the models considered in the study and it is shown that BDEM has capability to accommodates the data sets than other models. In Tables 1 and 2, we presented the nature of the data especially the excessive skewness and kurtosis. [18], [19] stated that normal coefficient of any normal data is 0 and coefficient of kurtosis is 3. However, the values of the skewness and kurtosis are greater than 0 and 3 which shows excessive skewness and kurtosis that requires more flexible and robust model to accommodate and capture the outrageous values in the data sets.

Furthermore, another way of detecting non normal data or values is depicted in Figs. 3 and 4 where normal QQ plots and the boxplots of the data used. It is then shows that the data is skewed in nature, while the Figures also presented the estimated pdf and cdf of the BDEM and its sub cases. Therefore, BDEM is flexible and robust than other considered models. Table 3 and 4 consist the MLEs of the model parameters, the associating SEs (starred) and the statistic Log-likelihood, AIC, BIC and CAIC of the new model and other models. Meanwhile, BDEM shows its flexibility and robustness over others.

7. CONCLUSION

We investigate a new model generated by convoluting double exponential distribution and logit of beta link function introduced by [12] In the work, some of the properties of the propose model emanated include some special cases like exponentiated double exponential distribution, Lehmann Type II double exponential distribution and two-parameter exponential. In this study, we also obtain the following: distribution function, survival function, hazard rate function, moments, generating function, skewness and kurtosis. Maximum likelihood is discussed. An application to COVID-19 deaths cases data set of 124 global and 53 African countries were used and the fit of the BDEM proved flexible and robust to the fits than other distributions despite the fact that the data is skewed. The model is not limited in fitting only COVID-19 data set, also it can be used in modelling data from various infectious disease such as Ebola, Polio, Cholera, Norovirus, Measles, HIV/AIDs, Hepatitis etc. Finally, the proposed BDEM gives better fits to the data sets than its competitors.

CONSENT

It is not applicable.

ETHICAL APPROVAL

It is not applicable.

ACKNOWLEDGEMENT

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. WHO. Novel Coronavirus –China; 2020. Coronavirus-china/en/ Accessed 20th April, 2020. Available:https://www.who.int/csr/don/12-january-2020-novel-

2. Smarat KD, Mahbubur Rahman MD, Umme RS, Arpita Howlader. Analyzing the epidemiological outbreak of COVID19: A visual exploratory data analysis approach. Journal of Medical Virology. 2020;92(6): 632-638. Available:https://doi.org/10.1002/jmv.25743

3. Yoo JH. The fight against the nCoV outbreak: An arduous march has just begun. Journal of Korean Medical Science. 2020;35(56). eISSN 1598 - 6357. ISSN 1011 - 8934. Accessed 22nd April, 2020. Available:https://doi.org/10.3346/jkms.2020.35.e56

4. ISARIC. A global federation of clinical research networks, providing a proficient,

coordinated and research response to outbreak prone infectious diseases. COVID-19reports.
Available:https://isaric.tghn.org/covid-19-data-management-hosting-contributors/

5. Zea LM, Rodrigo BS, Marcelo B, Santos AM, Corderio GM. The beta exponentiated pareto distribution with application to bladder cancer susceptibility. International Journal of Statistics and Probability. 2012; 1(2): 8-19.

6. Hamdy MS, Mahmoud AS. The Generalized Weibull-Exponential distribution: Properties and applications. International Journal of Statistics and Applications. 2014;4(2):102–112.

7. Badmus NI, Bamiduro TA. Some statistical properties of exponentiated weighted Weibull distribution. global Journal of Science Frontier Research: For Mathematics and Decision Sciences. 2014;14(2):12 -22.

8. Handique Laba, Chakraborty Subrata. Beta generated Kumaraswamy-G and other new families of distribution. arXiv preprint arXiv: 1603.00634. 2016;1 – 34.

9. Scooper Online News. Latest on coronavirus outbreak (official figures from represent active countries and regions); 2020.
Accessed 16th April, 2020.
Available: https://m.scoopernews.news

10. Suk-Bok K, Young-Suk C, Jun-Tae H. Estimation for the double-exponential distribution based on type-ii censored samples. Journal of Korean Data and Information Science Society. 2005;16(1): 115 – 126.

11. Jagannathan K, Gupta AK, Nguyen TT. Skewed double exponential distribution and its stochastic representation. European Journal of Pure and Applied Mathematics. 2009;2,1:1 – 20.
ISSN 1307 – 5543
Available:www.ejpam.com

12. Jones MC. Families of distributions arising from distribution of order statistics test. 2004;13: 1-43

13. Morais A, Cordeiro G, Audrey H C. The beta generalized logistic distribution. Brazilian Journal of Probability and Statistics. 2013;27(2):185 – 200.

14. Ortega EMM, Cordeiro GM, Kattan MW. The log-beta weibull regression model with application to predict recurrence of prostate cancer. Regular Article, Statistical Paper. 2013;54(1):113 – 132.

15. Pascoa MAR, de Paiva CMM, Cordeiro GM, Ortega EMM. The log-kumaraswamy generalized gamma regression model with application to chemical dependency data. Journal of Data Science. 2013;11:781-818.

16. Cordeiro GM, Alexandra C, Ortega EMM Sarabia JM. Generalized beta generated distributions. ICMA Centre, Discussion Papers in Finance DP. 2011;05:1 – 29

17. NCDC. Advisory on the use of masks by members of the public without respiratory symptoms. COVID-19.NCDC.GOV.NG. 2020;1-4.
Accessed April 19, 2020.
Available:https://covid19.ncdc.gov.ng

18. Ezequiel U. Relaxing the assumptions in the linear classical model. University of Valencia. 2013;9: 1 – 45.

19. Delignette-Muller ML, Dutang C. An R Package for Fitting Distributions. 2014;1 – 23.
Available: http://cran.r-project.org/package