



Modeling the Daily Number of Reported Cases of Infection from the COVID-19 Pandemic in Nigeria: A Stochastic Approach

Patrick Osatohanmwon^{1,*}, Francis O. Oyegun² and Sunday M. Ogbonmwon³

¹ Department of Statistics, University of Benin, Benin City, Nigeria
e-mail: profpato2014@gmail.com

² Department of Statistics, University of Benin, Benin City, Nigeria

³ Department of Statistics, University of Benin, Benin City, Nigeria

Abstract

The focus of this paper is to present a stochastic model to capture the random behavior of the number of reported daily infections due to the Corona Virus (COVID-19) in Nigeria. The model expressed in form of a distribution function has five parameters. The model was fitted to the logarithm of the reported daily number of infection cases for the time period March 18th - June 11th, 2020. While the results obtained established the adequacy of the model in fitting and explaining the random behavior of the number of reported daily infections, it was also possible to use the model to study the situation of the number of infections exceeding certain thresholds. The procedure for the determination of these thresholds was established and a number of them were estimated for some given return periods.

1 Introduction

Infectious disease insurgency such as those resulting from an epidemic or a pandemic is one with a very long history with the most elaborate documentation

Received: July 26, 2020; Accepted: August 27, 2020

2010 Mathematics Subject Classification: 62B15, 60E05, 62F10, 62N05.

Keywords and phrases: COVID-19, distribution function, infectious disease, pandemic, return period, stochastic model.

*Corresponding author

Copyright © 2021 the Authors

starting in the Middle Ages. It began with the Black Death plague in the 14th century where in Europe alone were recorded some 25 million deaths in a population of about 100 million. Towns and villages were completely ravaged with the plague. Similar report was also obtained from China, where up to two-thirds of many local populations were said to have died in widely scattered settlements [1]. Closer to our time is the Spanish Flu pandemic which surfaced in the twilight of World War I in 1919 with global mortality hitting some 20 million in twelve months. Since this last pandemic, several infectious disease outbreaks have plagued our world coming with their attendant health, economic and social challenges. To name a few, the world have had to endure infectious diseases like Rabies, Lassa Fever, Smallpox, Infectious Cancers, Genital Herpes, HIV/AIDS, Ebola Virus etc. and sandwiched between all of these are diseases induced by parasite like Malaria all having their toll on the well-being of humans.

The current COVID-19 pandemic is one that has placed the entire world on a stand-still. With no known vaccine for the treatment of the disease yet, government of nations and the World Health Organization (WHO) have put in place various measures to curb the spread of the viral infection. There is no doubt that an understanding of the stochastic nature of the infection trajectory of the disease will help a great deal in strengthening the existing control framework while also offering useful clues into the long-term trend or behavior of the disease. Several mathematical models have been put forward recently to understand the dynamics of the virus (see [2-6]). The focus of this paper is to develop a stochastic model in the form of a distribution function to capture the random behavior of the daily number of infections from COVID-19 in Nigeria.

The rest of the paper is organized as follows. In Section 2 preliminary arguments for the development of a stochastic model are presented. The model and its estimation procedure are presented in Section 3. In Section 4 the application of the model to the reported daily number of COVID-19 infection cases in Nigeria is carried out. The paper closes in Section 5 with a conclusion.

2 Preliminary Arguments for a Stochastic Model

The central focus of a stochastic model is usually on a random variable X which represents a quantity whose outcome is uncertain. This quantity can be the number of infected persons, the number of deaths or the number of recoveries from a disease per unit time in a population. Suppose F is the true model which generates all the outcomes of X in the population. In this regard, F is taken to be a distribution function for calculating the probability that X will be less than or equal to some x . The model F is usually unknown. Suppose there exist some sample data X_1, \dots, X_n of size n which are independent realizations of the random variable X , the most plausible thing to do would be to use these independent realizations of X to find an approximation for F say \hat{F}_n . Observe that the approximation \hat{F}_n will be a model-free estimate of F since it is entirely based on the realized sample. In fact, if $X_{(1)}, \dots, X_{(n)}$ denote the ordered sample of the sample of independent observations from F such that $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$, for any one of the $X_{(i)}$, exactly i of the n observations will have a value less or equal to $X_{(i)}$. So, a sample-based estimate of the probability of an observation being less than or equal $X_{(i)}$ is $\hat{F}_n(X_{(i)}) = i/n$ while to avoid $\hat{F}_n(X_{(n)}) = 1$, a slight adjustment is made and thus

$$\hat{F}_n(x) = \frac{i}{n+1}, \quad X_{(i)} \leq x \leq X_{(i+1)} \tag{1}$$

is the data-based approximation of F . It follows that the sample-based quantile $X_{(i)}$ can be realized from (1) using the relation

$$X_{(i)} = \hat{F}_n^{-1}\left(\frac{i}{n+1}\right), \quad i = 1, 2, \dots, n \tag{2}$$

where \hat{F}_n^{-1} is the inverse \hat{F}_n .

Given that \hat{F}_n the sample-based estimate of F is without parameter(s), essential statistical properties of F such as shape, moment and asymptotic behaviors will be difficult to ascertain from \hat{F}_n . It is usually the practice to develop a parameterized model \hat{F} and its inverse \hat{F}^{-1} that can best be

used to approximate F and F^{-1} respectively, and for any given application to a sample, closely align with \hat{F}_n and \hat{F}_n^{-1} respectively. Several methods for generating families of parameterized distribution functions have appeared in the literature within the last three decades (see [7-12]). In the following section, a parametric distribution function is developed to model the number of reported daily infections from COVID-19 in Nigeria.

3 Model and Estimation

Here the new stochastic model for the reported daily infections from COVID-19 in Nigeria is presented and defined in form of a distribution function. The estimation of the model using the maximum likelihood method is also presented.

3.1 The model

Suppose X is a random variable representing the number of infections from COVID-19 on a given day. Let X_1, \dots, X_n be independent observations of the number of infections from day 1 to day n . In many cases the natural logarithm of these observations are taken and modeled for a number of reasons. Random variables of this nature tend to be highly skewed with the possibility of having outliers. If the skewness of the random variable is to the right such that its distribution is approximately a lognormal distribution, the distribution of the natural logarithm of the random variable will be a normal distribution and the readily available analysis and inference which holds for the normal distribution can be applied. However, there is hardly a data set in reality whether original or transformed with a perfect normal distribution. What obtains in reality are distributions which exhibit some form of asymmetry and in many cases with different tail and shape behaviors ranging from heavy-tailed, light-tailed, left-skewed, right-skewed and even bimodal shapes. Going forward, it is difficult to know which, among these characteristics of a distribution, a given data arising from random processes will assume and because of that, it will be helpful to develop a model which possesses all of these characteristics. Models of such nature

are classified as being flexible due to the fact that they can accommodate the behavior of almost all data types. Again, the need to model the logarithm of the actual data set rather than the actual data set itself, stems from the fact that the logarithmic transformation of a data set can be useful in reducing variability in the data set. Thus, Our goal is to develop an appropriate distribution function \hat{F} with a corresponding density function \hat{f} where \hat{f} is the first derivative of \hat{F} that can be used to efficiently model the natural logarithm of the number of reported daily infections from COVID-19 in Nigeria, and for a simple anti-logarithm transformation lead us to the results for the original sample. We believed that convoluting two or more distribution functions will be very useful in arriving at such a flexible model in form of a distribution functions.

Suppose T is a random variable following an extreme value distribution with distribution and density functions given respectively by

$$F_T(x) = \exp \left[-\exp \left(-\frac{x-k}{c} \right) \right],$$

$$f_T(x) = \frac{1}{c} \exp \left(-\frac{x-k}{c} \right) \exp \left[-\exp \left(-\frac{x-k}{c} \right) \right].$$

$$-\infty < x < \infty, \quad c > 0, \quad -\infty < k < \infty.$$

Let R be a Weibull random variable with distribution and density functions given respectively by

$$F_R(x) = 1 - \exp \left[-\left(\frac{x}{\lambda} \right)^\alpha \right],$$

$$f_R(x) = \frac{\alpha}{\lambda} \left(\frac{x}{\lambda} \right)^{\alpha-1} \exp \left[-\left(\frac{x}{\lambda} \right)^\alpha \right],$$

$$x > 0, \quad \alpha > 0, \quad \lambda > 0.$$

Let Y be a standard logistic random variable with quantile function

$$Q_Y(p) = \log \left(\frac{p}{1-p} \right),$$

$$0 < p < 1.$$

The distribution function

$$F(x) = \int_{-\infty}^{Q_Y(F_R(x))} f_T(t) dt = F_T(Q_Y(F_R(x))) \tag{3}$$

is a valid convoluted distribution function. If β is a positive real number, the distribution function

$$\hat{F}(x) = (F(x))^\beta = (F_T(Q_Y(F_R(x))))^\beta \tag{4}$$

is a more flexible distribution function than the one in (3) and for $\beta = 1$ reduces to the distribution function in (3). Using (4), we define a new model for the logarithm of the reported daily number of infections from COVID-19 in Nigeria expressed in terms of the distribution and density functions given respectively by

$$\hat{F}(x) = \exp \left\{ -\beta e^{k/c} \left(e^{(x/\lambda)^\alpha} - 1 \right)^{-1/c} \right\}, \tag{5}$$

$$\hat{f}(x) = \frac{\alpha \beta e^{k/c}}{\lambda c} \left(\frac{x}{\lambda} \right)^{\alpha-1} e^{(x/\lambda)^\alpha} \left(e^{(x/\lambda)^\alpha} - 1 \right)^{-1-1/c} \exp \left\{ -\beta e^{k/c} \left(e^{(x/\lambda)^\alpha} - 1 \right)^{-1/c} \right\}, \tag{6}$$

$$x > 0, \alpha > 0, \beta > 0, c > 0, \lambda > 0, -\infty < k < \infty.$$

Observe that the model in (5) has 5 parameters with the parameters α, β, c and k playing the role of shape parameters. These parameters allow the model to be flexible enough to accommodate various shape and tail behaviors arising from the data it is applied to. The parameter λ is a scale parameter. In Figures 1, 2 and 3 various shapes of the density in (6) are shown for different combinations of the parameters. The Figures clearly show that the density can be skewed to the right, skewed to the left and also exhibit bimodal shapes, a situation which points to its flexibility. Consequently, daily infection quantiles at some probability p is obtained from the relation

$$N(p) = \exp(Q(p)) = \exp \left\{ \lambda \left[\log \left(1 + \left(-\frac{e^{-k/c}}{\beta} \log(p) \right)^{-c} \right) \right]^{1/\alpha} \right\}, \tag{7}$$

$$Q(p) = \hat{F}^{-1}(p) = \inf \{ x : \hat{F}(x) \geq p \}, 0 < p < 1.$$

A surface plot of the daily infection quantiles $N(p)$ for some range of values of α and β with $p = 0.4, k = 2, c = 0.5$ and $\lambda = 1$ is shown in Figure 4. The same plot for some range of values of k and c is shown in Figure 5 for $p = 0.3, \alpha = 4$ and $\beta = 5$ and $\lambda = 3$.

3.2 Model for infection threshold

Suppose it is of interest to determine the possibility of a given day’s number of infections exceeding or equating a certain threshold $X(t)$. The quantity $X(t)$ is interpreted as the number of infections which is exceeded at least once in $t > 1$ days with probability p_t . In fact, $X(t)$ is expected to be exceeded on average once in $1/p_t$ days. Determining $X(t)$ is important for planning by the authorities given that the bed spaces for admitting infected individuals are limited and has a finite carrying capacity. It can also come handy when the interest is to know when the available isolation facilities will no longer have space for admitting newly infected individuals. Suppose that the event $X \geq X(t)$ have just occurred. That is, a given day’s infections number have just exceeded the threshold $X(t)$. The number of days it will take for it to happen again is the “recurrence interval” and the expected recurrence interval is the return period t of the event $X \geq X(t)$. Thus t is the average number of days that it will take for the event $X \geq X(t)$ to happen again. It follows that the probability p_t of the occurrence of the event $X \geq X(t)$ is related to t and is computed from the relation

$$p_t = 1 - \hat{F}(\log(X(t))) = \frac{1}{t}. \tag{8}$$

Consequently, using the proposed model, $X(t)$ for a given t can be obtained by solving for $X(t)$ in the equation

$$1 - t \left(1 - \hat{F}(\log(X(t))) \right) = 0. \tag{9}$$

Solving (9) for $X(t)$ gives

$$X(t) = \exp \left\{ \lambda \left[\log \left(1 + \left(-\frac{e^{-k/c} \log(1 - 1/t)}{\beta} \right)^{-c} \right) \right]^{1/\alpha} \right\}. \tag{10}$$

Thus (10) is used to estimate the infection threshold for any return period t . Observe that large values of t will result in large values of $X(t)$ and vice versa.

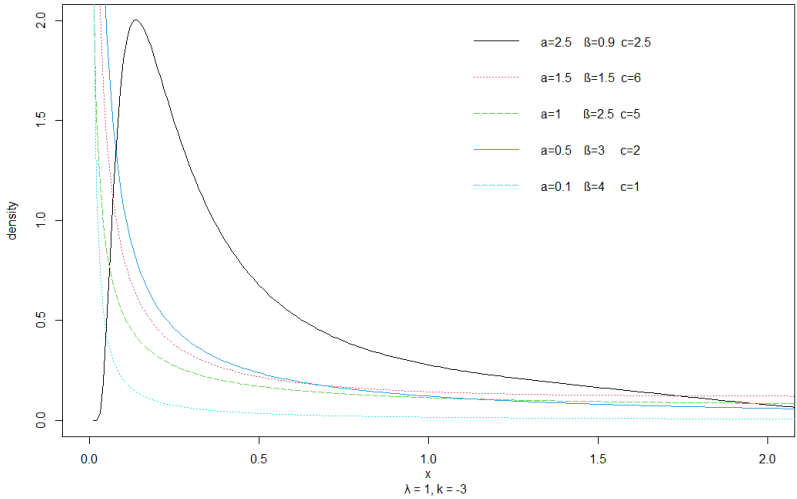


Figure 1: Model density for various parameters values.

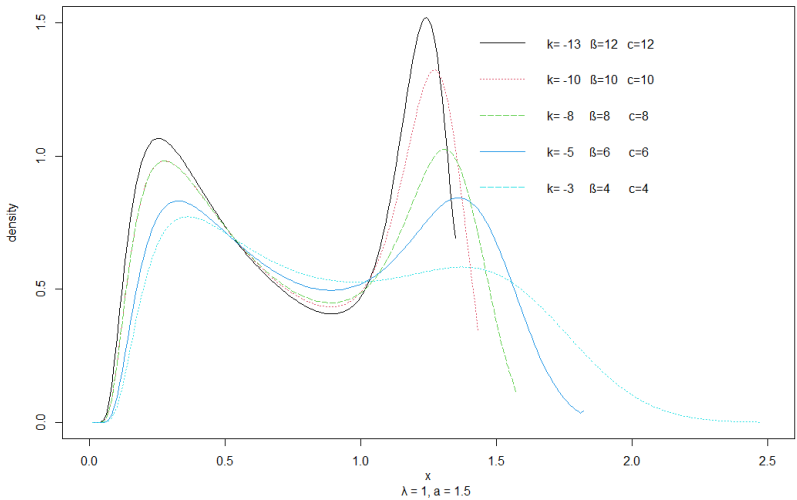


Figure 2: Model density for various parameters values.

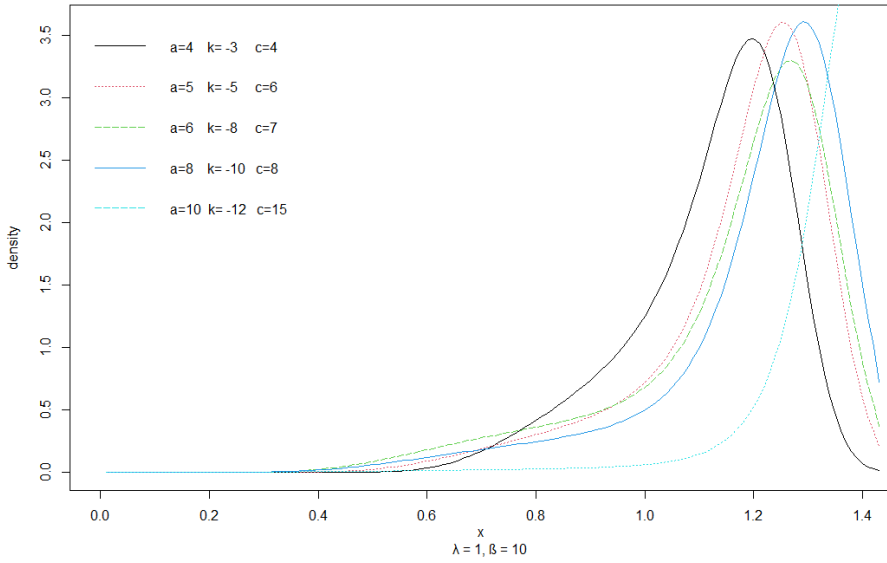


Figure 3: Model density for various parameters values.

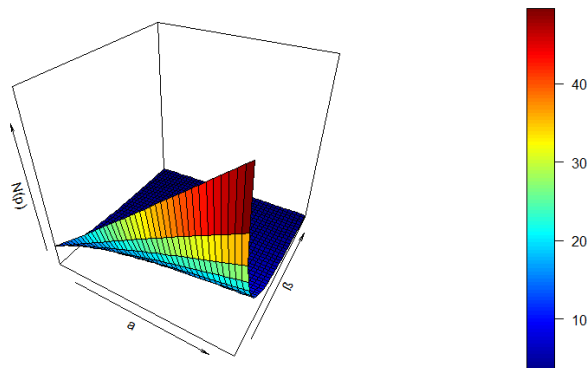


Figure 4: Surface plot of daily infection quantiles.

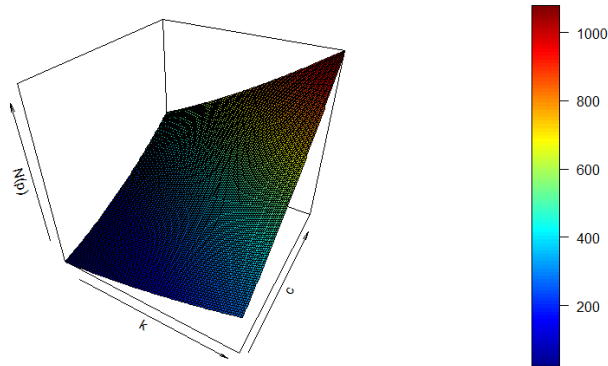


Figure 5: Surface plot of daily infection quantiles.

3.3 Model parameters estimation

For a random independent sample x_1, \dots, x_n of size n representing the logarithm of the daily number of infections, to estimate the parameters α, β, c, k , and λ of the model in (5) using the maximum likelihood method, we would have to maximize the log-likelihood function

$$L = \sum_{i=1}^n \log(\hat{f}(x_i)), \quad (11)$$

with respect to the parameters α, β, c, k , and λ . Let $\Theta = (\alpha, \beta, c, k, \lambda)^T$ be the unknown parameter vector. Then the associated score function is given by

$$\mathbf{U}(\Theta) = \left(\frac{\partial L}{\partial \alpha}, \frac{\partial L}{\partial \beta}, \frac{\partial L}{\partial c}, \frac{\partial L}{\partial k}, \frac{\partial L}{\partial \lambda} \right)^T,$$

where $\frac{\partial L}{\partial \alpha}, \frac{\partial L}{\partial \beta}, \frac{\partial L}{\partial c}, \frac{\partial L}{\partial k}$ and $\frac{\partial L}{\partial \lambda}$ are the first partial derivatives of the log-likelihood function in (11) w.r.t. each parameter. The maximum likelihood

point estimate of Θ can be obtained by solving the non-linear system of equations $\mathbf{U}(\Theta) = 0$. Since a simple analytic solution do not exist for all the equations in the system, the solutions can be obtained numerically using some specialized numerical optimization method like the Newton-type numerical methods. These numerical schemes can be implemented in software packages like R, Matlab and MATHEMATICA.

For interval estimation of the parameters, the Fisher Information Matrix (FIM) is required and it is the 5×5 symmetric matrix

$$\mathbf{I}(\Theta) = -E_{\Theta} \begin{pmatrix} I_{\alpha\alpha} & I_{\alpha\beta} & I_{\alpha c} & I_{\alpha k} & I_{\alpha\lambda} \\ I_{\beta\alpha} & I_{\beta\beta} & I_{\beta c} & I_{\beta k} & I_{\beta\lambda} \\ I_{c\alpha} & I_{c\beta} & I_{cc} & I_{ck} & I_{c\lambda} \\ I_{k\alpha} & I_{k\beta} & I_{kc} & I_{kk} & I_{k\lambda} \\ I_{\lambda\alpha} & I_{\lambda\beta} & I_{\lambda c} & I_{\lambda k} & I_{\lambda\lambda} \end{pmatrix},$$

where the elements $I_{ij}(\Theta) = \left[\frac{\partial^2 L}{\partial \Theta_i \partial \Theta_j} \right]$. The total FIM, $\mathbf{I}(\Theta)$, can be approximated by

$$\mathbf{J}(\hat{\Theta}) \approx \left[-\frac{\partial^2 L}{\partial \Theta_i \partial \Theta_j} \Big|_{\Theta=\hat{\Theta}} \right]_{5 \times 5}.$$

For real data, $\mathbf{J}(\hat{\Theta})$ is obtained after the maximum likelihood estimate of Θ is obtained at the point of convergence of the iterative numerical scheme involved in the estimation process.

Suppose $\hat{\Theta}$ is the maximum likelihood estimate of Θ . Under the usual regularity conditions and that the parameters are in the interior of the parameter space, but not on the boundary, we have: $\sqrt{n}(\hat{\Theta} - \Theta) \xrightarrow{d} N_5(\mathbf{0}, \mathbf{I}^{-1}(\Theta))$, where $\mathbf{I}^{-1}(\Theta)$ is the inverse of the expected FIM, which is also the variance-covariance matrix of the parameters. The asymptotic behavior is still valid if $\mathbf{I}^{-1}(\Theta)$ is replaced by the inverse of the observed information matrix evaluated at $\hat{\Theta}$, that is $\mathbf{J}(\hat{\Theta})$. The multivariate normal distribution with mean vector $\mathbf{0} = (0 \ 0 \ 0 \ 0 \ 0)^T$ and variance-covariance matrix $\mathbf{I}^{-1}(\Theta)$ can be used to

construct confidence intervals for the parameters. The approximate $100(1 - \theta)\%$ two-sided confidence interval for the estimates of the parameters α, β, c, k , and λ are given by

$$\hat{\alpha} \pm Z_{\theta/2} \sqrt{\mathbf{I}_{\alpha\alpha}^{-1}(\hat{\theta})}, \quad \hat{\beta} \pm Z_{\theta/2} \sqrt{\mathbf{I}_{\beta\beta}^{-1}(\hat{\theta})}, \quad \hat{c} \pm Z_{\theta/2} \sqrt{\mathbf{I}_{cc}^{-1}(\hat{\theta})},$$

$$\hat{k} \pm Z_{\theta/2} \sqrt{\mathbf{I}_{kk}^{-1}(\hat{\theta})}, \quad \hat{\lambda} \pm Z_{\theta/2} \sqrt{\mathbf{I}_{\lambda\lambda}^{-1}(\hat{\theta})},$$

respectively, where $\mathbf{I}_{\alpha\alpha}^{-1}(\hat{\theta}), \mathbf{I}_{\beta\beta}^{-1}(\hat{\theta}), \mathbf{I}_{cc}^{-1}(\hat{\theta}), \mathbf{I}_{kk}^{-1}(\hat{\theta})$ and $\mathbf{I}_{\lambda\lambda}^{-1}(\hat{\theta})$ are diagonal elements of $\mathbf{I}^{-1}(\hat{\theta})$ and $Z_{\theta/2}$ is the upper $(\theta/2)^{th}$ percentile of a standard normal distribution.

4 Model Application

Application of the model to the reported daily number of infections from COVID-19 in Nigeria is carried out in this section. The original data set was obtained from Worldometer (<https://www.worldometers.info/coronavirus/country/nigeria/>).

4.1 Analysis and results

The data set which is contained in Table 1 is for the period March 18th - June 11th, 2020. The mean, standard deviation, skewness and excess kurtosis of the original data set are 170.4405, 159.0805, 0.9139 and 0.5417 respectively. Observe that the positive value of skewness of the data implies that the data set is skewed to the right and the excess kurtosis value being positive implies that the distribution is peaked and possesses thick tails. A logarithmic transformation of the data set yields a mean value of 4.3655, a standard deviation value of 1.5337, a skewness value of -0.5780 and an excess kurtosis value of -1.0859. Observe that the log transformed data set possesses less variability as evident in its standard deviation value in comparison with that of the original data set. The skewness of the log transformed data set is to the left and the kurtosis value being negative implies that the distribution of the log-transformed data set is flat with thin tails. Scatter

plots and boxplots of the original and the log transformed data are shown in Figure 6(a-d). The histogram and the kernel density of both the original and the log transformed data sets are shown in Figure 7(a-d). The empirical cumulative probability and empirical quantile plots using (1) and (2) for both the original and log transformed data are shown in Figure 8(a-d).

Table 1: Reported daily infections from COVID-19 in Nigeria
March 18th - June 11th, 2020

5	4	10	8	10	4	7	14	5	28	14	20	4	39	10
26	4	18	6	16	22	12	17	13	5	20	30	34	35	51
49	85	38	117	91	108	114	87	91	64	195	196	204	238	218
170	244	148	195	381	386	239	248	146	184	193	288	176	338	216
226	276	339	245	265	313	229	276	389	182	387	553	307	416	241
347	350	328	389	253	315	663	409	681						

Using the model in (5) to fit the log transformed data and by employing the use of the R software package gives the following maximum likelihood estimate for the parameters (standard error of estimates in parenthesis): $\hat{\alpha} = 8.6624(0.0063)$, $\hat{\beta} = 1.0987(1.2520)$, $\hat{c} = 7.5255(0.6474)$, $\hat{k} = -0.8845(8.6022)$ and $\hat{\lambda} = 4.3155(0.0030)$. The estimate of the log-likelihood value is -130.6145, the value of the model's Kolmogorov - Smirnov (K-S) statistic is 0.0846 with a corresponding p-value of 0.5566. The theoretical cumulative probabilities based on the model alongside the data- or sample based counterpart, the Q-Q plot, theoretical density plot based on the model over the histogram of the data and the P-P plot are shown in Figure 9(a-d). The inverse of the observed information matrix which also corresponds to the variance-covariance matrix of the estimate of the parameters is given in (12). In Table 2, 95% confidence interval for the parameter estimates are presented. Some daily infection thresholds for some given return periods are computed and presented in Table 3.

$$\mathbf{J}^{-1}(\hat{\Theta}) = \begin{pmatrix} 3.91e-05 & 1.06e-05 & 7.28e-05 & -5.9e-05 & 2.59e-08 \\ 1.06e-05 & 1.567490 & 0.061173 & -10.0717 & -1.2e-05 \\ 7.28e-05 & 0.061173 & 0.419081 & -0.28717 & -8.1e-05 \\ -5.9e-05 & -10.0717 & -0.28717 & 70.3998 & 1.54e-05 \\ 2.59e-08 & -1.2e-05 & -8.1e-05 & 1.54e-05 & 9.28e-06 \end{pmatrix}. \tag{12}$$

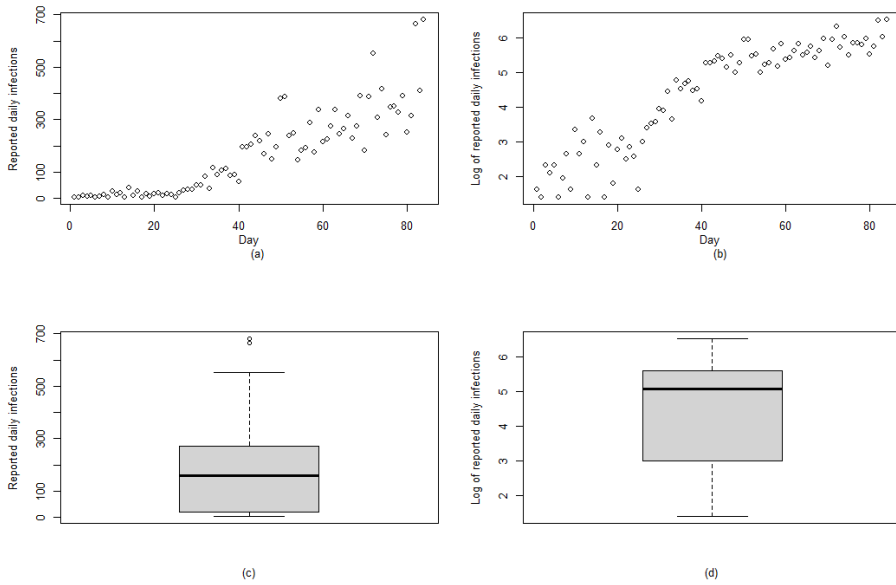


Figure 6(a-d): Scatter plots and boxplots of the dataset.

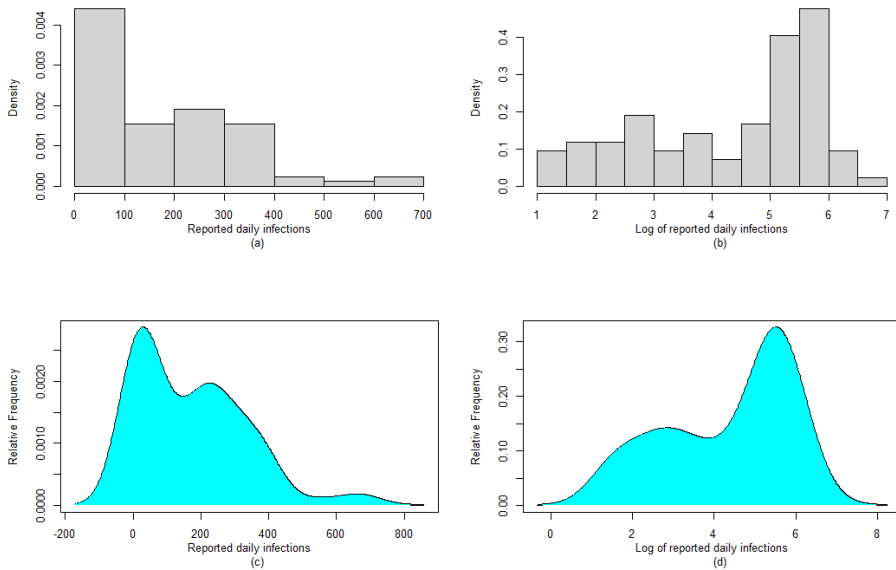


Figure 7(a-d): Histogram and kernel density of the dataset.

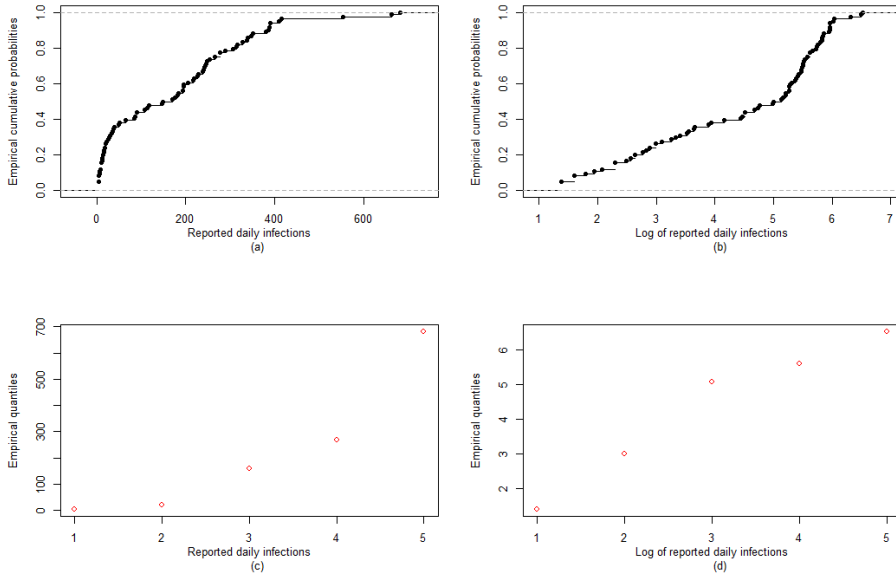


Figure 8(a-d): Empirical cumulative probabilities and quantile plots of the dataset.

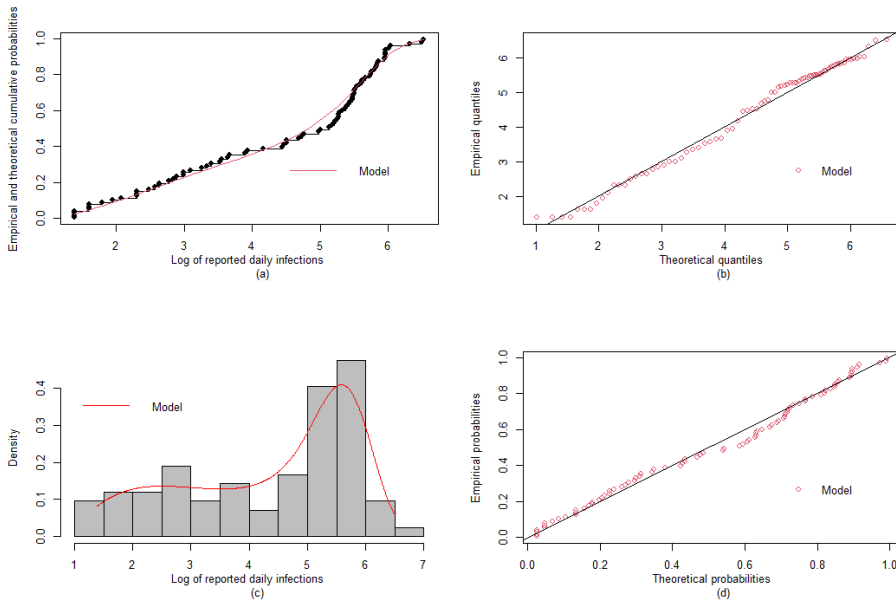


Figure 9(a-d): Summary plots of the fitted model.

4.2 Discussion

Results from the analysis clearly support the adequacy of the model in fitting the logarithm of the daily number of reported infections. This is evident in the fact that the p-value of the K-S statistic being 0.5566 is significantly greater than the nominal 0.05. Thus, we are 95% certain that the proposed model is not different from the true process the logarithm of the infections number is following. Put differently, we can confidently used (5) and (6) to determine the cumulative and exact probabilities respectively of observing the logarithm of a certain number of infections on a given day. This will no doubt be very helpful for planning by the Task Force Team saddled with the responsibility of managing the pandemic and indeed all health practitioners. To further buttress the adequacy of the model in fitting the data set, it can be observed in Figure 9 panel (a) that the empirical cumulative probabilities based on the data set closely aligns with the theoretical ones based on the model. Also, in panel (b) of Figure 9, it can also be observed that the empirical quantiles (the 45⁰ line) based on the data is almost the same as the theoretical ones generated by the model. The same holds true in panel (d) of Figure 9 where the empirical probabilities based on the data set closely aligns with the theoretical ones based on the model. In panel (c) of Figure 9, it can be observed that the fitted density based on the model adequately approximates the histogram of the data set while also closely aligning with the kernel density of the data in Figure 7 panel (d).

Table 2: Parameters and their 95% confidence interval

paramters	Estimates	95% confidence interval
α	8.6624	(8.6501 , 8.6747)
β	1.0987	(-1.3552 , 3.5526)
c	7.5255	(6.2566 , 8.7944)
k	-0.8845	(-17.7448 , 15.9758)
λ	4.3155	(4.3096 , 4.3214)

Table 3: Daily infection thresholds

t	2	5	7	10	14	21	30	40	50	60	70	80	90	100
$X(t)$	125	298	346	394	436	484	526	559	584	605	622	636	649	661

In Table 2, it can be observed that the maximum likelihood method of parameter estimation provided very good estimates for the model’s parameters. This is clearly shown by the short width of the confidence interval for each of the parameter estimates. In Table 3, some daily infection thresholds $X(t)$ are computed and shown for some given return period t . For example, for $t = 7$, the infection threshold is 346. This means that on average 346 number of reported cases will be recorded or exceeded at least once in 7 days. We also observed that as the return period t increases, the daily infection threshold increases also, but increases more sharply at lower return periods than at higher ones while for t within the neighborhood of 100 days, the daily infection threshold does not exceed 800 and this has been the reality of the pandemic in Nigeria since the first case was recorded with the highest reported daily infections number not exceeding 800.

5 Conclusion

Stochastic quantification of several real life problems have been very helpful in understanding the random nature of their incidence or occurrence. This has further helped in finding solutions to the problems arising from them either in form of minimization of their undesirability or maximizing their rewards. A pandemic is usually a situation of untold inconvenience and as such, all hands must be on deck to help bring it to a halt. Finding appropriate models that can help explain one or more of its effects will no doubt help in the provision of solutions. In this paper that has been carried out using Nigeria as a case study. This article is by no means an exhaustive exposition on the scheme of things but a contribution in an analytical and in a numerical way for the better understanding of the probabilistic trajectory of the pandemic in Nigeria.

References

- [1] N.T.J. Bailey, The role of statistics in controlling and eradicating infectious diseases, *The Statistician* 34 (1985), 3-17. <https://doi.org/10.2307/2987500>
- [2] Z. Liu, P. Magal, O. Seydi and G. Webb, Predicting the cumulative number of cases for the COVID-19 epidemic in China from early data, 2020. arXiv:2002.12298v1 <https://doi.org/10.20944/preprints202002.0365.v1>
- [3] K. Roosa, Y. Lee, R. Luo, A. Kirpich, R. Rothenberg, J.M. Hyman, P. Yan and G. Chowell, Real-time forecasts of the COVID-19 epidemic in China from February 5th to February 24th, 2020, *Infectious Disease Modelling* 5 (2020), 256-263. <https://doi.org/10.1016/j.idm.2020.02.002>
- [4] B. Tang, N.L. Bragazzi, Q. Li, S. Tang, Y. Xiao and J. Wu, An updated estimation of the risk of transmission of the novel corona virus (2019-nCov), *Infectious Disease Modelling* 5 (2020a), 248-255. <https://doi.org/10.1016/j.idm.2020.02.001>
- [5] B. Tang, X. Wang, Q. Li, N.L. Bragazzi, S. Tang, Y. Xiao and J. Wu, Estimation of the transmission risk of the 2019-nCov and its implication for public health intervention, *Journal of Clinical Medicine* 9(2) (2020b), 462. <https://doi.org/10.2139/ssrn.3525558>
- [6] J.T. Wu, K. Leung and G.M. Leung, Nowcasting and forecasting the potential domestic and international spread of the 2019-nCov outbreak originating in Wuhan, China: a modelling study, *The Lancet* 395(10225) (2020), 689-697. [https://doi.org/10.1016/S0140-6736\(20\)30260-9](https://doi.org/10.1016/S0140-6736(20)30260-9)
- [7] N. Eugene, C. Lee and F. Famoye, Beta-normal distribution and its applications, *Communications in Statistics - Theory and Methods* 31 (2002), 497-512. <https://doi.org/10.1081/STA-120003130>
- [8] M.C. Jones, Kumaraswamy's distribution: A beta-type distribution with tractability advantages, *Statistical Methodology* 6 (2009), 70-81. <https://doi.org/10.1016/j.stamet.2008.04.001>
- [9] G.M. Cordeiro and M. de Castro, A new family of generalized distributions, *Journal of Statistical Computation and Simulation* 81 (2011), 883-898. <https://doi.org/10.1080/00949650903530745>

-
- [10] A. Alzaatreh, C. Lee and F. Famoye, T - normal family of distributions: a new approach to generalize the normal distribution, *Journal of Statistical Distributions and Applications* 1 (2014), 16. <https://doi.org/10.1186/2195-5832-1-16>
- [11] P. Osatohanmwen, F.O. Oyegue, F. Ewere and B. Ajibade, A new family of generalized distributions on the unit interval: the T - Kumaraswamy family of distributions, *Journal of Data Science* 18(2) (2020), 218-236.
- [12] P. Osatohanmwen, F.O. Oyegue and S.M. Ogbonmwan, The $T - R\{Y\}$ power series family of probability distributions, *Journal of the Egyptian Mathematical Society* 28 (2020), 29. <https://doi.org/10.1186/s42787-020-00083-7>

This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted, use, distribution and reproduction in any medium, or format for any purpose, even commercially provided the work is properly cited.
