
EKE, Charles N. Ph.D.
Department of Mathematics and Statistics
Federal Polytechnic Nekede, Owerri, Imo State, Nigeria
ngomeeke@gmail.com

ABSTRACT
This study examined the statistical modeling of monthly returns from the Nigerian stock exchange using probability distributions (normal and logistic distributions) for the period, 1995 to 2014. The data were analyzed to determine the returns from the stock exchange, its mean and standard deviation in order to obtain the best suitable distribution of fit. Furthermore, comparative analysis was done between the logistic distribution and normal distribution. The findings proved that the logistic distribution is better than the normal distribution in modeling the returns from the Nigeria stock exchange due to its p-value is greater than the 5% significance level alpha and also because of its minimum variance.

Keywords: Stock Exchange, Minimum Variance, Statistical Distribution, Modeling

1.0 INTRODUCTION
Stock exchange market in any economy plays an important role towards economic growth. Importantly, stock exchange is an exchange where stock brokers and traders can buy and sell stocks (shares), bonds and other securities; this stock exchange also provides facilities for issue and redemption of securities and other financial instruments and capital events including the payment of income and dividends (Diamond, 2000).

In addition, the Nigeria stock exchange was established as the Lagos stock exchange, in the year 1960, while, the stock exchange council was constituted and officially started operation on August 25th 1961 with 19 securities that were listed for trading (Ologunde et al, 2006). It became officially known as the Nigeria stock exchange in December, 1997 with various branches established in major commercial cities of the country (Ologunde et al, 2006).

Interestingly, returns from stock exchange is through trading in the secondary market, this stock exchange returns are made through dividends that are announced by companies to its customers (investors) at the end of every quarter. More so, in the secondary market, an investor can earn stock exchange returns by selling it at higher price (Nouriel, 2010). It is pertinent to also note, that stock exchange returns are exposed to market hazards because they are not stable guaranteed returns, and they are expected to change from one stockholder to another. This is largely dependent on the amount of risk the stockholder is ready to take. It also depends on the worth of stock market scrutiny. This instability in the returns from Nigeria stock exchange market is captured in the graph below, fig. 1.0.
The variability of monthly returns from the Nigeria Stock Exchange makes forecasting of future monthly returns very difficult. As a result of this unpredictability, many research works have been done in to model monthly returns from the Nigeria Stock Exchange using various methods such as time series analysis, statistical regression analysis and econometric analysis. However, none of these methods used has been able to adequately capture the monthly returns of the Nigeria Stock Exchange. Hence, probability distributions method will be effectively used as the model to analyze the monthly returns from the Nigeria Stock Exchange. The normal and logistic distributions will be fitted empirically to the monthly data collated from statistical bulletin of the Central Bank of Nigeria, 2015, covering from January 1995 to December 2014. The objectives of this research study are in two folds; First, to investigate how normal and logistic distributions explain the pattern of monthly returns from the Nigeria Stock Exchange. Secondly, to compare the two distributions of modeling monthly returns from the Nigeria Stock Exchange in order to choose the best probability distribution that can be used to forecast future monthly returns from the market.

LITERATURE REVIEW
Several approaches have been used to model monthly returns from the Nigeria Stock Exchange. Agwuegbo et al (2010) studied the daily returns process of the Nigeria Stock Exchange using discrete time Markov chain (DTMC) and martingales. Their study provided evidence that the daily stock returns process follow a random process, but that the stock market itself is not efficient even in weak form. Serkan (2008) investigated the role of growth rate of industrial production index in explaining stock returns. He employed the growth rate of industrial production index from the period of July 1997 to June 2005. He found that growth rate of industrial production index do not appear to have significant effect on stock returns. Hassan and Malik (2007) used a multivariate GARCH model to simultaneously estimate mean and conditional variance using daily returns among different financial asset. Their result revealed that there is a significant bilateral transmission of the degree in variation of trading price series overtime as measured by the standard deviation of return between the sectors. These findings support the idea of cross market hedging and sharing of common information by investors.
Ologunde et al (2006) examined the relationship between the returns from the Nigeria Stock Exchange and interest rate. They found that prevailing interest rate exerts positive influence on the returns from the Nigeria stock exchange.

Fama and French (2004) stated that if returns changes from transaction to transaction and are independent, identically distributed random variables with finite variance and if transaction are fairly uniformly spaced through time, then the central limit theorem leads us to believe that returns changes across differencing intervals such as a day, a week, or a month will be normally distributed.

Chaudhuri and Smiles (2004) used multivariate cointegration methodology to investigate the long-run relationship between stock returns and gross domestic product (GDP). They used quarterly data from 1960:1 to 1984:4, they established existence of a long-run relation between stock returns and gross domestic product (GDP). The error correction technique indicated that the stock returns are related to the changes in GDP.

Eberlein and Keller (1995) stated that the hyperbolic distribution is far more powerful in modeling returns from stock exchange than the more widely accepted normal distribution. Their result revealed that the p value for the hyperbolic distribution was very large as compare to that of the normal and student’s T distribution that was very small.

Smith (1981) applied the geometric Brownian motion in modeling stock exchange returns. The result revealed that there is a continuous change in stock price. Hence, this kind of process is often referred to as a random process because it has no structure or statistical properties to give traders the opportunities of having expected returns.

Mandelbrot (1963), used the beta coefficient to measure the returns from the Nigeria stock exchange, the result revealed that the stock market returns is largely counter cyclical being larger in bad times than in good times. Accordingly, stock expected returns lower much less during expansions than they increase during recessions, the reason for this is because the investors required return is not only counter cyclical but also asymmetrically related to the development of the business cycle which happens when the investors expected return to invest in the stock market increase more in bad times than they decrease in good time.

2.0 METHODOLOGY

This study used a model description of normal and logistic distributions to achieve a detailed modeling of monthly returns from the Nigeria Stock Exchange. In addition, the study made use of secondary data sourced from the Statistical Bulletin of the Central Bank of Nigeria 2015.

3.1 Normal Distribution

The normal distribution is a continuous distribution. It is symmetric about its mean and is non-zero over the entire real line. As such it may not be a suitable model for variables that are inherently positive (Chukwu and Amuji, 2012).

3.2 Properties of the Normal Distribution

The properties of the normal distribution are:

- The normal curve is symmetrical about the mean ($\mu$)
- The mean is at the middle and divides the area into halves.
- The total area under the curve is equal to 1.
- $E(x) = \mu$ and $\text{var}(x) = \sigma^2$

3.3 Assumptions of Normal Distribution

- The data must come from a continuous set, that is, there is an infinite amount of values between any two numbers in the data set.
- The data must follow a uni-modal bell shaped curve that is symmetric about its mean.

3.4 Probability Density Function (P.D.F).

The p.d.f of a normal distribution is given by
Mean of the normal distribution

\[ \mu = \frac{1}{n} \sum x_i \]

Where, \( x \) is a random variable, \( n \) is the sample size, \( \mu \) is the mean

Variance of the normal distribution

\[ \sigma^2 = \frac{N}{N-1} S^2 \]

\[ S^2 = \frac{1}{N} \sum_{i=1}^{N} (x_i - \bar{x})^2 \]

\[ \text{var}(\bar{x}) = \frac{s^2}{N-1} \]

Where, \( \sigma^2 \) is the population variance, \( N \) is the population size, \( S^2 \) is the sample variance, \( x \) is the random variable, \( \bar{x} \) is the sample mean, \( \text{var}(\bar{x}) \) is variance of the sample mean

Kurtosis of the normal distribution

\[ kurt(x) = \frac{\text{E}(x^4) - 4\mu \text{E}(x^2) + 6\mu^2 \text{E}(x^2) - 3\mu^4}{\sigma^4} = \frac{\text{E}(x^4) - 4\mu \text{E}(x^2) + 6\mu^2 \text{E}(x^2) - 3\mu^4}{\sigma^4} \]

Skewness of the normal distribution

\[ skew(x) = \frac{\text{E}(x^3) - 3\mu \text{E}(x^2) - \mu^3 \sigma^2}{\sigma^3} = \frac{\text{E}(x^3) - 3\mu \text{E}(x^2) - \mu^3 \sigma^2}{\sigma^3} \]

3.5 Logistic Distribution

The logistic distribution is a continuous probability distribution. It resembles the normal distribution, but has higher kurtosis (Feller, 1971).

3.6 Probability Density Function (P.D.F)

The p.d.f of the logistic distribution is given by

\[ f(x; \mu, \sigma) = \frac{e^{-\frac{x-\mu}{\sigma}}}{\sigma \left(1 + e^{-\frac{x-\mu}{\sigma}}\right)^2} \]

\[ -\infty < x < \infty, -\infty < \mu < \infty \text{ and } 0 < \sigma < \infty \]

Where, \( x \) is a random variable, \( \mu \) is the mean of \( x \), \( \sigma \) is the standard deviation

Mean of the logistic distribution

\[ \mu = \frac{1}{\lambda} [\psi(\alpha) - \psi(\beta)] \]

Variance of the logistic distribution

\[ \text{var}(x) = \frac{\pi^2}{3} = \frac{1}{\lambda^2} [\psi(\alpha) + \psi(\beta)] \]

Where, \( \alpha \) is alpha function, \( \beta \) is beta function

Skewness of the logistic distribution

\[ skew(x) = E(x)^3 = 0 \]
Kurtosis of the logistic distribution
\[ kurt(x) = \frac{E(x^4)}{\text{var}(x^2)} = \frac{21}{5} \] (Feller, 1971).

3.7 Properties of Logistic Distribution
i. Uni-modality: The logistic distribution is unimodal, hence the logistic distribution has a close approximation to the normal distribution.
ii. As the \( \mu \) decreases, the pdf is shifted to the left, as \( \mu \) increases, the pdf is shifted to the right, Hence \( \mu \) is the location parameter of logistic distribution.
iii. The logistic distribution is bell shaped.

3.8 Assumption of Logistic Distribution
i. Data must come from a continuous set and its sample size must be large.
ii. Logistic distribution can handle ordinal and nominal data as independent variable, hence the independent variable do not need to be metric (interval or ratio scaled).

3.9 Maximum Likelihood Estimator (M.L.E)
Suppose that \( x_1, x_2, ..., x_n \) are iid random variables with common p.d.f \( f(x; \theta), \theta \in \Omega \), we assume that the parameter \( \theta \) is a scalar and unknown. The likelihood function is given by
\[ l(\theta; x) = \prod_{i=1}^{n} f(x_i; \theta), \theta \in \Omega \]

Where, \( x_i = x_1, x_2, ..., x_n \) is a random variable, \( \theta \) is a scalar and unknown, \( \Omega \) is sample space.
The log of the function is given by
\[ \log l(\theta) = \sum_{i=1}^{n} \log f(x_i; \theta), \theta \in \Omega \]

We say that \( \hat{\theta} = \hat{\theta}(x) \) is a maximum likelihood estimator (mle) of \( \theta \) if \( \hat{\theta} = \text{Argmax} \, l(\theta; x) \)
That is, if \( l(\theta; x) \) achieves its maximum value of \( \hat{\theta} \).
To determine the maximum likelihood estimator (MLE) of a distribution, we often take the log of the likelihood and determine its critical value. That is letting \( l(\theta) = \log l(\theta) \).

The MLE solves the equation
\[ \frac{\partial l(\theta)}{\partial \theta} = 0 \] (Hogg et al, 2005)

3.9 Maximum Likelihood Estimator for Normal Distribution
Suppose \( x \sim N(\mu, \sigma^2) \). The p.d.f of normal distribution is given by
\[ f(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \]
The likelihood function is given by
\[ l(x, \mu, \sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \]
\[ l(x, \mu, \sigma^2) = \left( \frac{1}{\sqrt{2\pi\sigma}} \right)^n e^{-\frac{\Sigma(x-\mu)^2}{2\sigma^2}} \]
The log - likelihood function is
\[ \log(x, \mu, \sigma^2) = n \log \frac{1}{\sqrt{2\pi\sigma}} - \frac{1}{2\sigma^2} \Sigma_{i=1}^{n} (x - \mu)^2 \]
To obtain the maximum likelihood estimate for \( \mu \), differentiate the log – likelihood function w.r.t \( \mu \)
To obtain the maximum likelihood estimate for $\sigma^2$, differentiate the log–likelihood function w.r.t $\sigma^2$

$$\frac{\partial \log(x, \mu, \sigma^2)}{\partial \sigma^2} = -\frac{n}{\sigma} + \frac{1}{\sigma^3} \sum_{i=1}^{n} (x - \mu)^2 = 0$$

$$-n\sigma^2 + \sum_{i=1}^{n} (x - \mu)^2 = 0$$

$$-n\sigma^2 = -\sum_{i=1}^{n} (x - \mu)^2$$

$$\sigma^2 = \frac{\sum(x-\mu)^2}{n} \quad \text{(Grete et al, 2003)}.$$  

### 3.10 Maximum Likelihood Estimator for Logistic Distribution

Let $x_1, x_2, \ldots, x_n$ are iid be with density

$$f(x, \theta) = \frac{\exp(-(x - \theta))}{1 + \exp(-(x - \theta))^2}$$

$-\infty < x < \infty, -\infty < \theta < \infty$

The log of the likelihood simplifies to

$$l(\theta) = \sum_{i=1}^{n} \log f(x_i, \theta) = n\theta - n\bar{x} - 2 \sum_{i=1}^{n} \log(1 + \exp - (x - \theta))$$

Using first derivative is

$$l'(\theta) = n - 2 \sum_{i=1}^{n} \frac{\exp(-(x-\theta))}{1+\exp(-(x-\theta))} \quad \text{(Hogg et al, 2005)}$$

### 3.11 Kolmogorov-Smirnov Goodness of Fit Test

The Kolmogorov-Smirnov goodness of fit test can be defined by the function

$$F_n(x) = \begin{cases} 
  \frac{k}{n}, & x_k \leq x < y_k + 1 \\
  1, & y_n \leq x 
\end{cases} \quad k = 1, 2, \ldots, n - 1.$$

The Kolmogorov-Smirnov test statistic is defined by

$$D_n = \sup_x \{|F_n(x) - F_o(x)|\}$$

Where

$D_n$ is the least upper bound of all point wise differences $|F_n(x) - F_o(x)|$.

The distribution of $D_n$ does not depend on the particular function $F_o(x)$ of the continuous type. The Kolmogorov-Smirnov statistics is used to test the hypothesis $H_0: F(x) = F_o(x)$ Against the alternatives $H_1: F(x) \neq F_o(x)$.

Where $F_o(x)$ is some distribution functions.
We accept $H_0$ if the function $F_n(x)$ is sufficiently close to $F_0(x)$. That is if the value of $D_n$ is sufficiently small. The hypothesized $H_0$ is rejected if the observed value of $D_n$ is greater than the critical value, where the critical value depends on the significance level and sample size (Hogg and Tanis, 2010).

### 3.12 Minimum Variance Unbiased Estimator (M.V.U.E)

This is an unbiased estimator that has lower variance than any other unbiased estimator for all possible values of the parameter. For a given positive integer $n$, $Y = u(x_1, x_2, \ldots, x_n)$ will be called a minimum variance unbiased estimator of the parameter $\theta$ if $Y$ is less than or equal to the variance of every other unbiased estimator of $\theta$. Therefore, if $x_1, x_2, \ldots, x_n$ are random samples from a normal population with mean ($\mu$) and variance ($\sigma^2$), and $T_1, T_2$ are both unbiased estimators of the population parameter ($\theta$). Then $T_1$ is MVUE if $\forall \theta \in \Omega$, $var(T_1) < var(T_2)$ (Chukwu and Amuji, 2012).

### 3.13 Stock Exchange Returns (SER)

Stock exchange returns are the returns that the investors generate out of the stock market.

$$SER = \frac{(Y_t - Y_{t-1})}{Y_t} \quad or \quad SER \ (100\%)$$

Where, SER = stock exchange return, $Y_t$ = Current value of all – share index, $Y_{t-1}$ = Previous value of the all – share index

### 4.0 ANALYSIS


**Table 4.0 Descriptive Statistics for the data**

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.0137</td>
</tr>
<tr>
<td>Variance</td>
<td>0.0046</td>
</tr>
<tr>
<td>Skewness (Pearson)</td>
<td>0.2968</td>
</tr>
<tr>
<td>Kurtosis (Pearson)</td>
<td>5.3208</td>
</tr>
</tbody>
</table>

Source: Author’s Computation

Table 4.0 shows the descriptive statistics for the monthly stock exchange returns. The table shows that from the 239 observations, the mean value of monthly returns is 0.0137 or 1.37% with variance to be 0.0046 or 0.46%. The skewness of 0.2968 showed that the data is skewed to the right; while, the kurtosis of 5.3208 showed that the distribution of the monthly returns is highly peaked with fat tail.

#### 4.2 Theoretical Distribution Descriptive Statistics

**Table 4.1: Theoretical Descriptive Statistics**

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Data</th>
<th>Logistic Distribution</th>
<th>Normal Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.0137</td>
<td>0.0121</td>
<td>0.0137</td>
</tr>
<tr>
<td>Variance</td>
<td>0.0046</td>
<td>0.0040</td>
<td>0.0046</td>
</tr>
<tr>
<td>Skewness (Pearson)</td>
<td>0.2968</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>Kurtosis (Pearson)</td>
<td>5.3208</td>
<td>1.2000</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

Source: Author’s Computation
Table 4.1 shows the descriptive statistics for each distribution. The logistic distribution presented a mean monthly return of 0.0121 or 1.21%, while the normal distribution gave a mean of 0.0137 or 1.37% which is the same for the original data. In terms of the variance, logistic distribution showed a variance of 0.0040 or 0.40%, while the normal distribution mirrored the variance of the original monthly returns data of 0.0046 or 0.46%. Both logistic and normal distributions showed a symmetric distribution with the skewness value of 0.0000. The monthly returns data did not exhibit high peakedness based on kurtosis value of 1.2000 and 0.0000 are for logistic and normal distributions respectively.

### 4.3 Goodness of Fit Test

**Table 4.2: Goodness of Fit – Summary for Logistic Distribution**

<table>
<thead>
<tr>
<th>D</th>
<th>0.0445</th>
</tr>
</thead>
<tbody>
<tr>
<td>p-value</td>
<td>0.7211</td>
</tr>
<tr>
<td>alpha</td>
<td>0.05</td>
</tr>
</tbody>
</table>

Source: Author’s computation

Table 4.2 reveals the Kolmogorov–Smirnov test for the logistic distribution. This table gave D as 0.00445, p-value as 0.7211 at significance level of 0.05%. As the computed p-value is greater than the significance level of 0.05%, this means that the monthly returns from the Nigeria Stock Exchange follows a logistic distribution. Hence for the logistic distribution the p.d.f is given as:

\[
f(x; \mu, \sigma) = \frac{1}{\sigma} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)\]

\[-\infty < x < \infty, -\infty < \mu < \infty and 0 < \sigma < \infty\]

**Table 4.3: Goodness of Fit – Summary for Normal Distribution**

<table>
<thead>
<tr>
<th>D</th>
<th>0.0677</th>
</tr>
</thead>
<tbody>
<tr>
<td>p-value</td>
<td>0.2154</td>
</tr>
<tr>
<td>alpha</td>
<td>0.05</td>
</tr>
</tbody>
</table>

Source: Author’s computation

Table 4.3 reveals the Kolmogorov–Smirnov test for the normal distribution. This table gave D as 0.0677, p-value as 0.2154 at significance level of 0.05%. As the computed p-value is greater than the significance level of 0.05%, this means that the monthly returns from the Nigeria Stock Exchange follows a normal distribution. Hence for the normal distribution the p.d.f is given as:

\[
f(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)\]

\[-\infty < x < \infty, -\infty < \mu < \infty and 0 < \sigma < \infty\]

### 4.4 Best Fit Distribution

The logistic distribution fits the monthly returns from the Nigeria Stock Exchange; and equally the normal distribution fits the monthly returns from the Nigeria Stock Exchange. Using the theorem of minimum variance unbiased estimator (MVUE) which states that if \( \text{var}(T_1) < \text{var}(T_2) \) then \( T_1 \) is more efficient than \( T_2 \). This means that the logistic distribution is better than the normal distribution in modeling the monthly returns from the Nigeria Stock Exchange because of its minimum variance.

### 5.0 CONCLUSION

In the final analysis, the future monthly returns from the Nigeria Stock Exchange can be modeled using the logistic distribution. A lot of models have been developed to help investors understand the variability inherent in that market. Therefore, this research study has revealed that logistic distribution can be used by investors and other market participants in the Nigeria Stock Exchange to predict the behavior of the future monthly returns.
REFERENCES