International Invention of Scientific Journal
Available Online at http://www.iisj.in
eISSN: 2457-0958
Volume 03 | Issue 08 | August, 2019 |

Original Article

OPEN ACCESS

Received: June 24, 2019
Accepted: August 03, 2019
Published: August 07, 2019

***Corresponding Author:**
*** BADMUS, N. IDOWU2**
Department of Statistics, Yaba College of Technology, Yaba, Lagos, Nigeria
Email: idoscomus2011@gmail.com

Application of Parametric Regression Model to Economic Data

Efuwape, Biodun T[1] and Badmus, N. Idowu[2]*

[1]Department of Mathematical Sciences, Olabisi Onabanjo University, Ago-Iwoye, Nigeria
[2]*Department of Statistics, Yaba College of Technology, Yaba, Lagos, Nigeria

Abstract

Parametric regression models have been in existence years ago in literature for problems in several areas such as: survival analysis, reliability, biology and hydrology. In this research, a new parametric regression model is developed as an important model to accommodate skewed economic data. Some statistical properties of the model distribution are derived, and estimation of parameters is also obtained maximum likelihood method. The random variable follows the distribution of the proposed model. Therefore, the proposed model is applying to economic data set.

Key Words: Economic Data, regression model, reliability, skewed, survival analysis

Introduction

Parametric model can calculate the distribution form of survival time. It can also capture the behaviour of a lifetime and any data that exhibited skewness in any field in economics, medicine, biology, engineering, quality control among others. There are advantages of parametric model to data that exhibited excess skewness include; full maximum likelihood is be used for estimating parameters and the distribution of the dependent variable follows the distribution of the residual term. Regression model allows the researcher with important material and allowing predictions to be done about the past, the present and the future events in different forms. Economic statistics are quantitative measures in nature pointing an actual economy, past or present. All these can be found in time series form that covers more than one time period. Many works have been done using parametric model in literature such as; Cancho et al. 2009 investigated on the log-exponentiated weibull regression models with cure rate, Ortega et al. 2013 studied the Kumaraswamy generalised Gamma regression model with application to chemical dependency data, Mahmoud et al. 2015 worked on log-beta log-logistic regression model, likewise, Badmus, et al. 2016 discussed and based their work on generalized modified weighted Weibull distribution among others. However, we propose a location-scale regression model, which will refer to as log-Lehmann Type II modified weighted Rayleigh regression model (LLMWRRM) based on an asymmetric continuous distribution proposed by (Badmus et al., 2017) an extension of the beta modified weighted Rayleigh (BMWR) distribution. Therefore, the paper is arranged as follows: Section two, discussed the propose log- log-Lehmann Type II modified weighted Rayleigh (LLMWR) distribution, some of its statistical properties and the log-Lehmann Type II modified weighted Rayleigh regression model. We estimate the model parameters in section three using maximum likelihood and the observed information matrix are also obtained. Section four carried out the analysis of the data (economic data) and selection of model: e.g the AIC, BIC and CAIC on the proposed model and its sub cases. Lastly conclusion is drawn in section 5.

The Log-Lehmann Type II Modified Weighted Rayleigh (LLMWR)

Here, we extend the BMWR distribution (Badmus et al., 2017) to LLMWR distribution. The BMWR distribution is developed from the Modified Weighted Rayleigh mentioned in the work of Aleem et al. 2013 where, the density and its corresponding distribution function are given as follows:

$$f_{MWR|\{\ \gamma,\alpha,\lambda\}}(x) = 2\lambda(\alpha\gamma^2+1)x.\exp(-\gamma(\alpha\gamma^2+1)x^2 \quad (1)$$

and

$$F_{MWR|\{\ \gamma,\alpha,\lambda\}}(x) = 1-\exp(-\gamma(\alpha\gamma^2+1)x^2 \quad (2)$$

where, λ is scale parameter, α and γ are shape parameters.

Then, both the pdf and cdf of BMWR are given respectively

$$f_{BMWR|\{\ \gamma,\alpha,\lambda,a,\mathrm{b}\}}(x) = B(a,\mathrm{b})^{-1}[1-\exp(-\gamma(\alpha\gamma^2+1)x^2]^{a-1}$$

$$`[\exp(-\gamma(\alpha\gamma^2+1)x^2]^{b-1}\ 2\lambda(\alpha\gamma^2+1)x\exp(-\gamma(\alpha\gamma^2+1)x^2 \quad (3)$$

and

$$F_{BMWR}(x) = \frac{B(x;\,a,\ \ b)}{B(a,b)} \quad (4)$$

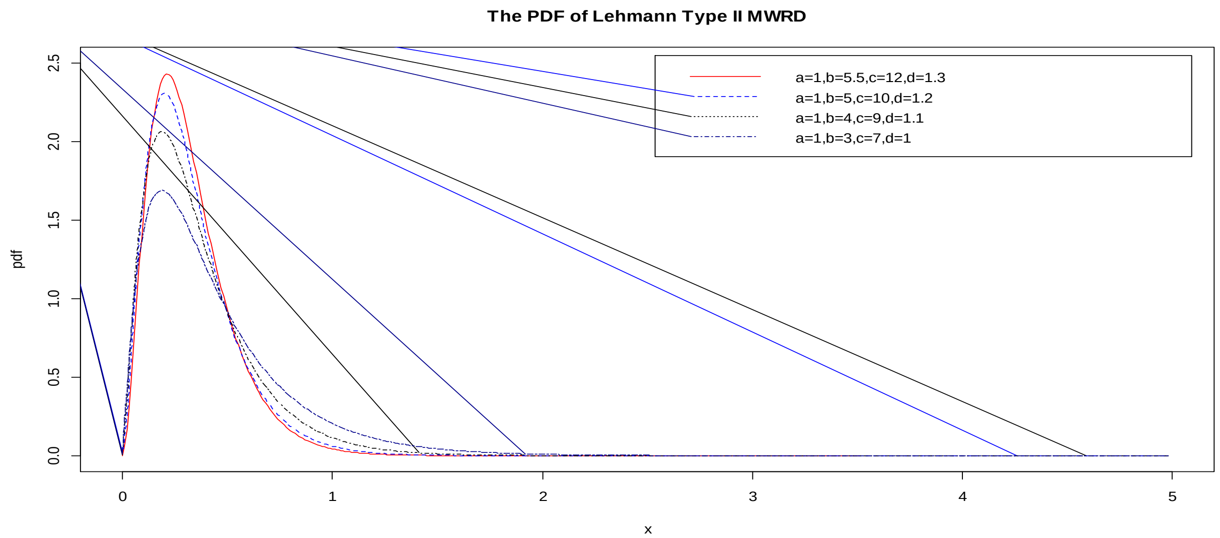
Now, if $a = 1$ in (3) above, the expression becomes (5) which is called the density function of the Lehmann Type II Modified Weighted Rayleigh (LMWR) distribution.

$$f_{LMWR|\{\ \gamma,\alpha,\lambda,1,\mathrm{b}\}}(x) = \mathrm{b}[\exp(-\gamma(\alpha\gamma^2+1)x^2]^{b-1}\ 2\lambda(\alpha\gamma^2+1)x\exp(-\gamma(\alpha\gamma^2+1)x^2 \quad (5)$$

and the associate corresponding distribution function is gives as

$$F_{LMWR|\{\ \gamma,\alpha,\lambda,1,\mathrm{b}\}}(x) = [1-(1-x)^b] \quad (6)$$

where, $x\ > 0, f(x) = \frac{d}{dx}F(x)$ and $b > 0$ is a shape parameter added to the existing one in the parent (MWR) distribution.



Also, the LLMWR distribution is defined from logarithm of the LMWR random variable to give a better fitting of skewed data. The MWR density function in (1) with parameters $(\lambda, \gamma, \alpha,) > 0$ can be re-written in a simplified version of Rayleigh as stated below:

$$f(x) = 2\beta^{\alpha} + 1\left[\left(\frac{\alpha}{\gamma}\right)^{\alpha} x\, exp\left(-\left(\frac{x}{\gamma}\right)^{\alpha}\right)\right] \quad (7)$$

We got (7) LLMWR distribution using transformation method by setting $Y = \log(x)\, i.e.\ x = e^{y}$, $\alpha = \frac{1}{\sigma}\ and\ \mu = \log(\gamma)\ i.e. \gamma = e^{\mu}$ and substituting the outcome into (7), we get

$$f(y) = 2\lambda^{\frac{1}{\sigma}} + 1\left(\frac{\frac{1}{\sigma}}{e^{\mu}}\right)^{\frac{1}{\sigma}} (e^{y})\, exp\left(-\left(\frac{e^{y}}{e^{\mu}}\right)^{\frac{1}{\sigma}}\right).e^{y}$$

$$= \frac{2\lambda^{\frac{1}{\sigma}+1}}{\sigma}\left(\frac{e^{y}}{\sigma} - \frac{e^{\mu}}{\sigma}\right).exp\left(-\left(\frac{e^{y}}{\sigma} - \frac{e^{\mu}}{\sigma}\right)\right)$$

$$f(y) = (2\lambda + 1).exp\left(\frac{y-\mu}{\sigma}\right).exp\left(-exp\left(\frac{y-\mu}{\sigma}\right)\right) \quad (8)$$

(8) now becomes the pdf of the LLMWR distribution; and can also be written as the LMWR distribution by convoluting the beta function in equation (8), then gives

$$g(y) = 2\lambda^{\alpha} + 1\left[\left(\frac{\alpha}{\gamma}\right)^{\alpha} y\, exp\left(-\left(\frac{y}{\gamma}\right)^{\alpha}\right)\right]\left[2\lambda^{\alpha} + 1\left(exp\left(-\left(\frac{y}{\gamma}\right)^{\alpha}\right)\right)\right]^{b-1} \quad (9)$$

$y \sim LMWR\ (b, \lambda,\ \gamma, \alpha)$ distribution; λ, γ and α are shape, weight and scale parameter in the existing distribution while, b is the shape parameter added to it.

If X is a random variable with the LMWR density function (5). Properties of the proposed (LLMWR) distribution are obtained; and defined the random variable $Y = \log(X)$. ((Ortega *et al*. (2013), Badmus *et al*. 2016. Hence, the density function of Y had been transformed in (7); and is defined as

$$f_{LLMWR}(y; b, \lambda, \mu, \sigma) = b(2\lambda + 1)exp\left(\frac{y-\mu}{\sigma}\right).exp\left(-exp\left(\frac{y-\mu}{\sigma}\right)\right)$$

$$\left[(2\lambda + 1)exp\left(-exp\left(\frac{y-\mu}{\sigma}\right)\right)\right]^{b-1} \quad (10)$$

where $-\infty < y < \infty,\ \sigma > 0\ and\ -\infty < \mu < \infty$.

Equation (10) is the LLMWR distribution; μ, σ, λ and b is location, dispersion, weight and shape parameter; and $Y = \log(X) \sim LLWMR\ (b, \lambda, \mu, \sigma)$.

The survival function to (10) is given by

$$S_{LLMWR}(y) = 1 - b\int_0^{F(y)}(1-x)^{b-1}dx = 1 - I_{F(y)}(1,b) \tag{11}$$

where, $F_{(y)} = \left[1 - b(2\lambda+1)exp\left(-exp\left(\frac{y-\mu}{\sigma}\right)\right)\right]$

Moments and Generating Function:

The rth ordinary moment of the LLMWR distribution is defined as

$$\mu'_{LLMWR(r)} = E(S^r) = b\int_{-\infty}^{\infty} Z^r\left[(2\lambda+1)exp(z).exp\left(-exp(z)\right)\right]$$

$$\left[(2\lambda+1)exp\left(-exp(z)\right)\right]^{b-1}dz$$

Therefore, $q = e^z$ is the expanded binomial term that gives

$$\mu'_{LLMWR(r)} = b\sum_{i=0}^{\infty}(-1)^i\binom{b-1}{i}\int_0^{\infty} log(s)^r$$

$$\left\{(2\lambda+1)exp(q).exp\left(-exp(q)\right)\left[(2\lambda+1)exp\left(-exp(q)\right)\right]^{(i+1)-1}q\right\}dq$$

$$I_{(r,(i+1))} = \left(\frac{\partial}{\partial p}\right)^r\left[\left((i+1)\right)^{-p}\Gamma(p)\right]|_{p=1}$$

Pascoa *et al*. 2013 and Badmus *et al*. 2016.

and thus

$$\mu'_r = b\sum_{i=0}^{\infty}(-1)^i\binom{b-1}{i}I_{(r,(i+1))} \tag{12}$$

(12) describes the moments of the LLMWR distribution.

The moment generating function (MGF) of X, such that $M(t) = E(e^{tx})$ is given by

$$M_{LBMWW}^{(t)} = b\sum_{i=0}^{\infty}(-1)^i\binom{b-1}{i}\int_0^{\infty} W^t$$

$$\left\{\left[(2\lambda+1)exp(q).exp\left(-exp(q)\right)\right]\left[(2\lambda+1)exp\left(-exp(q)\right)\right]^{(i+1)-1}q\right\}dq$$

Hereafter,

$$M(t) = \Gamma(t+1)b\sum_{i=0}^{\infty}(-1)^i\binom{b-1}{i}\left[\left((i+1)-1\right)\right]^{-(t+1)} \tag{13}$$

The first-four moments, skewness and kurtosis of the LLMWR distribution are obtained from the rth ordinary moment of the LLMWR as expressed in (12).

$$\mu'_{LLMWR(r)} = \int_0^\infty \log Q^r\{b(2\lambda+1)[1-X(q)]^{b-1}dx\,(q)\} \tag{14}$$

where,

$$M(a) = \left(1-(\theta)exp\left(-exp(q)\right)\right)$$

$$\theta = (2\lambda+1), \qquad k(q) = \exp\left(\frac{y-\mu}{\sigma}\right)$$

$$\mu'_{LLMWR(r)} = \left[\frac{2(\theta)\lambda^{-\frac{r}{2}}\Gamma\left(\frac{r}{2}+1\right)(\theta)^{-\frac{r}{2}-1}}{\sigma}\right] b\sum_{i=0}^{\infty}(-1)^i\binom{b-1}{i}$$

$$\left\{\int_0^\infty log\left[1-(\theta)exp\left(-k(q)\right)\right]^{(i+1)-1}dq\right\} \tag{15}$$

where,

$$S = \frac{b\sum_{i=0}^{\infty}(-1)^i\binom{b-1}{i}\int_0^\infty log\left[1-(\theta)exp\left(-k(q)\right)\right]^{(i+1)-1}dq}{\sigma}$$

Furthermore, the 1st to 4th non-central moments μ'_r by substituting for r = 1, 2, 3 and 4 respectively in equation (15) it's resulted as given below:

$$\mu'_{LLMWR(r)} = E_{LLMWR}(x) = \left[\frac{2(\theta)\lambda^{-\frac{r}{2}}\Gamma\left(\frac{r}{2}+1\right)(\theta)^{-\frac{r}{2}-1}}{\sigma}\right] b\sum_{i=0}^{\infty}(-1)^i\binom{b-1}{i} \tag{16}$$

The 1st moment of the LLMWR is obtained from (16). Then, the mean, second, third and fourth moments of the LLMWR distribution are given as follows:

$$\mu = \mu'_1$$

$$\mu_2 = \mu'_2 - \mu^2$$

$$\mu_3 = \mu'_3 - 3\mu\mu'_2 + 2\mu^3 \text{and}$$

$$\mu_4 = \mu'_4 - 4\mu\mu'_3 + 6\mu^2\mu'_2 - 3\mu^4$$

Where

$$\mu'_1 = S\left[2(\theta)\lambda^{-\frac{1}{2}}\Gamma\left(\frac{1}{2}+1\right)(\theta)^{-\frac{1}{2}-1}\right] = S[2(\theta)\lambda^{-\frac{1}{2}}\Gamma\left(\frac{3}{2}\right)(\theta)^{-\left(\frac{3}{2}\right)} \tag{17}$$

$$\mu'_2 = S\left[2(\theta)\lambda^{-\frac{2}{2}}\Gamma\left(\frac{2}{2}+1\right)(\theta)^{-\frac{2}{2}-1}\right] = 2S[2(\theta)\lambda^{-1}\Gamma(2)(\theta)^{-(2)} \tag{18}$$

$$\mu_3' = S\left[2(\theta)\lambda^{-\frac{3}{2}}\Gamma\left(\frac{3}{2}+1\right)(\theta)^{-\frac{3}{2}-1}\right] = 6S[2(\theta)\lambda^{-\frac{3}{2}}\Gamma\left(\frac{5}{2}\right)(\theta)^{-\left(\frac{5}{2}\right)} \quad (19)$$

$$\mu_4' = S\left[2(\theta)\lambda^{-\frac{4}{2}}\Gamma\left(\frac{4}{2}+1\right)(\theta)^{-\frac{4}{2}-1}\right] = 24S[2(\theta)\lambda^{-2}\Gamma(3)(\theta)^{-(3)} \quad (20)$$

Measures of Skewness ω_1 and excess kurtosis, ω_2 are given below respectively

$$\omega_1 = \frac{\mu_3}{\mu_2^{3/2}} \quad (21)$$

$$\omega_2 = \frac{\mu_4}{\mu_2^2} - 3 \quad (22)$$

The Log-Lehmann Type II Modified Weighted Rayleigh Regression Model:

We linked the response variable y_i and vector $\boldsymbol{X}_i^T = (x_{i1}, \dots, x_{ip})$ of explanatory variables x; and the location-scale regression model as given below

$$y_i = X_i^T\beta + \sigma z_i\,, \qquad i = 1, 2, \dots, n \quad (23)$$

Pescim *et al*. 2013, Mahmoud *et al*. 2015, Badmus *et al*. 2016 among others have used the model in their work in literature; and the random error z_i has density function

$$\psi(z;\ b,\lambda) = b\left[(2\lambda+1)exp(z).exp\left(-exp(z)\right)\right]\left[(2\lambda+1)exp\left(-exp(z)\right)\right]^{b-1} \quad (24)$$

with parameters $\boldsymbol{\beta} = (\beta_1, \dots \dots .., \beta_v)^T$, (b, λ, σ) > 0, are unknown parameters. The parameter $\mu_i = \boldsymbol{X}_i^T\boldsymbol{\beta}$ is the location of y_i. The location parameter vector $\boldsymbol{\mu} = (\mu_1, \dots, \mu_n)^T$ stands for a model $\boldsymbol{\mu} = \boldsymbol{X}^T\boldsymbol{\beta}$where $\boldsymbol{X} = (X_1, \dots, X_n)^{\mathrm{T}}$ is a known model matrix. The LLMWR regression model (10) made it easy for fitting many difficult and skewed data.

Estimation of Model Parameter Distribution

The likelihood function (LF) for the vector of parameters $\tau = (b, \lambda, \sigma, \beta^T)^T$from model (23) has the form $l(\tau) = \sum_{i\epsilon F} log[f(y_i)] + \sum_{i\epsilon C} log[s(y_i)]$, where $f(y_i)$ is the density function (10) and $S(y_i)$ is the survival function (11) of Y_i.

The log-likelihood (LL) function for τ becomes:

$$l(\tau) = -nlog\{log(\sigma) + logb\}\left[(2\lambda+1)\left\{\sum_{i\epsilon F} exp(z).exp\left(-exp(z)\right)\right\}\right]$$

$$+ (b-1)\sum_{i\epsilon F} log\left[(2\lambda+1)\sum_{i\epsilon F}\left\{\left\{exp\left(-exp(z)\right)\right\}\right\}\right]$$

$$+\sum_{i\epsilon C} log\left\{1 - I_{\left[(2\lambda+1)\sum_{i\epsilon C}\left\{\left\{exp(-exp(z))\right\}\right\}\right]}\right\}^{(b)} \quad (25)$$

the number (n) of uncensored observations and $z_i = \frac{(y - X_i^T \beta)}{\sigma}$.

Then, MLE $\hat{\tau}$ of the vector τ of unknown parameters can be computed by maximising the LF in (25); and fitted LLMWR model leads to the estimated survival function of Y for any individual with explanatory vector **x**

$$R_{LLMWR}\left(y;\ \hat{b}, \hat{\lambda}, \hat{\sigma}, \hat{\beta}^T\right) =$$

$$1 - I_{\left[(2\hat{\lambda}+1)\left\{\left(1-exp\left(-exp\left(\frac{y-x^T\hat{\beta}}{\hat{\sigma}}\right)\right)\right)+exp\left(\left((2\hat{\lambda}+1)exp\left(\frac{y-x^T\hat{\beta}}{\hat{\sigma}}\right)\right)\right)\right\}\right]^b} \tag{26}$$

Data Analysis (Application to Economic Data)

We applied the model to secondary data (economic data) extracted from Nigeria statistical bulletin. The data contain GDP and external reserve from 1981 to 2016 (35 years). The GDP is the response and external reserve is explanatory variable. The model is

$$y_i = \beta_0 + \beta_1 x_{i1} + \sigma z_i, \tag{27}$$

and variable $y_i = log(t_i)$ follow the log LLMWR distribution given in (10), and the random errors s_i has the density function (23); $i = 1, \ldots, 35$. For MLEs, we used R code and procedure NLMixed in SAS in computing model parameters and data exploratory analysis. Iterative maximization of the logarithm of the LF (25) taken initial values for λ, also σ taken from the fit of the LMWR regression model with b = 1.

Exploratory Data Analysis (EDA):

Table 1: summary: Descriptive Statistics of GDP (Blood Pressure)

Mini	FirstQut.	Median	Mean	Third Qut.	Maxi	Skewness	Kurtosis
94.33	427.50	4189	17830	19610	94140	1.6882	4.4055

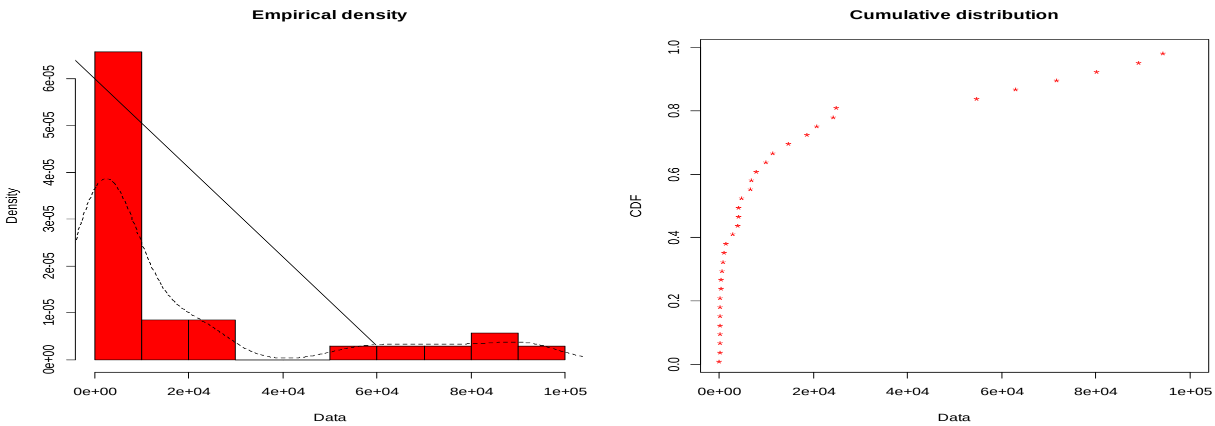
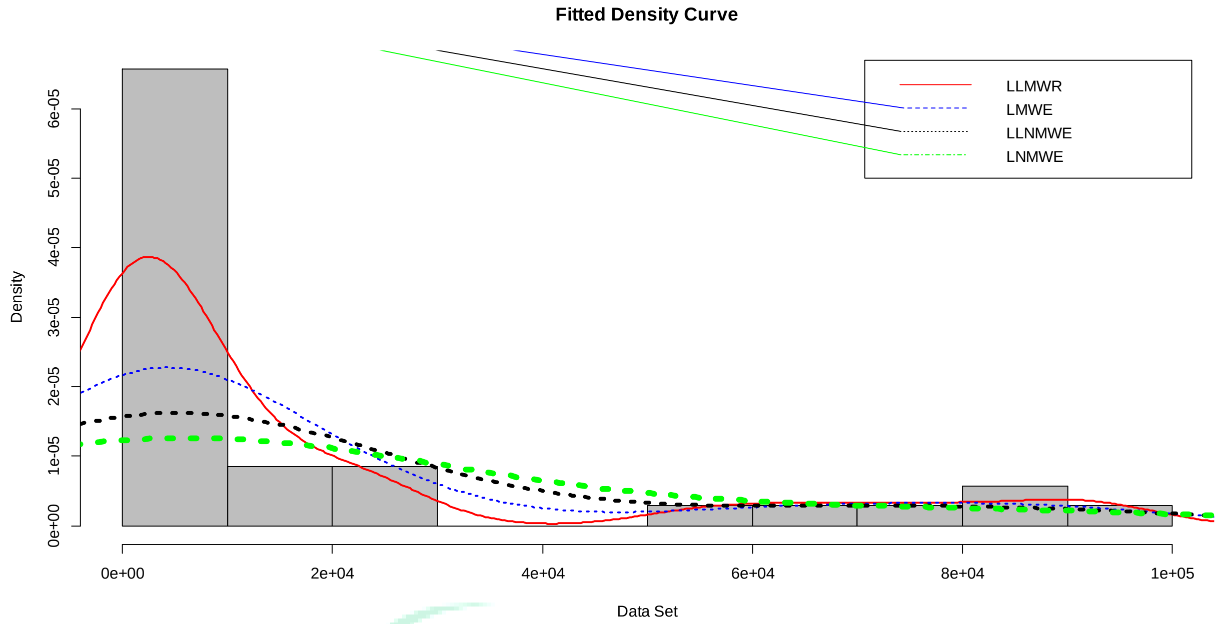


Figure 1: Plots of GDP (cumulative distribution)



Summary of the Analysis:

Table 2: MLEs and their standard error (in brackets) for data set.

Model	Estimates				
	$\hat{b}$	$\hat{\lambda}$	$\hat{\sigma}$	$\hat{\beta}_0$	$\hat{\beta}_1$
LLMWR	401.800 (0.0080)	100.500 (0.0000)	196.430 (0.2210)	-61.910 (1.5540)	38.590 (1.5540)
LMWR	--	500.500 (0.0000)	400.050 (0.0000)	-68.500 (0.0000)	32.000 (0.0000)
LNMWR	--	--	299.720 (0.1997)	-89.980 (2.0142)	60.523 (2.0142)
LLNMWR	10.500 (0.000)	--	349.790 (0.2037)	-141.330 (2.0142)	109.170 (2.0142)

Table 3: The Statistics $-2l$, AIC, CAIC and BIC statistics.

Model	$-2l$	*AIC*	*CAIC*	*BIC*
LLMWR	176.8	186.8	188.9	194.6
LMWR	189.8	197.8	199.1	204.0
LNMWR	184.9	190.9	191.7	195.6
LLNMWR	181.2	189.2	190.5	195.4

Discussion/Result:

In figure 1 and 2, we displayed the GDP line plot, Normal Q-Q plot, Boxplot, histogram plot, density plot, ecdf plot, showed the empirical density plots and cumulative distribution plot respectively.

Table 1 contains the exploratory data analysis that shows the degree level of skewnes of the GDP. Table 2 consists of MLEs for LLMWR, LMWR, LNMWR and LLNMWR regression models fitted to the economic data. Then, LR statistic is used for testing the hypotheses H_0: $b = 1$ versus H_1: H_0 is not true. The LMWR is compared with LLMWR regression model; it yields $d = 2\big(-176.8 - (-189.8)\big) = 26.0$ with (p-value < 0:0001). However, this gives the potentiality of the LLMWR model.
Thereafter, Table 3 contains the log-likelihood, Akaike Information Criterion (AIC), Corrected Akaike's Information Criterion (CAIC) and Bayesian Information Criterion (BIC) to compare with the LLMWR, LMWR, LNMWR and LLNMWR. Then, lower values of the *AIC, AICc and BIC* indicate the most efficient

model. Hence, the LLMWR regression model has better performance than other competing models (LMWR, LNMWR and LLNMWR) in all the criteria listed in table above.
The regression model was estimated as

$$y = -61.91 + 38.59x_1 \tag{28}$$

Equation (28) is interpreted as follows:

The estimated βs (parameters of the regression model) of the proposed model in table 2 above are positive except the intercept. There will be an increase of 38.59 in y for a unit change in x_1. (i.e external reserve has positive impact on Nigeria economy).

Conclusion

A new log-Lehmann Type II modified weighted Rayleigh (LLMWR) distribution and some of it properties were properly obtained. This was extended to LLMWR regression model using location-scale regression model method. Also, we investigated and derived the estimation procedure using method (MLEs).
The model was applied to economic data and the values of AIC, CAIC and BIC for the proposed LLMWR Regression Model are less than LMWR, LNMWR and LLNMWR regression models. Therefore, the developed LLMWR regression model provides better fit, flexible and performs more than other competing Regression Models mentioned above.

References

1) Aleem, M., Sufyan M., Khan, N. S. (2013). A class of Modified Weighted Weibull Distribution and its Properties. America Review of Mathematics and Statistics, 1(1) 29 – 37.
2) Badmus, Nofiu. I., Bamiduro, Adebayo T., Akingbade, Adekunle A. and Rauf-Animasaun, Monsuru A. (2016). Estimation of Parameters in Generalised Modified Weighted Weibull Distribution: International Journal of Science and Society, Yabatech, Vol.4, No. 2 11– 23.
3) Badmus, N. I.; Alakija, T. O and Olanrewaju, G. O (2017). The Beta-Modified Weighted Rayleigh Distribution. Application to Virulent Tubercle Disease. International Journal of Mathematical Analysis and Optimization: Theory and Application. Vol. 2017, pp. 154 – 167.
4) Cancho, V. G., Ortega, E. M. M. and Bolfarine, H. (2009). The Log-Exponentiated-Weibull regression models with cure rate: Local influence and residual analysis. Journal of Data Science7: 433-458.
5) Mahmoud, R. M., EL- Sheikh, A. A., Morad, N. A. and Ahmad, M. A. M. (2015). Log-Beta Log-Logistic Regression Model. International Journal of Sciences: Basic and Applied Research 22.2:389 – 405.
6) Ortega,E. M. M., Cordeiro, G. M. and Kattan, M. W. (2013). The log-beta Weibull regression model with application to predict recurrence of prostate cancer. Regular Article, Statistical Paper 54.1: 113 – 132.
7) Pascoa, M. A. R., dePaiva,C. M. M.,Cordeiro,G. M. and Ortega,E. M. M. (2013). The Log-Kumaraswamy Generalized Gamma Regression Model with Application to Chemical Dependency Data. Journal of Data Science 11: 781-818.
8) Pescim, R. R. Demetrio, C. G. B. Cordeiro, G. M. Ortega, E. M. M. and Urbano, M. R. (2013). The log-beta generalized half-normal distribution. Journal of Statistical Theory and Applications12.4: 330-347.