American Journal of Applied Mathematics and Statistics, 2023, Vol. 11, No. 1, 13-21
Available online at http://pubs.sciepub.com/ajams/11/1/3
Published by Science and Education Publishing
DOI:10.12691/ajams-11-1-3

Log Normal Distribution Approach for Forecasting the Spread of Covid-19 in Nigeria

Fawei Tonbra Erepamo Lawrence[1], Orumie Ukamaka Cynthia[2,*], Biu Emmanuel Oyinebifun[2]

[1]Department of Mathematics/Statistics, Ignatius Ajuru University of Education, Rivers State, Nigeria
[2]Department of Mathematics and Statistics, University of Port Harcourt, Choba, Rivers State, Nigeria
*Corresponding author: amakaorumie5@gmail.com

Received February 02, 2023; Revised March 04, 2023; Accepted March 10, 2023

Abstract The Nigeria National Weekly Covid-19 confirmed cases were used in this study work to anticipate the predicted number of cases by Lagos and Abuja areas using the log normal distribution technique. Using trend analysis, this study was able to determine the trend component in both Lagos and Abuja Covid-19 confirmed cases (i.e. Quadratic Trend). According to the findings, the number of confirmed cases of covid-19 increased at the start of the pandemic before starting to decline slightly as the number of cases increased in both areas on the Nigeria National Weekly Covid-19 report each week. The distribution of Covid-19 confirmed cases in both locations was then modelled using the quadratic trend, which was found to be appropriate. For simple examination of the expected (average) distribution and variation of both FCT and Lagos on the Nigeria National Weekly Covid-19 confirmed cases, log-normal distribution techniques were converted to a quadratic form. The results showed that the Nigeria National Weekly Covid-19 confirmed cases in the Federal Capital Territory (FCT) had an expected value of 27 cases per week, while the Lagos State Covid-19 confirmed cases had an expected value of 19 cases per week. Additionally, this study established a mathematical relationship between the Nigeria National Weekly Confirmed cases and the FCT and Lagos COVID-19 confirmed cases. This was accomplished by constructing their mathematical functions for forecasting purposes: probability density function and cumulative density function. It was also able to predict the anticipated pandemic cases of both FCT and Lagos on the Nigeria National Weekly Covid-19 cases and established a quadratic trend for both FCT and Lagos on the Nigeria National Weekly Covid-19 confirmed cases (NWC).

Keywords: *log normal distribution, trend analysis, expected number of cases, Nigeria National Weekly Covid-19 confirmed cases*

Cite This Article: Fawei Tonbra Erepamo Lawrence, Orumie Ukamaka Cynthia, and Biu Emmanuel Oyinebifun, "Log Normal Distribution Approach for Forecasting the Spread of Covid-19 in Nigeria." *American Journal of Applied Mathematics and Statistics*, vol. 11, no. 1 (2023): 13-21. doi: 10.12691/ajams-11-1-3.

1. Introduction

The distribution of a random variable whose logarithm is normally distributed is known as the log normal distribution, and it is typically formulated with two parameters, mean and variance, or and. The log normal distribution has been widely used in economics, finance, and risk analysis as well as biology, ecology, geology, and meteorology, among other areas of life science. In cosmology and astrophysics, it is significant.

In survival analysis, the analysis of the time until the occurrence of any event is a key topic.

A common statistical technique, linear regression has been utilized successfully in numerous fields, including survival analysis. Because it offers a more straightforward interpretation than the cox regression proportional hazards model, a log-transformation of the response variable in survival analysis transforms a traditional linear model into an accelerated failure time model [1].

The "Censored" data in which the time to event cannot be totally seen but instead indicates the lower bound of the time to occurrence is a common feature of survival analysis. There are numerous statistical techniques available for estimating survival functions, comparing survival curves between groups, and modeling survival data using regression to examine associations with risk factors. Non-parametric statistical inference is increasingly frequently employed in survival analysis to estimate the survival function, compare the survival function, and compare the survival curves between two or more groups.

The employment of a log-normal regression model is a well-known and often used technique Haipeng and Zhengyuan [2].

In probability theory, a log-normal distribution is a continuous probability distribution of a random variable whose logarithm is normally distributed. These distributions are self-replicating when multiplied and divided. It is helpful for modeling skewed data with a small mean value and high variance. According to

Stahel et al., the log-normal distribution is the life distribution model that is utilized the most frequently for technological applications [3].

Numerous studies have been conducted on the log-normal distribution, log-normal regression, and distribution of COVID-19 in Nigeria and across the globe. The Extended Log normal distribution; characteristics and application was the focus of Bachioua et al [4]'s research. The approximation of a Log normal sum (in the complete support) and the approximate Laplace transformation of the Log normal sum via transform inversion were two closely connected issues that Asunussen et al. [5] examined in relation to the Log normal distribution. Christopher and Mitchell [6] found that log-normal random variables naturally occur in several engineering areas, such as finance, reliability theory, and wireless communications. In their paper, Akbar et al. [7] provide a brand-new test based on a computational approach to examine the equality of several log-normal means. Using log normal distribution, Orumie and Biu [8] calculated the time to failure rate and the number of successful transactions of chosen banks in Nigeria. A few parametric survival distributions with applications in data on loan default and tuberculosis were compared by Orumie et al. in [9]. Aronu [10] looked at the early epidemiological analysis of the first 45 days of the COVID-19 epidemic in Nigeria while Guillermo et al [11] worked on a survival analysis of COVID-19 in the Mexican community. The survival scenario of COVID-19 patients in India was examined by Kundu et al. in 2020 using a survival analysis of patient demographics. In their work titled "Online Forecasting of covid-19 cases in Nigeria using restricted data," Abdulmajeed et al. [12] examined the scope of the disease's distribution and the efficacy of containment measures to stop the spread of the illness. Research on the Geographical Trend Analysis of the Covid-19 Pandemic's Inception in Africa was piloted by Onafeso et al. in [13]. Forecasting the spread of COVID-19 in Nigeria using the Box-Jenkins Modeling Procedure was the focus of Rauf and Oladipo's work in [14]. In Nigeria, the coronavirus (Covid-19) was investigated by Aronu, et al. in [10]. Log-normal distributions were the focus of Limpert et al [16]'s study on science's key and hints.

In order to demonstrate the trend component in both cities' covid-19 confirmed cases, this research aims to create predictability distribution models for the region between Lagos and Abuja utilizing trend analysis and the National Weekly Confirmed Cases (NWC) of Nigeria. The goals are to calculate the trend effect between cases confirmed in Abuja/Lagos and those confirmed by the Nigeria National Weekly, and the objectives are;

To obtain a log normal distribution model and estimate the parameters between Abuja/Lagos and Nigeria National Weekly Confirmed.

To obtain the probability density function and cumulative density function for the purpose of prediction.

And using log normal distribution models, to forecast the anticipated COVID-19 pandemic based on the National Weekly Confirmed Cases.

2. Materials and Methods

2.1. Scope and Delimitation

The National Weekly Confirmed Cases (NWC) on Lagos and the Federal Capital Territory are the main subjects of this study's trend analysis and log normal distribution technique (Abuja). The components of the states with documented laboratory-confirmed COVID-19 cases are intended to be used. The data that the researcher pulled from the NCDC website between March 16, 2020, and April 11, 2021, constitutes the sole subject of this study. The information includes cases reported in Nigeria as a whole as well as in two of its most populous regions, where economic and political activity is intense.

2.2. Nature and Source of Data

From 2020 to 2021, two data sets were obtained from the NCDC (National Center for Disease Control) website. Covid-19 cases between the Federal Capital Territory and Lagos according to the National Weekly Confirmed cases (NWC) (Abuja). The secondary data used in this study was taken from the National Centre for Disease Control (NCDC) website and pertains to the daily/weekly confirmed cases of COVID-19 in Nigeria. cases that were reported and confirmed on a daily and weekly basis between March 16, 2020, and May 9, 2021. The analyses were conducted using Minitab and Gretl statistical software.

2.3. Trend Analysis

Technical analysis methods like trend analysis aim to forecast future movements using trend data that has just been observed.

1. The foundation of trend analysis is the notion that what has occurred in the past can predict what will occur in the future.

2. Short-, intermediate-, and long-term time frames are the main emphasis of trend analysis.

Trend analysis is a type of comparative analysis that entails examining present trends in order to forecast future ones.

However, the trend analysis in this study will provide a general notion of what has occurred in the past and what is likely to occur in the future for the Covid-19 cases.

2.3.1. Quadratic Trend Model

One of the models utilized in regression studies is the quadratic trend. Due to its complexity, this kind of regression is typically unable to be properly conveyed using a linear trend. Take into account the regression equation below, which is an illustration of a quadratic trend model:

$$Y_i = B_1 + B_2 X_{2i} + B_3 X_{3i}^2 + u_i$$

2.3.2. Log Normal Distribution

A continuous probability distribution of a random variable whose logarithm is normally distributed is known in probability theory as a log-normal (or lognormal) distribution. Therefore, Y = ln(X) has a normal distribution if the random variable X is log-normally distributed.

The methodologies for parametric survival analysis for log-normal distributions with several covariates are covered in this article. Here, the Covid-19 model comparison on national weekly confirmed cases in Lagos and the Federal Capital Territory (Abuja) using Log-Normal Distribution Approach is applied.

2.4. The Model

A continuous probability distribution function whose logarithm is normally distributed is known as a log-normal distribution. Therefore, z = ln(t) has a normal distribution if the random variable t has a log-normal distribution. The exponential function of z, t = exp(z), also has a log-normal distribution if z has a normal distribution. A log-normally distributed random variable only accepts positive real values.

A positive random variable *t* is log-normally distributed if the logarithm of *t* is normally distributed, and the pdf is thus;

$$f(t) = f(t, \square, \delta)$$
$$= \frac{1}{t} \cdot \frac{1}{\delta\sqrt{2\pi}} \exp\left(-\frac{(lnt - \mu)^2}{2\delta^2} \right), t \le 0, t < 0 \quad (1)$$

where: μ and σ is mean and variance respectively.

2.4.1. The Log Normal Distribution

Let $x_1, x_2, \dots, x_n$ be independent positive random variable such that

$$T_n = \prod_{i=1}^{n} x_i \quad (2)$$

Then the log of their product is equivalent to the sum of their logs

$$In\ T_n = \sum_{i=1}^{n} In(x_i) \quad (3)$$

The log-normal distribution has two parameters which are, μ and σ.

$$S(t) = 1 - \Phi\left(\frac{\log t - \mu}{\sigma}\right) \quad (4)$$

$$h(t) = \frac{\phi\left(\frac{\log t}{\sigma}\right)}{\left[1 - \Phi\left(\frac{\log t}{\sigma}\right)\sigma t\right]} \quad (5)$$

where $\phi(x)$ is the PDF and $\Phi(x)$ is the CDF

The Log-normal distributions have a constant variance, are normally distributed, and have a zero mean.

Consequently, the distribution for x is known as a log-normal distribution and the pdf is given as; if Z = log(x) is normally distributed.

$$f(x) = \frac{1}{t\sigma\sqrt{2\pi}} \exp \frac{-\left[In(t) - \mu\right]^2}{2\sigma^2}; \quad (6)$$
$$\sigma^2 \mu \ \epsilon (-\infty, \infty) > 0, t \in (0, \infty)$$

where x = t.

$$Mean = \exp^{\left(\mu + \sigma^2/2\right)}$$

$$Variance = \exp\left[\left(\sigma^2\right) - 1\right] \exp^{\left[2\mu + \sigma^2\right]}$$

The cumulative density function (CDF) is given as

$$CDF = \frac{1}{2} + \frac{1}{2} \exp^{\frac{\left[In(t) - \mu\right]}{\sqrt{2\delta}}} \quad (7)$$

Let equation (4) be written term of t, then

$$f(t) = \frac{1}{t\sigma\sqrt{2\pi}} \exp^{\frac{-(In(t) - \mu)^2}{2\delta^2}}.$$

Taking natural logarithm to base e on both sides

$$In\ f(t) = In\left(t\delta\sqrt{2\pi}\right)^{-1} - \frac{(Int - \mu)^2}{2\delta^2}$$

$$In\ f(t) = -In\left(t\delta\sqrt{2\pi}\right) - \frac{(Int - \mu)^2}{2\delta^2}$$

$$In\ f(t) = -In(t) - In\left(\delta\sqrt{2\pi}\right) - \frac{1}{2\delta^2}\left[(Int)^2 - 2\mu In(t) + \mu^2\right]$$

$$Inf(t) = -In(t) - In\left(\delta\sqrt{2\pi}\right) - \frac{(Int)^2}{2\delta^2} + \frac{\mu}{\delta^2} + \frac{\mu}{\delta^2} In(t) - \frac{\mu^2}{2\delta^2}$$

Collect like terms

$$Inf(t) = -\left(\delta\sqrt{2\pi}\right) - \frac{\mu^2}{2\delta^2} + \frac{\mu}{\delta^2} In(t) - In(t) - \frac{(Int)^2}{2\delta^2}$$
$$= -\left[In\left(\delta\sqrt{2\pi}\right) + \frac{\mu^2}{2\delta^2}\right] + \left(\frac{\mu}{\delta^2} - 1\right) In(t) - \frac{1}{2\delta^2}(Int)^2 \quad (8)$$

We have a quadratic function in Equation (8), which can be expressed as

$$Inf(t) = \beta_0 + \beta_1 In(t) + \beta_2 (In(t))^2 \quad (9)$$

$$y = Inf(t)$$

$$x = In(t)\ and\ x^2 = \left[In(t)\right]^2$$

Then, by comparing equation (9) with (8), obtain that

$$\beta_0 = -\left[In\left(\delta\sqrt{2\pi}\right) + \frac{\mu^2}{2\delta^2}\right]$$

$$\beta_1 = \left(\frac{\mu}{\delta^2} - 1\right)$$

$$\beta_2 = -\frac{1}{2\delta^2}$$

Equation (9) represents a curved or quadratic regression model. Natural logarithms have a normal distribution, and the log-Normal distribution is a continuous distribution of those numbers.

For instance, x = log(y) has a normal distribution if the random variable y = exp(x) has a log-normal distribution. The most frequently used log-normal distribution in finance is to simulate stock price, index values, asset returns, exchange rates derivatives, etc.

Log-Normal Distribution Statistics

Notation	In	$N(\mu, \delta^2)$
		$-\infty < \mu < \infty$
Parameter		$\delta^2 > 0$
		$x > 0$

Distributions

PDF

$$\frac{1}{\sqrt{2\pi\delta x}}\exp\left[-\frac{\left(In(x)-\mu\right)^2}{2\delta^2}\right] \tag{10}$$

CDF

$$\frac{1}{2}\left[1 + erf\left(\frac{In(x-\mu)}{\delta}\right)\right] \tag{11}$$

Mean

$$e\left(\mu + \tfrac{1}{2}\delta^2\right) \tag{12}$$

Variance

$$\left(e^{\delta^2} - 1\right)e^{2\mu+\delta^2} \tag{13}$$

Skewness

$$\left(e^{\delta^2} + 2\right)\sqrt{\left(e^{\delta^2} - 1\right)} \tag{14}$$

Kurtosis

$$e^{4\delta^2} + 2e^{3\delta^2} + 23e^{2\delta^2} - 6 \tag{15}$$

where erf is the Gaussian error function.

However, this study applied log-normal distribution approach.

2.4. Models Accuracy Measures

Many selection criteria, such as Mean Absolute Percentage Deviation (MAPD), Mean Absolute Deviation (MAD), and Mean Squared Deviation (MSD), have been given in recent years (MSD).

2.4.1. Mean Absolute Percentage Deviation (MAPD)

The accuracy of fitted trend curves for the propagation of the Covid-19 pandemic in Nigeria is measured by mean absolute percentage deviation. A better match is indicated by the smaller MAPD values.

$$MAPD = \frac{\Sigma\left|(Y_t - \hat{Y}_t)/Y_t\right|}{n} X100\ (Y_t \neq 0) \tag{16}$$

where $\hat{Y}_t$ is the fitted values of the trend curves considered, Y_t is the actual national weekly confirmed cases series values at time t and n is the number of observations.

2.4.2. Mean Absolute Deviation (MAD)

The average distance between the values of each fitted trend curve and the actual series is known as mean absolute deviation. A lower MAD is more consistent.

$$MAD = \frac{\sum_{t=1}^{n}\left|(Y_t - \hat{Y}_t)\right|}{n} \tag{17}$$

where $\hat{Y}_t$, Y_t and n as defined in Equation (16).

2.4.3. Mean Squared Deviation (MSD)

The square average distance between each fitted trend curve's values and the actual series is calculated using mean square deviation. The smaller MSD values suggest a better fit, and as a result, MSD is more sensitive than MAD to the most accurate forecast.

$$MSD = \frac{\sum_{t=1}^{n}(Y_t - \hat{Y}_t)^2}{n} \tag{18}$$

where $\hat{Y}_t$, Y_t and n as defined in Equation (16).

3. Results and Discussion

3.1. Descriptive Statistics

The descriptive statistics of the Covid-19 Pandemic in the Federal Capital Territory (Abuja) and Lagos National Weekly Confirmed cases (NNWC).

Table 1 shows the mean of the Covid-19 Pandemic in Nigeria, where the expected value is 980 case for Lagos area, 335 for FCT and 2792 for National Weekly Confirmed cases.

Table 1. Descriptive Statistics of Covid-19 Pandemic in the Nigeria

Variable	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum	Skewness	Kurtosis
LAGOS(x)	980	142	1099	57	297	623	1281	4780	2.27	5.01
FCT (w)	335	51.4	398.5	7	71.3	148.5	449	1727	1.76	2.42
NNWC (y)	2792	353	2738	81	976	1703	3956	11179	1.55	1.89

3.2. Trend Analysis

Table 2 and the figure below show the patterns of the Covid-19 Pandemic in Nigeria's two most populous regions compared to the number of confirmed cases reported in the National Weekly; (Figure 1 and Figure 2). The fitted trends have their R-squares and are linear, quadratic, cubic, and quartic (or polynomial of degree four).

Table 2. Trend Analysis Result for Both States on NNWC

Coefficients	Trend Analysis of NWC (y) against FCT (w)				Trend Analysis of NWC (y) against Lagos (x)			
	Linear	Quadratic	Cubic	Quartic	Linear	Quadratic	Cubic	Quartic
Constant (β_0)	3.396 (0.001**)	**2.466 ((0.000)****	1.28 (0.461)	-4.79 (0.339)	0.931 (0.002)**	**-2.16 (0.101)**	-6.56 (0.265)	-77.7 (0.022**)
β_1	0.7897 (0.000**)	**1.191 (0.000)****	2.05 (0.089)	8.21 (0.098)*	1.0176 (0.000)**	**2.013 (0.000)****	4.20 (0.147)	52.5 (0.021)**
β_2		**-4.04×10^{-2} (0.000)****	-0.233 (0.373)	-2.41 (0.160)		**-/7.81×10^{-2}(0.000)****	-0.433 (0.350)	-12.46 (0.027)
β_3			1.34×10^{-2} (0.458)	0.335 (0.182)			0.0187 0.442)	1.332 (0.030)*
β_4				-0.017 (0.199)				-0.0519 (0.032**)
R^2	85.7%	**86.3%**	86.5%	86.8%	89.9%	**90.9%**	91.0%	91.7%
$\bar{R}^2$	85.5%	**85.8%**	85.7%	85.9%	89.8%	**90.6%**	90.5%	91.1%
Analysis of Variance	Source P-Value Regression 0.000** FCT (w) 0.000**	Source P-Value Regression 0.000** FCT (w) 0.000** w^2 0.128	Source P-Value Regression 0.000** FCT (w) 0.089* w^2 0.373 w^3 0.459	Source P-Value Regression 0.000** FCT (w) 0.020** w^2 0.356 w^3 0.331 w^4 0.235	Source P-Value Regression 0.000** LAGOS x 0.000**	Source P-Value Regression 0.000** LAGOS x 0.000** x^2 0.000**	Source P-Value Regression 0.000** LAGOS x 0.000** x^2 0.977 x^3 0.503	Source P-Value Regression 0.000** LAGOS x 0.011** x^2 0.925 x^3 0.996 x^4 0.911

Footnote: **= p-values less than 5%.

The identified trend model are

$$Y_t = C_0 + C_1 W_t + C_2 W_t^2 = 2.466 + 1.191 W_t - (4.04\times10^{-2}) W_t^2 \quad (19)$$

$$Y_t = C_0 + C_1 X_t + C_2 X_t^2 = -2.16 + 2.013 X_t - (7.91\times10^{-2}) X_t^2 \quad (20)$$

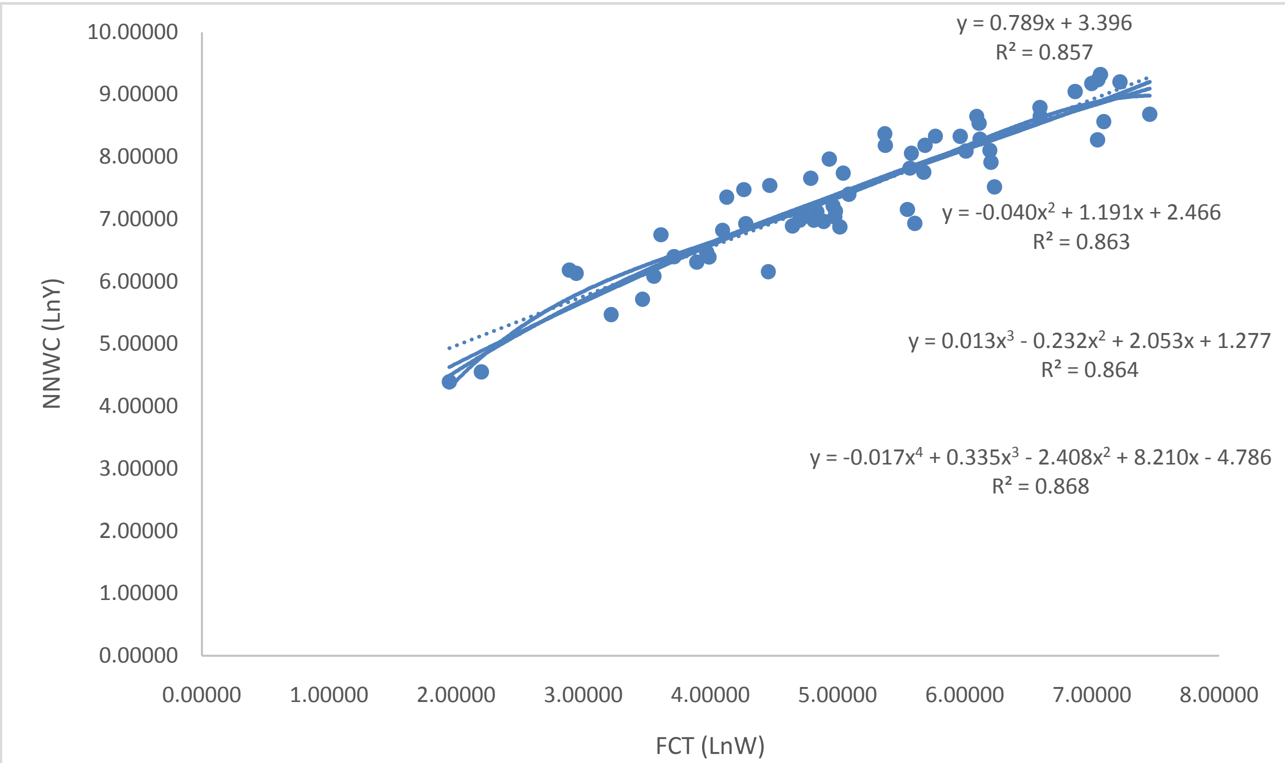

Figure 1. Scattered Plot with Linear, Quadratic, Cubic and Quartic Trend between FCT (lnW) and NNWC (lnY)

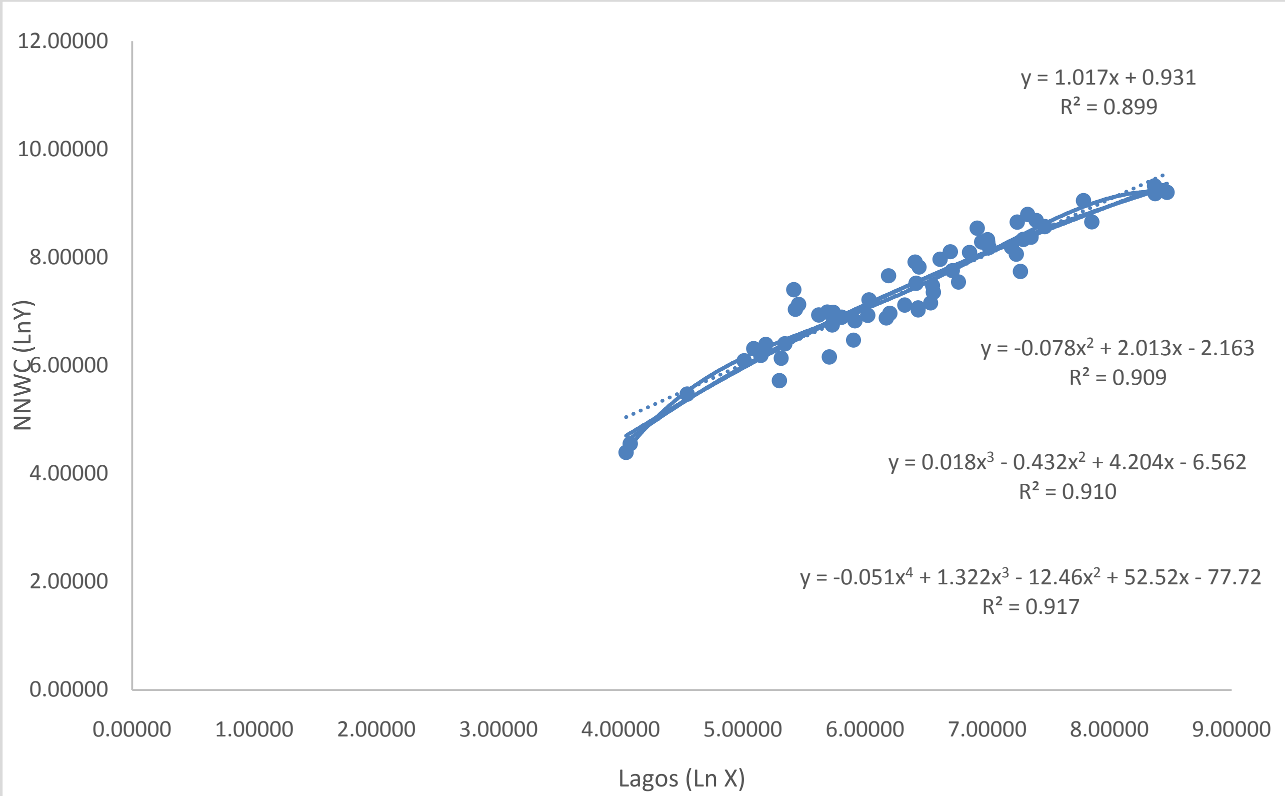

Figure 2. Scattered Plot with Linear, Quadratic, Cubic and Quartic Trend between Lagos (Ln X) and NNWC (Ln Y)

Table 3. Trend Curves Models Accuracy Measures Summary

Models Accuracy Measures	**Trend Analysis of NWC (y) against FCT (w)**			**Trend Analysis of NWC (y) against Lagos (x)**		
	Linear	Quadratic	Cubic	Linear	Quadratic	Cubic
MAPE	0.056	**0.038**	0.039	0.052	**0.051**	0.053
MAD	0.391	**0.278**	0.283	0.367	**0.362**	0.381
MSD	0.252	**0.130**	0.130	0.310	**0.220**	0.224

According to the findings in Table 2, Figure 1 and Figure 2, the quadratic trend is the best trend among trend curves to explain the Covid-19 Pandemic in the two densely populated regions of Nigeria in comparison to the National Weekly Confirmed Cases. Given that, when approximated to three decimal places, the cubic and quartic trend coefficients of powers three and four are identical to zero.

The trend predicted by the regression process is significant at an alpha level of 0.05, according to the analysis of p-value for quadratic trends, which are (0.000). This suggests that at least one of the coefficients is not equal to zero.

Both of the computed quadratic trend coefficients' p-values are 0.000, showing that they are significantly related to the Covid-19 Pandemic. The other trend after the second order has a p-value greater than 0.05, which means it is not connected to the Covid-19 Pandemic at a level of 0.05.

According to the R-square value, the predictor [Abuja (FCT) Covid-19 Pandemic] accounts for 79.8% of the variance in the Nigerian National Weekly Covid-19 Pandemic in Figure 1 and the predictor [Lagos Covid-19 Pandemic] accounts for 92.5% in Figure 2. In Figures 1 and 2, the Adjusted R-squared is 79.1% and 92.2%, respectively, accounting for the number of predictors in the model. Both results show that the quadratic trend model well fits the data sets. As a result, it appears that a quadratic trend of order two polynomial is preferable. The quadratic trend has sufficient predictive power and doesn't seem to be overfit.

The trend behavior between the national weekly confirmed cases in FCT and Lagos state was also supported by the trend curves models accuracy measures results in Table 3. These findings show that the dissemination of covid-19 in both cases increased initially before decreasing as the illness progressed over the next days or weeks (quadratic trend).

3.3. Parameter Estimates and Model Identification of Log-Normal Distribution

Estimating the mean, variance, skewness, and kurtosis of the national weekly Covid-19 confirmed cases in Lagos and the Federal Capital Territory using the log-normal distribution technique (Abuja).

To determine the statistics for the Log-normal distribution, we compare Equations (19) and (20) with Equation (6). (i.e. mean, variance, skewness, PDF, CDF, etc.).

3.3.1. Log-Normal Distribution Statistics for FCT against the Nigeria National Weekly Confirmed Cases

The identified trend model for FCT against NNWC is

$$\mathrm{Y}_t = C_0 + C_1 W_t + C_2 W_t^2$$
$$= 2.466 + 1.191 W_t - (4.04\times10^{-2}) W_t^2$$

Equation (19) compare with Equation (6).

$$= -\left[In\left(\delta\sqrt{2\pi}\right) + \frac{\mu^2}{2\delta^2}\right] + \left(\frac{\mu}{\delta^2} - 1\right) In(w) - \frac{1}{2\delta^2}(Inw)^2$$

We have a quadratic function in Equation (6), which can be expressed as

$$Inf\,(w) = C_0 + C_1 In(w) + C_2\left(In(w)\right)^2$$

$$y = Inf\,(w),\ x = In(w)\ and\ x^2 = \left[In(w)\right]^2.$$

Then, we obtain the mean and variance using the equations below

$$C_0 = -\left[In\left(\delta\sqrt{2\pi}\right) + \frac{\mu^2}{2\delta^2}\right]$$

$$C_1 = \left(\frac{\mu}{\delta^2} - 1\right)$$

$$C_2 = -\frac{1}{2\delta^2}$$

That is

$$\delta^2 = -\frac{1}{2C_2}$$

$$\mu = (C_1 + 1)\delta^2$$

Note: $C_0 = 2.466, C_1 = 1.191$ and $C_2 = -4.04\times10^{-2}$.

Substitute C_2

$$\delta^2 = -\frac{1}{2C_2} = \frac{1}{0.0808} = 12.37$$

Substitute C_1 and δ^2 to obtain the mean

$$\mu = (C_1 + 1)\delta^2 = (1.191 + 1)\times 12.37 = 27.10;$$

that is 27 approximate cases

Log-normal distribution statistics for FCT (Abuja)

Notation	$\mathrm{N}(\mu, \delta^2) = \mathrm{N}(27, 12.4)$ cases
	$-\infty < \mu < \infty$
Parameter	$\delta^2\ (12) > 0$
FCT Distribution	$w > 0$
PDF	

$$\frac{1}{\sqrt{2\pi(12)}\,w}\exp\left[-\frac{\left(In(w) - 27\right)^2}{2(12)}\right]$$

CDF

$$\frac{1}{2}\left[1 + erf\left(\frac{In(w - 27)}{3.46}\right)\right]$$

Mean

$$e\left(27 + \tfrac{1}{2}(12)\right)$$

Variance

$$\left(e^{12} - 1\right)e^{2(27)+(12)}$$

Skewness

$$\left(e^{12} + 2\right)\sqrt{\left(e^{12} - 1\right)}$$

Kurtosis

$$e^{4(12)} + 2e^{3(12)} + 23e^{2(12)} - 6$$

3.3.2. Log-Normal Distribution Statistics for Lagos against the Nigeria National Weekly Confirmed Cases

Similarly, the identified trend model for Lagos against NNWC is

$$\mathrm{Y}_t = C_0 + C_1 X_t + C_2 X_t^2$$
$$= -2.16 + 2.013 X_t - (7.91\times10^{-2}) X_t^2$$

Equation (20) compare with Equation (6).

$$= -\left[In\left(\delta\sqrt{2\pi}\right) + \frac{\mu^2}{2\delta^2}\right] + \left(\frac{\mu}{\delta^2} - 1\right) In(X) - \frac{1}{2\delta^2}(InX)^2$$

We have a quadratic function in Equation (6), which can be expressed as

$$Inf\,(X) = C_0 + C_1 In(X) + C_2\left(In(X)\right)^2$$

$$y = Inf\,(X),\ x = In(X)\ and\ x^2 = \left[In(X)\right]^2.$$

Then, we obtain the mean and variance using the equations below

$$C_0 = -\left[In\left(\delta\sqrt{2\pi}\right) + \frac{\mu^2}{2\delta^2}\right]$$

$$C_1 = \left(\frac{\mu}{\delta^2} - 1\right)$$

$$C_2 = -\frac{1}{2\delta^2}$$

That is

$$\delta^2 = -\frac{1}{2C_2}$$

$$\mu = (C_1 + 1)\delta^2$$

Note: $C_0 = -2.16, C_1 = 2.013$ and $C_2 = -7.91\times10^{-2}$
Substitute C_2

$$\delta^2 = -\frac{1}{2C_2} = \frac{1}{0.1582} = 6.32$$

Substitute C_1 and δ^2 to obtain the mean

$$\mu = (C_1+1)\delta^2 = (2.013+1)\times 6.32 = 19.07;$$

that is 19 approximate cases

Log-normal distribution statistics for Lagos

Notation	$N(\mu, \delta^2) = N(19, 6.32)$ cases
	$-\infty < \mu < \infty$
Parameter	$\delta^2\ (6) > 0$
Lagos Distribution	$x > 0$
PDF	

$$\frac{1}{\sqrt{2\pi(6)X}} \exp\left[-\frac{(In(X)-19)^2}{2(6)}\right]$$

CDF

$$\frac{1}{2}\left[1+erf\left(\frac{In(x-19)}{2.45}\right)\right]$$

Mean

$$e\left(19+\tfrac{1}{2}(6)\right)$$

Variance

$$\left(e^6-1\right)e^{2(19)+(6)}$$

Skewness

$$\left(e^6+2\right)\sqrt{\left(e^6-1\right)}$$

Kurtosis

$$e^{4(6)}+2e^{3(6)}+23e^{2(6)}-6$$

Table 4 shows that FCT log-normal distribution contributed an average of twenty seven (27) cases on the Nigeria national weekly Covid-19 confirmed cases, while Lagos contributed an average of nineteen (19) cases with variance of six. In addition, this study established its probability and cumulative density functions (PDF and CDF) for both FCT and Lagos on the Nigeria national weekly Covid-19 confirmed cases. It will be suitability for modelling the distribution of Covid-19 confirmed cases in Nigeria.

Table 4. Summary of Log-normal distribution for FCT and Lagos on the Nigeria National Weekly Covid-19 Confirmed Cases

Distribution	FCT	LAGOS
$N(\mu, \delta^2)$	N(27, 12) cases	N(19, 6) cases
PDF	$\frac{1}{\sqrt{12\pi w}} \exp\left[-\frac{(In(w)-27)^2}{24}\right]$, w > 0	$\frac{1}{\sqrt{2\pi(6)X}} \exp\left[-\frac{(In(x)-19)^2}{12}\right], x > 0$
CDF	$\frac{1}{2}\left[1+erf\left(\frac{In(w-27)}{3.46}\right)\right]$	$\frac{1}{2}\left[1+erf\left(\frac{In(x-19)}{2.45}\right)\right]$

Footnote: erf is the Gaussian error function.

4. Summary and Conclusion

4.1. Summary

The trend analysis and Log-normal distribution function on covid-19 pandemic spread between FCT/Lagos and the National Weekly confirmed pandemic cases were carried. The linear, quadratic, cubic and quartic trends were considered for both FCT and Lagos on the Nigeria National Weekly confirmed cases. Then the quadratic trend was determine to be suitability for modelling the distribution of Covid-19 confirmed cases in Nigeria (i.e. between the national weekly confirmed cases on FCT and Lagos state). These result indicate that the spread of covid-19 in both cases was appreciating at the beginning of the disease, then started depreciating as the days/weeks goes (quadratic trend).

Furthermore, Log-normal distribution approach was considered and transformed to a quadratic form for easy evaluation of the expected (or average) distribution and variation of both FCT and Lagos on the Nigeria National Weekly Covid-19 confirmed cases.

4.2. Conclusion

Probability and cumulative density functions (PDF and CDF) were established for both FCT and Lagos, thereby determine which one affect the national weekly covid-19 pandemic reported in Nigeria. The Log-normal distribution model forecasts between the National Weekly Covid-19 confirmed cases and FCT pandemic cases was twenty seven (27) cases weekly. This implied that an expected values of 27 cases weekly on the Nigeria National Weekly Covid-19 confirmed cases by FCT. Likewise, the Log-normal distribution model forecasts between the National Weekly Covid-19 confirmed cases and Lagos pandemic cases was nineteen (19) cases weekly. This implied that an expected values of 19 cases weekly on the Nigeria National Weekly Covid-19 confirmed cases by Lagos state. This implies that the effects of the number of expected pandemic cases will be more on FCT based on the Log-normal distribution approach.

References

[1] Reid F. (1994). A conversation with Sir David Cox. Statistical Science. 1994; 9(3): 439-455.
[2] Haipeng S, and Zhengyuan Z. (2008). Efficient mean estimation in lognormal linear models. *Journal of Statistical Planning and Interence*. Department of statistics and operation research, University of North Carolina, USA. 138: 552-567.
[3] Stahel W, Limper E, and Abbt M. (2001). Lognormal distribution across the science. Journal on Bioscience; 51(5).
[4] Bachioua, L. (2020). Extended Lognormal Distribution: Properties and Applications. World Scientific News Journal (WSN); 145

(2020) 16-30 EISSN 2392-2192. Available online at www.worldscientificnews.com.

[5] Asmussen, S., Jens, L. J., and Leonardo, R. (2014). A Literature Review on Lognormal Sums. Department of Mathematics, Aarhus University; School of Mathematics and Physics, University of Queensland.

[6] Christopher J. and Mitchell C. (2015). Approximating the sum of correlated lognormal. Department of System Engineering of Stevens' Institute of Technology School of Systems and Enterprises. Hoboken.

[7] Akbar A, Kamel A, Farhad Y. (2016). Testing the equality means of several log-normal distribution. Available: https://www.tandfonline.com.

[8] Orumie, U. C. and E. O. Biu (2019). Comparative Study of Failure Rate of Bank's ATM: Log Normal Distribution Approach. Asian Journal of Probability and Statistics, 3(4): 1-19, 2019; Article no.AJPAS.47921.

[9] Orumie, U. C., Biu, O. E. and Nvene, S. (2021). Comparative study of Some Parametric Survival Functions. *Academic Journal of Statistics and Mathematics (AJSM)* Vol. 7(4); ISSN (5730 – 7151); p –ISSN 4052 – 392X.

[10] Aronu, C. O., Ekwueme, G. O., Sol-Akubude, V. I., & Okafor, P. N. (2021). Coronavirus (COVID-19) in Nigeria: Survival rate. Scientific African, 11, e00689.

[11] Guillermo, S. E., Carrillo-Vega, M. F., Victor, G. G., Silvia, M. V., Filiberto, T. T., and Juan, G. E. (2020). Survival Analysis of Covid-19 in the Mexican Population. *BMC Public Health*, Vol. 20, no. 1616.

[12] Abdulmajeed, K., Adeleke, M., & Popoola, L. (2020). Online forecasting of covid-19 cases in Nigeria using limited data. Journal of Science Direct, 30, 105683.

[13] Onafeso, O. D., Onafeso, T. E., Olumuyiwa-Oluwabiyi, G. T., Faniyi, M. O., Olusola, A. O., Dina, A. O., . . . Adagbasa, E. (2021). Geographical trend analysis of COVID-19 pandemic onset in Africa. Social Sciences & Humanities Open, 4(1), 100137.

[14] Rauf, R. I., & Oladipo, H. O. (2020). Forecasting the spread of covid-19 in Nigeria using box-Jerking Modeling Procedure, the Journal of.

[15] Ajagbe, O. B., Zubair, K. and Tarry, O. (2014). Survival Analysis of Adult Tuberculosis Disease. *National Library of Medicine*. 9(11): e112838.

[16] Eckhard Limpert, Werner A. Stahel, Markus Abbt, (2001). Log-normal Distributions across the Sciences: Keys and Clues: On the charms of statistics, and how mechanical models resembling gambling machines offer a link to a handy way to characterize log-normal distributions, which can provide deeper insight into variability and probability—normal or log-normal: That is the question, BioScience, Volume 51, Issue 5, May, Pages 341-352.