European Journal of Statistics and Probability
Vol.8, No.2, pp, 25-40, September 2020
Published by ECRTD-UK
Print ISSN: 2055-0154(Print), Online ISSN 2055-0162(Online)

APPLICATION OF CUBE ROOT TRANSFORMATION OF ERROR COMPONENT OF MULTIPLICATIVE TIME SERIES MODEL

[1]Dike A. O.,[2] Chikezie D. C.,& [3]Otuonye E. L,.
[1]Department of Mathematics/Statistics, Akanu Ibiam Federal Polytechnic, Unwana Nigeria
[2,3]Department of Statistics, Abia State University, Uturu, Nigeria

ABSTRACT: *This paper makes use of cube transformation of the error component of multiplicative time series model. Data from federal road safety commission (FRSC) Nigeria on road accident were collected and analyzed by fitting the regression line of log mean (logmean) against log standard deviation (logstdev). This gave a fitted slope* $\beta = 0.666977$ *which agrees with the required value of 0.6666 this gives a transformation of 1-0.666977= 0.333023 (1-*β*) which is the cube root transformation. Data were later decomposed into time series components. Recommendations on areas of application of cube root transformation were equally given.*

KEYWORDS: Logmean, Cube Transformation, Components, Multiplicative, Time Series.

INTRODUCTION

The normal distribution (Gaussian distribution) is the best well known and most frequently used in probability distribution theory. It is widely used in natural and social sciences to represent real-valued random variables whose distributions are not known.The normal distribution derived its usefulness from the central limit theorem (Hogg and Craig,1978). The central limit theorem states that the averages of random variables independently drawn from independent distributions converge in distribution to the normal, that is, become normally distributed when the number of random variables is sufficiently large. Physical quantities that are expected to be the sum of many independent processes (such as measurement errors) often have distributions that are nearly normal. In practice the random variable X which has a $N(1,\sigma^2)$ distribution do not admit values less than or equal to zero. We therefore disregard or truncate all values of $x \leq 0$ to take care of the admissible region $x > 0$. Now if the values of x below or equal to zero cannot be observed due to censoring or truncation, then the resulting distribution is a left-truncated normal distribution. The

European Journal of Statistics and Probability
Vol.8, No.2, pp, 25-40, September 2020
Published by ECRTD-UK
Print ISSN: 2055-0154(Print), Online ISSN 2055-0162(Online)

restriction necessitated the derivation of distributions under the truncated environment. One such example is the error component of the multiplicative time series model which assume to be $N(1, \sigma^2)$

According to Spiegel and Stephens [4] the general time series model is always considered as a mixture of four major components, namely the Trend, Seasonal movements, cyclical movements and irregular or Random Movements. Hence classifications of the time series model are

Multiplicative model: $X_t = T_t S_t C_t e_t$ (5)

Additive model: $X_t = T_t + S_t + C_t + e_t$ (6)

Mixed model: $X_t = T_t S_t C_t + e_t$ (7)

In short term series the trend and cyclical components are merged to give the trend-cycle component; hence equation (5) through (7) can be rewritten as

$$X_t = M_t S_t C_t e_t \tag{8}$$

$$X_t = M_t + S_t + C_t + e_t \tag{9}$$

$$X_t = M_t S_t C_t + e_t \tag{10}$$

where M_t is the trend cycle component and e_t is independent identically distributed (iid) normal errors with mean 1 and variance $\sigma^2 > 0$ [$e_t \sim N(1, \sigma^2)$]

The probability density function of the left truncated normal distribution is given in Uche (2003) as

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \; e^{-\frac{1}{2}\left(\frac{x-1}{\sigma}\right)^2}, x \geq 0, \sigma^2 > 0 \qquad (1)$$

The error component e_t of the multiplicative time series model has a pdf $N(1, \sigma^2)$ where $e_t > 0$, Iwueze (2007) established the distribution of the left-truncated normal distribution and is given by

$$f(x) = \frac{e^{-\frac{1}{2}\left(\frac{x-1}{\sigma}\right)^2}}{\sigma\sqrt{2\pi}\left[1 - \Phi\left(\frac{-1}{\sigma}\right)\right]}, x \geq 0, \sigma^2 > 0 \qquad (2)$$

With mean $E(X)$ and variance $Var(X)$ given by

$$E(X) = 1 + \frac{\sigma e^{-\frac{1}{2}\sigma^2}}{\sigma\sqrt{2\pi}\left[1-\Phi\left(\frac{-1}{\sigma}\right)\right]}, x \geq 0, \sigma^2 > 0 \qquad (3) \text{ and}$$

$$Var(X) = \frac{\sigma^2}{2\left(1-\Phi\left(-\frac{1}{\sigma}\right)\right)}\left(\left[1 + P_r\left(\chi^2_{(1)} \leq \frac{1}{\sigma^2}\right)\right]\right) - \frac{\sigma e^{-\frac{1}{2}\sigma^2}}{\sqrt{2\pi}\left[1-\Phi\left(\frac{-1}{\sigma}\right)\right]} - \left[\frac{\sigma e^{-\frac{1}{2}\sigma^2}}{\sqrt{2\pi}\left[1-\Phi\left(\frac{-1}{\sigma}\right)\right]}\right]^2 \qquad (4)$$

respectively.

Iwueze (2007) examined some implications of truncating $N(1, \sigma^2)$ to the left. Which include:

(i) that the truncated values are always greater or equal to the non-truncated values for all values of σ. However, the two stochastic variables behave alike in the interval $\sigma<0.30$. It follows from the analysis that the 0.001 limits may be used to give practical assurance that the truncated (truncation at zero) values from the $N(1, \sigma^2)$ distribution are all positive. In the interval $\sigma<0.30$, the truncated and the non-truncated variables have the same mean equal to 1 and variance equal to σ^2.

(ii) The most important implication of truncating the $N(1, \sigma^2)$ distribution to the left at zero is in descriptive modelling of time series data, where the logarithmic transform of the truncated distribution is equally assumed to have mean zero and some finite variance. It was noted that the logarithmic transform will have mean zero and the same variance as both the original $N(1, \sigma^2)$ distribution and its truncated distribution in the interval $\sigma<0.10$.

Vol.8, No.2, pp, 25-40, September 2020
Published by ECRTD-UK
Print ISSN: 2055-0154(Print), Online ISSN 2055-0162(Online)

The cube root transformation of the error component of the multiplicative time series model assumes that the error $e_t > 0$ to take care of transformations where the values must necessarily be greater than zero.

The truncated normal distribution has gained much acceptance in various fields of human endeavours, these include inventory management, regression analysis, operation management, time series analysis and so on. Johnson and Thomopoules (2004) considered the use of the left truncated distribution for improving achieved service levels. They presented the table of the cumulative distribution function of the left truncated normal distribution and derived the characteristic parameters of the distribution, and also presented the table of the partial expectation of the left truncated normal distribution

In situation where the cube transformation is to be applied the following steps should be adopted for it to be successful and serve the need for which it was adopted.

This would be achieved by ensuring that

(i) The model used to analyze the error component is multiplicative
(ii) It fits the transformation to be adopted
(iii) The untransformed and the transformed meet the conditions for normality. These measures will guarantee that data satisfy the assumption inherent in the statistical inference to be applied and ensure improved interpretation as expressed by Osborne(2002); who expressed that caution should be exercised on the choice of transformation to be used so that the fundamental structure of the series is not altered

European Journal of Statistics and Probability
Vol.8, No.2, pp, 25-40, September 2020
Published by ECRTD-UK
Print ISSN: 2055-0154(Print), Online ISSN 2055-0162(Online)

This paper makes us of the nth power transformation given by Dike et al (2016) with probability density function (pdf) of

$$f(y) = \frac{\frac{1}{n}.y^{\frac{1}{n}-1}\, e^{-\frac{1}{2}\left(\frac{y^{\frac{1}{n}}-1}{\sigma}\right)^2}}{\sigma\sqrt{2\pi}\left[1-\Phi\left(\frac{-1}{\sigma}\right)\right]}$$

As a general rule to all transformation, but did not show any practical application and this what this paper is trying to show. Bartlett (1947) gave conditions for different transformation

i. If the variance is not stable the appropriate transformation is determined using Bartlett (1947) as was applied by Akpanta and Iwueze (2009);

$$Y = \begin{cases} \log_e X, & \beta = 1 \\ X^{1-\beta}, & \beta \neq 1 \end{cases} \quad \textbf{(7)}$$

The linear relationship between the natural log of periodic standard deviations (log*eσi*) and natural log of the periodic means (log*eμi*) is given as

$$\log_e \sigma_i = \alpha + \beta \log_e \mu_i \quad (8)$$

European Journal of Statistics and Probability
Vol.8, No.2, pp, 25-40, September 2020
Published by ECRTD-UK
Print ISSN: 2055-0154(Print), Online ISSN 2055-0162(Online)

The value of slope β according to Bartlett (1947) should be approximately

0.3333for the inverse square root transformation (see Table 1)

Table 1. Bartlett's transformations for some values of β

	0	$\frac{1}{2}$	**1**	$\frac{3}{2}$	**2**	**3**	**-1**
Transformation	No transformation	$\sqrt{X}$	$\log_e X$	$\frac{1}{\sqrt{X}}$	$\frac{1}{X}$	$\frac{1}{X^2}$	X^2

Practical example to illustrate the effect of the cube root transformation.

In this section, real data would be used to illustrate the effect of cube root transformation on the error component of a multiplicative time series model. The following procedure would be adopted in carrying out the analysis.

(i)Validation of the multiplicative time series model and data decomposition to obtain the residual series (error component $\left(e_t\right)$) of the original data $\left(X_t\right)$)

(ii)Calculation of the mean and variance of$\left(e_t\right)$) and assessment of its normality using Kolmogorov- Smirnov test.

(iii)Justification of the cube root transformation of the original data $\left(X_t\right)$

(iv)Cube root transformation of the error series to obtain$\left(e_t^*\right)$)

(v)Calculation of the mean and variance of $\left(e_t^*\right)$

(vi) Test for normality of the transformed error component using Kolmogorov- Smirnov test.

(vii) Comparison of $\left(e_t\right)$ and $\left(e_t^*\right)$ using the means and variances, ratio of variances and Kolmogorov- Smirnov test statistics.

European Journal of Statistics and Probability
Vol.8, No.2, pp, 25-40, September 2020
Published by ECRTD-UK
Print ISSN: 2055-0154(Print), Online ISSN 2055-0162(Online)

Data for this example is the monthly reported cases of road accidents in Imo state, Nigeria for the period 2009 – 2014. The data is presented in table 1 a table of the periodic means and standard deviation and their natural logarithms is presented in table 2

Table 1: Buys-Ballot table of road accidents.

Year	Jan	Feb	March	April	May	June	July	Aug	Sept	Oct	Nov	Dec
2009	21	25	35	55	53	37	40	37	38	34	35	29
2010	36	34	22	37	33	37	57	29	38	41	47	38
2011	38	60	57	52	39	60	32	34	49	28	39	52
2012	37	31	38	45	40	36	30	34	34	38	42	41
2013	35	33	40	36	38	45	56	39	42	36	31	29
2014	10	16	13	20	14	14	17	26	15	17	26	16

Validation of the application of the multiplicative model

Validation of the multiplicative model was done using the plots of the periodic means and standard deviations Iwueze *et al* 2011).The appropriate data transformation was assessed using the Bartlett's transformation procedure Bartlett (1947) as applied by Akpanta and Iwueze (2009). This method was achieved by plotting the natural logarithm of the periodic (yearly) standard deviations against the logarithm of the periodic means and the slope of the linear relationship obtained was used to determine the nature of the transformation to be adopted. For the cube root transformation, a slope of 0.66 or its approximate value is required. All other calculations involved in the above outlined steps were done using Minitab software.

Table 2: Periodic means and standard deviations/ natural logarithm of the means and standard deviations.

Year	Mean	Standard deviation	Logmean	Logstdev
2009	36.5833	9.8577	3.59959	2.28825
2010	37.4167	8.6913	3.62212	2.16232
2011	45.0000	11.3137	3.80666	2.42602
2012	37.1667	4.4687	3.61541	1.49711
2013	38.3333	7.1647	3.64632	1.96917
2014	17.0000	4.8617	2.83321	1.58139

Validation of the Application of the Multiplicative model.

A plot of the periodic mean and standard deviation is given on Figure 1 to show whether the multiplicative model is applicable to the data. Based on the figure, it can be generally inferred that both plots are increasing with each other which suggest the use of multiplicative model. (Iwueze 2011)

European Journal of Statistics and Probability
Vol.8, No.2, pp, 25-40, September 2020
Published by ECRTD-UK
Print ISSN: 2055-0154(Print), Online ISSN 2055-0162(Online)

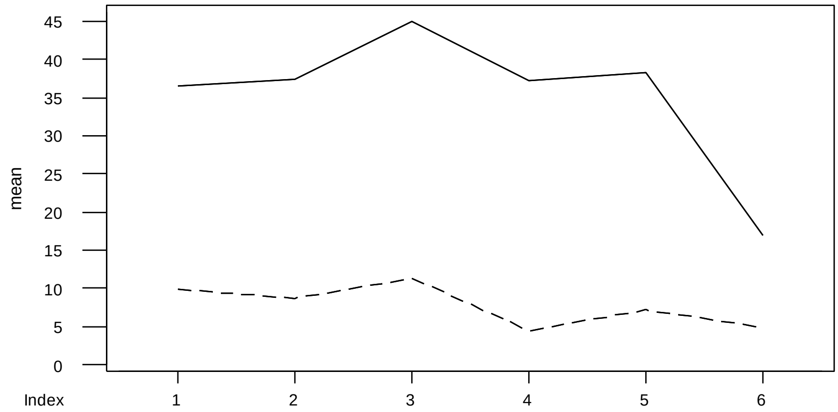

Fig. 1: Plot of the periodic mean and standard deviation

Identification of the appropriate transformation.

A linear plot of the periodic natural logarithm of the mean and standard deviation of the period.

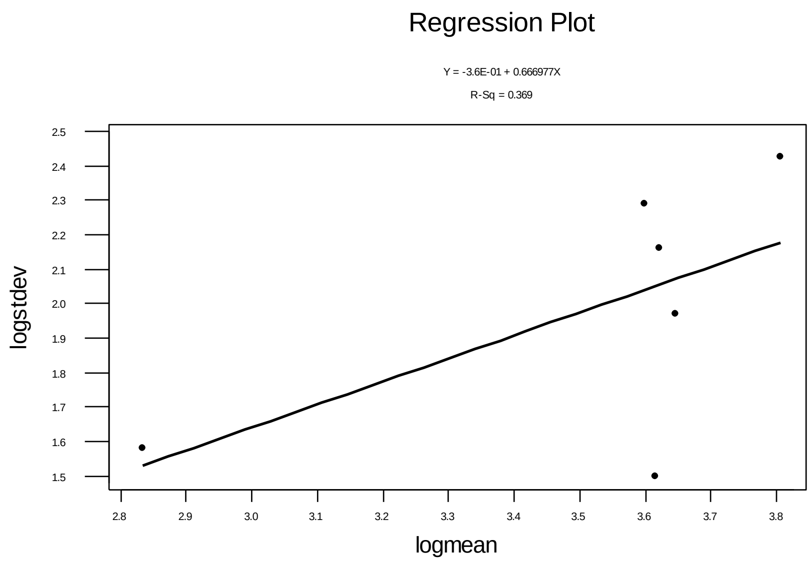

Fig. 2: Plot of log mean against log standard deviation

Because the slope β is 0.666977, it is not necessary to test its significance because it is about the same value as the required value of 0.6666. This gives the transformation of 1 - 0.666977 = 0.333023 (1- β) which is the cube root transformation.

European Journal of Statistics and Probability
Vol.8, No.2, pp, 25-40, September 2020
Published by ECRTD-UK
Print ISSN: 2055-0154(Print), Online ISSN 2055-0162(Online)

Data Decomposition

A time plot of the series is shown on Figure 3. In order to determine the appropriate trend to adopt, three accuracy measures were considered. The measures are; Mean Absolute percentage Error (MAPE), Mean Absolute Deviation (MAD) and Mean squared Deviation.

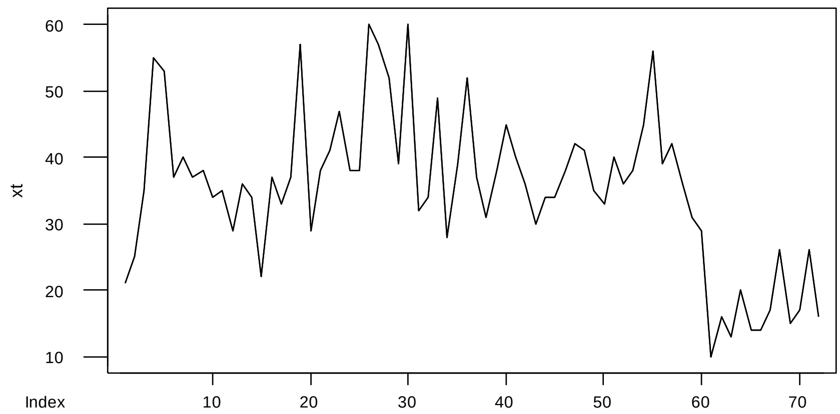

Figure 3: Time Plot of the Reported Cases of Road Accidents

In order to determine the appropriate trend line to fit, the two likely fits (linear and quadratic) were considered.

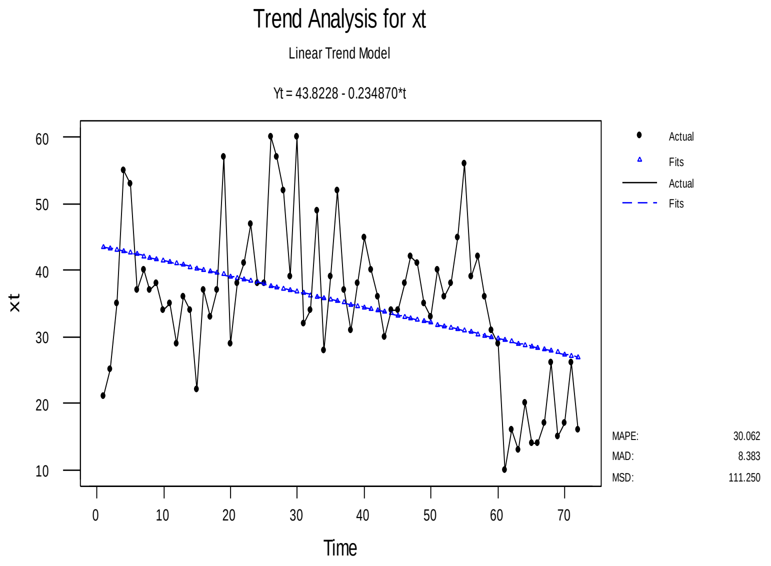

Figure 4: Linear trend fit of road accidents

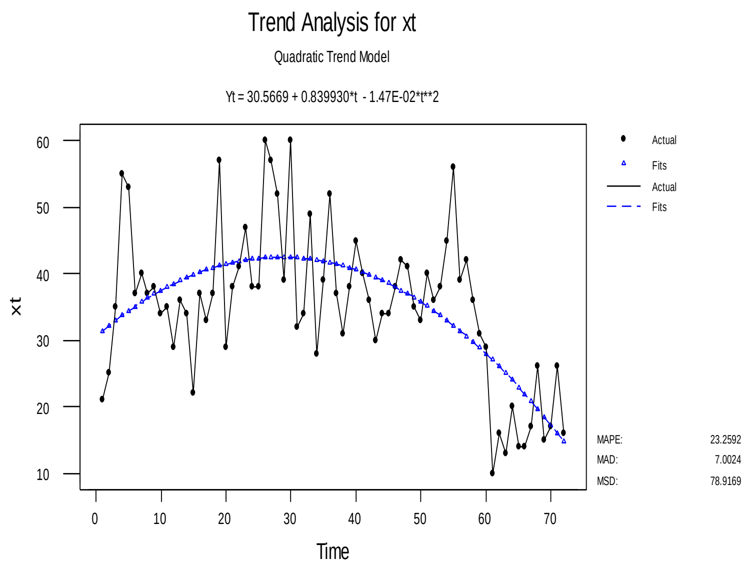

Figure 5: Quadratic trend fit of road accidents

From the accuracy measures, it is observed that the quadratic trend is the most appropriate whose trend curve is given on Figure 5 as $Y_t = 30.5669 + 0.839930*t - 0.0147*t**2$

European Journal of Statistics and Probability
Vol.8, No.2, pp, 25-40, September 2020
Published by ECRTD-UK
Print ISSN: 2055-0154(Print), Online ISSN 2055-0162(Online)

The decomposition of the time series data on road accidents into the trend, seasonal and irregular components is shown in table 3**.** Also the transformed error series is shown on the table.

Table 3: Decomposition of Time Series Data of road accidents

X_t	T	Trend	Seasonal	Untransformed error series (e_t)	Transformed error series(e_t^*)
21	1	31.3921	0.93812	0.71308	0.89341
25	2	32.1880	0.85562	0.90775	0.96826
35	3	32.9544	1.02310	1.03809	1.01254
55	4	33.6914	1.07233	1.52236	1.15036
53	5	34.3991	0.93546	1.64703	1.18094
37	6	35.0773	1.02408	1.03001	1.00990
40	7	35.7261	1.10215	1.01586	1.00526
37	8	36.3455	0.93999	1.08300	1.02693
38	9	36.9356	1.02411	1.00460	1.00153
34	10	37.4962	0.97096	0.93388	0.97746
35	11	38.0274	1.05065	0.87602	0.95684
29	12	38.5293	1.06343	0.70778	0.89119
36	13	39.0017	0.93812	0.98392	0.99461
34	14	39.4447	0.85562	1.00742	1.00247
22	15	39.8584	1.02310	0.53949	0.81409
37	16	40.2426	1.07233	0.85741	0.95002
33	17	40.5974	0.93546	0.86894	0.95426
37	18	40.9228	1.02408	0.88288	0.95933
57	19	41.2189	1.10215	1.25469	1.07856
29	20	41.4855	0.93999	0.74367	0.90601
38	21	41.7227	1.02411	0.88933	0.96166
41	22	41.9306	0.97096	1.00705	1.00234
47	23	42.1090	1.05065	1.06234	1.02036
38	24	42.2580	1.06343	0.84560	0.94564
38	25	42.3777	0.93812	0.95585	0.98506
60	26	42.4679	0.85562	1.65124	1.18194
57	27	42.5287	1.02310	1.31001	1.09418
52	28	42.5601	1.07233	1.13939	1.04445
39	29	42.5622	0.93546	0.97952	0.99313
60	30	42.5348	1.02408	1.37744	1.11264
32	31	42.4780	1.10215	0.68351	0.88089
34	32	42.3919	0.93999	0.85325	0.94848
49	33	42.2763	1.02411	1.13176	1.04212
28	34	42.1313	0.97096	0.68447	0.88130
39	35	41.9570	1.05065	0.88471	0.96000
52	36	41.7532	1.06343	1.17113	1.05406
37	37	41.5200	0.93812	0.94992	0.98302

31	38	41.2574	0.85562	0.87817	0.95762
38	39	40.9655	1.02310	0.90667	0.96787
45	40	40.6441	1.07233	1.03249	1.01072
40	41	40.2933	0.93546	1.06121	1.02000
36	42	39.9132	1.02408	0.88075	0.95856
30	43	39.5036	1.10215	0.68904	0.88326
X_t	t	Trend	Seasonal	Untransformed error series (e_t)	Transformed error Series(e_t^*)
34	44	39.0646	0.93999	0.92592	0.97467
34	45	38.5963	1.02411	0.86018	0.95104
38	46	38.0985	0.97096	1.02725	1.00900
42	47	37.5713	1.05065	1.06398	1.02089
41	48	37.0147	1.06343	1.04160	1.01368
35	49	36.4288	0.93812	1.02415	1.00799
33	50	35.8134	0.85562	1.07693	1.02501
40	51	35.1686	1.02310	1.11170	1.03592
36	52	34.4945	1.07233	0.97325	0.99100
38	53	33.7909	0.93546	1.20214	1.06328
45	54	33.0579	1.02408	1.32924	1.09950
56	55	32.2956	1.10215	1.57327	1.16304
39	56	31.5038	0.93999	1.31698	1.09611
42	57	30.6826	1.02411	1.33663	1.10154
36	58	29.8320	0.97096	1.24285	1.07515
31	59	28.9521	1.05065	1.01912	1.00633
29	60	28.0427	1.06343	0.97245	0.99073
10	61	27.1039	0.93812	0.39329	0.73268
16	62	26.1358	0.85562	0.71549	0.89442
13	63	25.1382	1.02310	0.50547	0.79660
20	64	24.1112	1.07233	0.77354	0.91798
14	65	23.0549	0.93546	0.64914	0.86587
14	66	21.9691	1.02408	0.62228	0.85376
17	67	20.8539	1.10215	0.73964	0.90437
26	68	19.7093	0.93999	1.40339	1.11958
15	69	18.5354	1.02411	0.79021	0.92452
17	70	17.3320	0.97096	1.01018	1.00338
26	71	16.0992	1.05065	1.53713	1.15407
16	72	14.8371	1.06343	1.01406	1.00466

Normality test of the original error series and the transformed series.

The error components of the original data and the transformed series were tested for normality using Kolmogorov-Smirnov normality test as shown on Figure 7 and 8

Print ISSN: 2055-0154(Print), Online ISSN 2055-0162(Online)

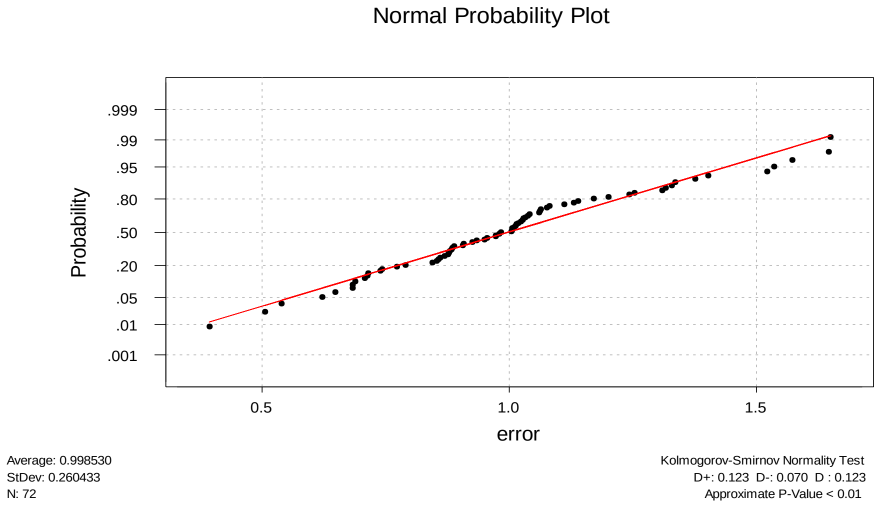

Fig 6: Normality plot of the original error series

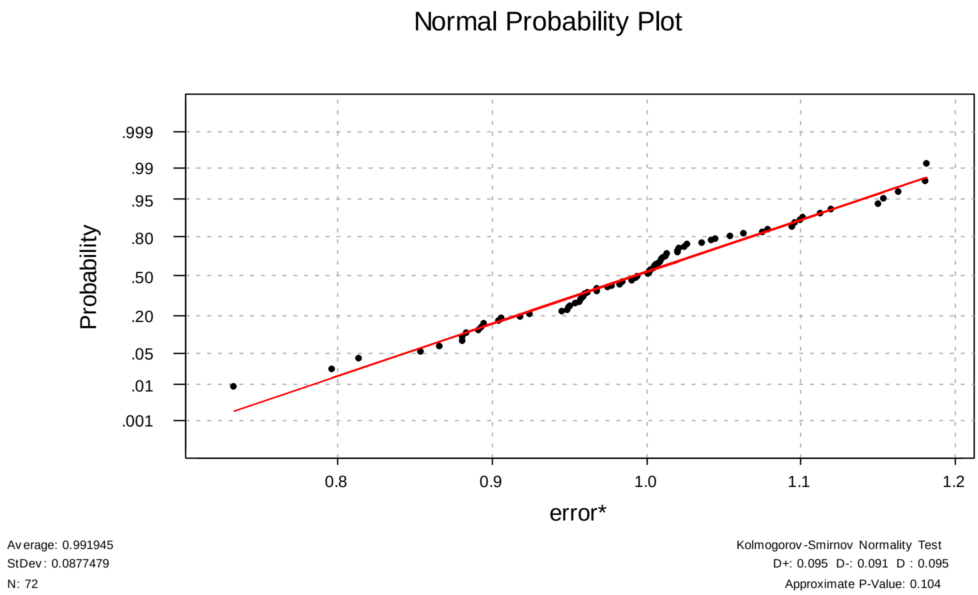

Fig 7: Normality plot of the cube root transformed error series.

From the normality plot and the Kolmogorov-Smirnov normality test, the approximate p-value is 0.01 which falls in the rejection region that the original error component is not normally distributed. However, the cube root transformed error series has a p-value of 0.104 which falls in the acceptance region which shows that it has normalized the original series.

The summary of the descriptive statistics of e_t and e_t^* is given in table 3

Table 4: Summary of the descriptive statistics of e_t and e_t^*

Error Component	Mean	Standard deviation	p-value of KS test	Decision	Standard deviation ratio	Variance ratio
e_t	0.998530	0.260433	<0.01	Reject Normality	2.97 to two decimal places	8.82 or 9 to the nearest whole number.
e_t^*	0.991945	0.0877479	0.104	Accept normality		

KS = Kolmogorov - Smirnov

In this practical example the variance ratio is approximately 9.0. This validates our theoretical and simulated results on the variance ratio earlier obtained.

CONCLUSION

In this work, the probability density function of the cube root transformed truncated error component of the multiplicative time series model was obtained.

The mean and variance of the cube root transformed series were also derived.

Comparison was made between the means and variances of the transformed and untransformed series to see whether the transformation maintained the original structure of the series as advised by Osborne (2002). It was found that the means are approximately 1 and the variance of the original series is 9 times the variance of the transformed series for $0 < \sigma \leq 0.22$.

This implies that the error cube root transformation reduced the error variance and that a successful cube root transformation is achieved for the same range of σ.

REFERENCES

Ajibade B. F, Nwosu C. R and Mbegdu J. I. (2015). The Distribution of the Inverse Square Root Transformed Error Component of the Multiplicative Time Series Model. *Journal of Modern Applied Statistical Methods.* (14)2:171 – 199.

Akpanta A. C. and Iwueze I. S (2009). On Applying the Bartlett Transformation Methods to Time Series Data. Journal of Mathematical Sciences (2)3:227 – 243.

Bartlett, M. S. (1947). *The Use of Transformations,* Biometrics, 3, 39-52.

Box, G. E. P. and Cox, D. R. (1964). An Analysis of Transformations. J. Roy. Statistics Soc., B-26, 2 11-243, discussion 244-252.

Chartfield C. (1980). The Analysis of Time Series; an Introduction, Chapman & Hall, London.

Dike A. O., Otuonye E. L. & Chikezie D. C (2016). The nth power Transformation of the Error Component of the Multiplicative Time Series Model. *British Journal of Mathematics and Computer Science* 18(1)

Ibeh G. C and Nwosu C. R. (2013). Study on the Error Component of Multiplicative Time Series Model under Inverse Square Transformation. *American Journal of Mathematics and Statistics* (396): 362 – 374.

Iwueze Iheanyi S. (2007). Some Implications of Truncating the N $(1,\sigma^2)$ Distribution to the left at Zero. *Journal of Applied Sciences*. 7(2): 189-195.

Iwueze I. S, Nwogu E.C., Ohakwe J and Ajaraogu J.C, (2011). New Uses of Buys-Ballot Table. Applied Mathematics, (2): 633 — 645.

Iwueze, I. S., Nwogu, E. C. and Ajaraogu, J. C. (2011). Uses of the Buys-Ballot table in Time Series Analysis. *Journal of Applied Mathematics*, 2, 633 – 645.

Iwueze, I. S. and Nwogu, E. C. (2004). Buys-Ballot for Time Series Decomposition. *Global Journal of Mathematical Science, 3(2): 83 – 89*

Johnson A. C. and Thomopolous N. T.(2004). *Characteristics and Tables of the Partial Expectation of the left-Truncated Normal Distribution*, Midweek Decision Science Institute, Annual conference proceedings.

Nwosu C. R, Iwueze I. S. and Ohakwe J. (2010). Distribution of the Error Term of the Multiplicative Time Series Model under Inverse Transformation. Advances and Applications in Mathematical Sciences. (7)2:119 – 13

Ohakwe, J., Dike, O. A. and Akpanta, A. C (2012). The Implication of Square Root Transformation on a Gamma Distribution Error Component of a Multiplicative Time Series Model, *Proceedings of African Regional Conference on Sustainable Development,* (6)4: 65 – 78.

Ohakwe J, Iwuoha O. and Otuonye E. L (2013). Condition for successful square Transformation in Time series modelling. *Journal of Applied Mathematics* (4):. 680 – 687.

Okereke O. E. and Omekare C. O. (2010) on the Properties of the Truncated Normal Distribution and its Square Root Transformation. *Journal of the Nigerian Statistical Association.* (22):37 – 43.

Osborne, J. (2002). Notes *on the Use of Data Transformations*, J. Practice Practical Assessment, Research &Evaluation, 8(6).

Otuonye E. L, Iwueze I.S. and Ohakwe J. (2011). The Effect of Square Root Transformation on the Error Component of the Multiplicative Time Series Model . *International Journal of Statistics and Systems.* (6)4;(461 – 476)

Uche P. I. (2003). *Probability Theory and Applications*. Lagos Longman Publishers.

Wei, W.W. (1990). *Time Series Analysis*: *Multivariate Methods*. Redwood City, Addison-Wesley Publishing Company