



# Journal of Applied Sciences

ISSN 1812-5654

**science**  
alert

**ANSI***net*  
an open access publisher  
<http://ansinet.com>

## Some Implications of Truncating the $N(1, \sigma^2)$ Distribution to the left at Zero

Iheanyi S. Iwueze

Department of Statistics, Faculty of Biological and Physical Sciences,  
 Abia State University, P.M.B. 2000, Uturu, Abia State, Nigeria

**Abstract:** In this research, we want to present the  $N(1, \sigma^2)$  distribution when its values do not admit values less than or equal to zero. The 0.001 probability limits were used to give practical assurance that the values from  $N(1, \sigma^2)$  are all positive. The truncated values are greater than zero if and only if  $0 < \sigma < 0.30$ . The expression for the mean of the left-truncated  $N(1, \sigma^2)$  distribution is in terms of the cumulative distribution function of the standard normal distribution, while that of the variance is in terms of the Chi-square distribution with one degree of freedom. The logarithmic transform of the left-truncated  $N(1, \sigma^2)$  distribution was found to be approximately normal when  $\sigma < 0.10$ .

**Key words:** Truncation, probability limits, skewness, kurtosis

### INTRODUCTION

Consider a normally distributed random variable  $X$  with a probability density function  $f(x)$  specified as

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(x-1)^2}{2\sigma^2}\right\}, \quad (1)$$

$$-\infty < x < \infty, \sigma^2 > 0$$

In many applications, the random variable  $X$ , which has a  $N(1, \sigma^2)$  distribution do not admit values less than or equal to zero. We must therefore disregard or truncate all values of  $X \leq 0$  to take care of the admissible region of  $X > 0$ .

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 with probability density function  $f^*(x)$  given by Uche (2003).

$$f^*(x) = \begin{cases} 0, & -\infty < x \leq 0 \\ \frac{k}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(x-1)^2}{2\sigma^2}\right\}, & 0 < x < \infty \end{cases} \quad (2)$$

The distribution (2) is said to be truncated to the left at zero,  $k$  is the normalizing quantity and its value is obtained by noting that

$$\int_0^{\infty} f^*(x) dx = 1 \quad (3)$$

Following Hald's conventions (Hald, 1952; Rohatgi, 1976) the resulting distribution (2) that satisfies (3) can be written as

$$f^*(x) = \begin{cases} 0, & -\infty < x \leq 0 \\ \frac{f(x)}{\int_0^{\infty} f(x) dx}, & 0 < x < \infty \end{cases} \quad (4)$$

where  $f(x)$  is as defined in (1). For purposes of generality, Eq. (4) can be re-stated in terms of the standard normal distribution (Cohen, 1991; Johnson and Kotz, 1970; Schneider, 1986). It is important to note that Eq. (2) and (4) are the same and all computations in this research are based on (2).

For a review of prior research on estimation procedures/algorithms of the parameters of the truncated normal distribution (Johnson and Thomopoulos, 2004). Johnson and Thomopoulos (2004) after looking at the characteristics of the truncated population, also provided tables of the cumulative distribution function of the left-truncated normal distribution.

The focus of this research is on the characteristic parameters (mean and standard deviation) of the truncated population with a view to achieving positive values only. The restriction to be placed on  $\sigma$  for the achievement of positive values will be investigated. The  $N(1, \sigma^2)$  assumption is always placed on the residuals of the multiplicative time series models in descriptive time series analysis. The residuals are also assumed to be positive. To linearize the multiplicative model, the logarithmic transformation is taken. A brief discussion of the logarithmic transformation of the left-truncated  $N(1, \sigma^2)$  distribution is also presented.

### MEAN AND VARIANCE OF THE TRUNCATED $N(1, \sigma^2)$ DISTRIBUTION

Using (2), we obtain

$$\int_0^\infty f^*(x)dx = \frac{k}{\sigma\sqrt{2\pi}} \int_0^\infty e^{-\frac{1}{2\sigma^2}(x-1)^2} dx$$

$$dx = \frac{k}{\sigma\sqrt{2\pi}} \int_{\left(\frac{-1}{\sigma}\right)}^\infty \sigma e^{-\frac{z^2}{2}} dz = 1$$

$$\Rightarrow k \int_{\left(\frac{-1}{\sigma}\right)}^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz = 1 \Rightarrow k \Pr\left[Z > \frac{-1}{\sigma}\right]$$

$$= k \left[1 - \Pr\left(Z < \frac{-1}{\sigma}\right)\right] = 1$$

(were,  $Z = \frac{x-1}{\sigma}$ )

$$\Rightarrow k \left[1 - \Phi\left(\frac{-1}{\sigma}\right)\right] = 1, \text{ Thus } k = \frac{1}{1 - \Phi\left(\frac{-1}{\sigma}\right)} \quad (5)$$

Thus,

$$f^*(x) = \begin{cases} 0, & -\infty < x \leq 0 \\ \frac{1}{\sigma\sqrt{2\pi} \left[1 - \Phi\left(\frac{-1}{\sigma}\right)\right]} e^{-\frac{(x-1)^2}{2\sigma^2}}, & 0 < x < \infty \end{cases} \quad (6)$$

Note that

$$R(\sigma) = \int_{\left(\frac{-1}{\sigma}\right)}^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz = 1 - \Phi\left(\frac{-1}{\sigma}\right) \quad (7)$$

and Eq. (6) is a proper p.d.f. The expectation from is

$$E^*(X) = \frac{1}{\sigma\sqrt{2\pi} \left[1 - \Phi\left(\frac{-1}{\sigma}\right)\right]} \int_0^\infty x e^{-\frac{(x-1)^2}{2\sigma^2}} dx$$

$$\frac{1}{\left[1 - \Phi\left(\frac{-1}{\sigma}\right)\right]} \left[ \frac{1}{\sqrt{2\pi}} \int_{\left(\frac{-1}{\sigma}\right)}^\infty e^{-\frac{z^2}{2}} dz + \frac{\sigma}{\sqrt{2\pi}} \int_{\left(\frac{-1}{\sigma}\right)}^\infty z e^{-\frac{z^2}{2}} dz \right]$$

$$\frac{1}{\left[1 - \Phi\left(\frac{-1}{\sigma}\right)\right]} \left[ \left[1 - \Phi\left(\frac{-1}{\sigma}\right)\right] + \frac{\sigma e^{-\frac{-1}{2\sigma^2}}}{\sqrt{2\pi}} \right]$$

Thus,

$$E^*(X) = 1 + \frac{\sigma e^{-\frac{-1}{2\sigma^2}}}{\sqrt{2\pi} \left[1 - \Phi\left(\frac{-1}{\sigma}\right)\right]} = 1 + G(\sigma) \quad (8)$$

where

$$G(\sigma) = \frac{\sigma e^{-\frac{-1}{2\sigma^2}}}{\sqrt{2\pi} \left[1 - \Phi\left(\frac{-1}{\sigma}\right)\right]} \quad (9)$$

Note that,

$$\int_{\left(\frac{-1}{\sigma}\right)}^\infty z e^{-\frac{z^2}{2}} dz = e^{-\frac{-1}{2\sigma^2}}$$

$$E^*(X^2) = \frac{1}{\sigma\sqrt{2\pi} \left[1 - \Phi\left(\frac{-1}{\sigma}\right)\right]} \int_0^\infty x^2 e^{-\frac{(x-1)^2}{2\sigma^2}} dx \quad (10)$$

If we let  $Z = \frac{X-1}{\sigma}$ , we obtain from (6) that

$$E^*(X^2) = \frac{1}{\left[1 - \Phi\left(\frac{-1}{\sigma}\right)\right]} \left\{ \begin{aligned} & \frac{1}{\sqrt{2\pi}} \int_{\left(\frac{-1}{\sigma}\right)}^\infty e^{-\frac{z^2}{2}} dx + \\ & \frac{2\sigma}{\sqrt{2\pi}} \int_{\left(\frac{-1}{\sigma}\right)}^\infty z e^{-\frac{z^2}{2}} dx + \\ & \frac{\sigma^2}{\sqrt{2\pi}} \int_{\left(\frac{-1}{\sigma}\right)}^\infty z^2 e^{-\frac{z^2}{2}} dx \end{aligned} \right\}$$

Using (7) and (10), we obtain

$$E^*(X^2) = \frac{1}{\left[1 - \Phi\left(\frac{-1}{\sigma}\right)\right]} \left\{ \begin{aligned} & \left[1 - \Phi\left(\frac{-1}{\sigma}\right)\right] \\ & + \frac{2\sigma e^{-\frac{-1}{2\sigma^2}}}{\sqrt{2\pi}} \\ & + \frac{\sigma^2}{\sqrt{2\pi}} \int_{\left(\frac{-1}{\sigma}\right)}^\infty z^2 e^{-\frac{z^2}{2}} dz \end{aligned} \right\}$$

$$= 1 + \frac{2\sigma e^{-\frac{-1}{2\sigma^2}}}{\sqrt{2\pi} \left[1 - \Phi\left(\frac{-1}{\sigma}\right)\right]} + \frac{\sigma^2 A}{\sqrt{2\pi} \left[1 - \Phi\left(\frac{-1}{\sigma}\right)\right]} \quad (11)$$

where,

$$A = \int_{\left(\frac{-1}{\sigma}\right)}^{\infty} z^2 e^{\frac{-z^2}{2}} dz = \int_{\left(\frac{-1}{\sigma}\right)}^0 z^2 e^{\frac{-z^2}{2}} dz + \int_0^{\infty} z^2 e^{\frac{-z^2}{2}} dz \quad (12)$$

The following equations are correct:

$$\int_0^{\infty} z^2 e^{\frac{-z^2}{2}} dz = \frac{\sqrt{2\pi}}{2} \quad (13)$$

$$\int_0^{\frac{1}{\sigma^2}} u^{\frac{-1}{2}} e^{\frac{-u}{2}} du = \sqrt{2\pi} \Pr \left[ \chi_{(1)}^2 < \frac{1}{\sigma^2} \right] \quad (14)$$

$$\begin{aligned} \int_0^{\frac{1}{\sigma^2}} u^{\frac{1}{2}} e^{\frac{-u}{2}} du &= \frac{-2e^{\frac{-1}{2\sigma^2}}}{\sigma} + \int_0^{\frac{1}{\sigma^2}} u^{\frac{-1}{2}} e^{\frac{-u}{2}} du \\ &= \frac{-2e^{\frac{-1}{2\sigma^2}}}{\sigma} + \sqrt{2\pi} \Pr \left[ \chi_{(1)}^2 < \frac{1}{\sigma^2} \right] \end{aligned} \quad (15)$$

$$\begin{aligned} \int_{\left(\frac{-1}{\sigma}\right)}^0 z^2 e^{\frac{-z^2}{2}} dz &= \frac{1}{2} \int_0^{\frac{1}{\sigma^2}} u^{\frac{1}{2}} e^{\frac{-u}{2}} du \\ &= \frac{-e^{\frac{-1}{2\sigma^2}}}{\sigma} + \left( \frac{\sqrt{2\pi}}{2} \right) \Pr \left[ \chi_{(1)}^2 < \frac{1}{\sigma^2} \right] \end{aligned} \quad (16)$$

Using (13) and (14), we obtain

$$A = \left( \frac{\sqrt{2\pi}}{2} \right) \left[ 1 + \Pr \left[ \chi_{(1)}^2 < \frac{1}{\sigma^2} \right] \right] - \frac{e^{\frac{-1}{2\sigma^2}}}{\sigma} \quad (17)$$

$$D(\sigma) = \Pr \left[ \chi_{(1)}^2 < \frac{1}{\sigma^2} \right] \quad (18)$$

Using (8) and (11), we obtain

$$\begin{aligned} \text{var}^*(X) &= E^*(X^2) - [E^*(X)]^2 \\ &= \frac{\sigma^2 A}{\sqrt{2\pi} \left[ 1 - \Phi \left( \frac{-1}{\sigma} \right) \right]} - \left[ \frac{\sigma e^{\frac{-1}{2\sigma^2}}}{\sqrt{2\pi} \left[ 1 - \Phi \left( \frac{-1}{\sigma} \right) \right]} \right]^2 \end{aligned} \quad (19)$$

Finally, substituting (17) into (18), we obtain

$$\text{var}^*(X) = \frac{\sigma^2}{2 \left[ 1 - \Phi \left( \frac{-1}{\sigma} \right) \right]} \left[ 1 + \Pr \left[ \chi_{(1)}^2 < \frac{1}{\sigma^2} \right] \right]$$

$$- \frac{\sigma e^{\frac{-1}{2\sigma^2}}}{\sqrt{2\pi} \left[ 1 - \Phi \left( \frac{-1}{\sigma} \right) \right]} - \left[ \frac{\sigma e^{\frac{-1}{2\sigma^2}}}{\sqrt{2\pi} \left[ 1 - \Phi \left( \frac{-1}{\sigma} \right) \right]} \right]^2 \quad (20)$$

Our computations have shown that the expression for the mean of the left-truncated  $N(1, \sigma^2)$  distribution is in terms of the cumulative distribution function (c.d.f.) of the standard normal Eq. (8), while that of the variance is in terms of the cumulative distribution function of the standard normal and the c.d.f. of the Chi-Square distribution with one degree of freedom Eq. 20. Our results for the variance agree with those of Barr and Sherill (1999).

**PROPERTIES OF THE MEAN AND VARIANCE OF THE TRUNCATED  $N(1, \sigma^2)$  Distribution**

In most applications, the requirements needed are (i)  $f^*(x), x < 0$ , (ii)  $E^*(X) = E(X)$  and  $\text{var}^*(X) = \text{var}(X)$ . Our derivations have shown that all these depend on the parameter  $\sigma$ . To investigate these requirements, we look for the values of  $\sigma$  for which (i)  $f^*(0) \approx 0$ , (ii)  $E^*(X) - E(X) \approx 0$  and (iii)  $\text{var}(X) - \text{var}^*(X) \approx 0$ .

Figure 1 gives the shape of the distribution curve for some values of  $\sigma$ . The two curves for the original (non-truncated) and the truncated populations coincide for some values of  $\sigma$ , while for some values of  $\sigma$  the truncated curve do not admit values greater than zero when they are approximately the same. Generally, the truncated values are always greater or equal to the non-truncated values for all values of  $\sigma$ . We would, therefore, expect the two stochastic variables to behave alike in some interval  $\sigma \in (0, b)$ . The quantity  $b$  will be determined by use of the three requirements stated above. Our simulation results

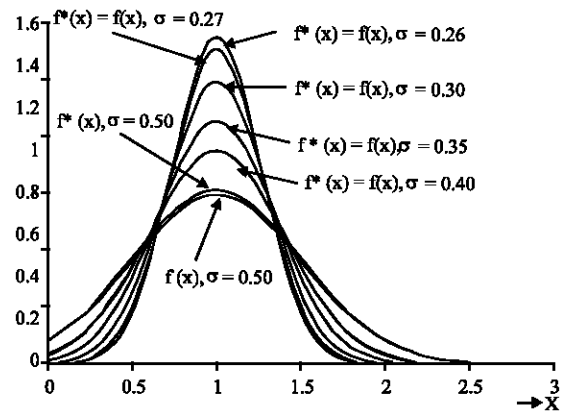


Fig. 1: The curve shapes of the left-truncated  $N(1, \sigma^2)$  distribution

Table 1: Summary of results for  $f^*(0)$ ,  $E^*(X) - E(X)$ ,  $\text{var}(X) - \text{var}^*(X)$

No. of decimal places	$f^*(0) \approx 0$	$E^*(X) - E(X) \approx 0$	$\text{var}(X) - \text{var}^*(X) \approx 0$	$y_0 \approx 0$
4	$\sigma \leq 0.21$	$\sigma \leq 0.25$	$\sigma \leq 0.25$	N/A
3	$\sigma \leq 0.24$	$\sigma \leq 0.29$	$\sigma \leq 0.29$	$\sigma \leq 0.02$
2	$\sigma \leq 0.29$	$\sigma \leq 0.38$	$\sigma \leq 0.38$	$\sigma \leq 0.07$
1	$\sigma \leq 0.40$	$\sigma \leq 0.56$	$\sigma \leq 0.56$	$\sigma \leq 0.23$

(all simulations in this paper were performed using MINITAB) will help determine the level of accuracy required for the two stochastic variables to behave alike in a prescribed interval.

Using Eq. 5, 8 and 20, results were obtained for  $0.01 \leq \sigma \leq 2.0$  and the results obtained are summarized in Table 1.

From Table 1 and using the three requirements stipulated, it is clear that the desired interval is the interval  $\sigma < 0.30$ . It follows from our analysis that the 0.001 limits may be used to give practical assurance that the truncated values from the  $N(1, \sigma^2)$  are all positive. The 0.001 probability limits (99.9% confidence limits) are practically equivalent to  $3.2905\sigma$  limits and for the lower confidence limit to be positive we have  $1 - 3.2905\sigma > 0$ . This implies that  $\sigma < 0.30$  with the truncated values lying between 0.01 and 1.99 ( $0.01 < X < 1.99$ ). It is also clear from this analysis that the admissible region,  $\sigma < 0.30$ , is needed to ensure that the mean and variance of the truncated distribution remains, respectively, at the assumed values of 1 and  $\sigma^2$ .

**TIME SERIES APPLICATION**

A general multiplicative model for descriptive time series analysis would be of the form

$$X_t = T_t S_t C_t I_t, \quad t = 1, 2, 3, \dots, n \tag{21}$$

where, for time  $t$ ,  $X_t$  denotes the observed value of the series,  $T_t$  is the trend,  $S_t$  the seasonal term,  $C_t$  the cyclic term and  $I_t$  is the irregular or random component of the series. The model (21) will be regarded to be adequate when the irregular component is purely random. If short period of time are involved, the cyclical component is superimposed into the trend (Chartfield, 1980), p.13 and we obtain a trend-cycle component denoted by  $M_t$ . For our purposes,  $I_t$  to be denoted by  $e_t$  are independent, identically distributed normal errors with mean 1 and variance  $\sigma_1^2 > 1 (I_t = e_t \sim N(1, \sigma_1^2))$ . Model (21) can then be written as

$$X_t = M_t S_t e_t \tag{22}$$

Descriptive modeling of the multiplicative time series model (22) generally involve the logarithmic transformation. Upon linearization of (22), we obtain

$$\ln X_t = \ln M_t + \ln S_t + \ln e_t \tag{23}$$

or

$$Y_t = M^*_t + S^*_t + e^*_t \tag{24}$$

where  $Y_t = \ln X_t, M^*_t = \ln M_t, S^*_t = \ln S_t, e^*_t = \ln e_t$ . It is important to note that (24) is the additive time series model. For model (24), the variant assumption on the random component  $e^*_t$  is that they are independent, identically distributed normal errors with mean zero and variance  $\sigma_2^2 > 0 (e^*_t \sim N(0, \sigma_2^2))$ .

It is evident that for the left-truncated  $N(1, \sigma^2)$  distribution given by (6), we need to investigate the characteristics of

$$Y = \log_e X \tag{25}$$

The probability density function of  $Y$  is given by

$$f(y) = f(x) \left| \frac{dx}{dy} \right|. \text{ Thus,}$$

$$f(y) = \frac{e^y e^{-\frac{(e^y - 1)^2}{2\sigma^2}}}{\sigma \sqrt{2\pi} \left[ 1 - \Phi\left(\frac{-1}{\sigma}\right) \right]}, \quad -\infty < y < \infty \tag{26}$$

Equation (26) is a proper probability density function and its curve shapes for some values of  $\sigma$  are given in Fig. 2.

It is clear from Fig. 2 that the distribution of  $Y = \log_e X$  is skewed to the left with some symmetric and bell-shaped conditions for small values of  $\sigma$ . For all values of  $\sigma$ , the curve  $f(y)$  has one maximum point ( $y_0$ ) and one maximum value ( $f(y_0)$ ). To obtain the maximum point for a given value of  $\sigma$ , we consider  $f'(y) = 0$ . Now,

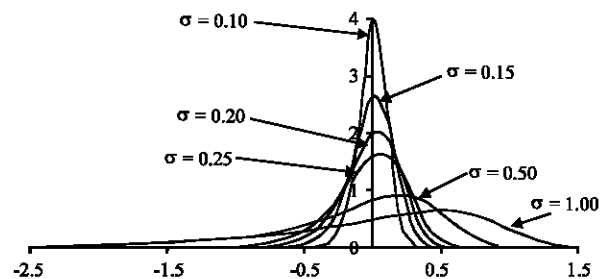


Fig. 2: Curve shapes of the logarithm of the left-truncated

$$f'(y) = \frac{1}{\sigma\sqrt{2\pi}\left[1 - \Phi\left(\frac{-1}{\sigma}\right)\right]} e^y e^{-\frac{(e^y - 1)^2}{2\sigma^2}} \left[1 - \frac{(e^{2y} - e^y)}{\sigma^2}\right] \quad (27)$$

Equating  $f'(y) = 0$ , we obtain

$$\left[1 - \frac{(e^{2y} - e^y)}{\sigma^2}\right] = 0 \Rightarrow e^{2y} - e^y - \sigma^2 = 0 \quad (28)$$

If we let  $w = e^y$ , we obtain

$$w^2 - w - \sigma^2 = 0 \quad (29)$$

That is,

$$w = \frac{1 \pm \sqrt{1 + 4\sigma^2}}{2}$$

Note that

$$\frac{1 - \sqrt{1 + 4\sigma^2}}{2} < 0$$

and

$$\frac{1 + \sqrt{1 + 4\sigma^2}}{2} > 0$$

Hence

$$e^{y_0} = \frac{1 + \sqrt{1 + 4\sigma^2}}{2}$$

and

$$y_0 = \log_e \left[ \frac{1 + \sqrt{1 + 4\sigma^2}}{2} \right] \quad (30)$$

For our purposes (investigation of bell-shaped conditions),  $y_0 \approx 0$ . Using the number of significant decimal places to determine the level of accuracy required, we obtain the summary listed in the last column of

Table 1. We use 2 significant decimal places as the level of accuracy required to obtain that  $\sigma \leq 0.07$  for symmetry and bell-shaped conditions. However, if and when the accuracy requirement is 1 significant decimal place, we use  $\sigma \leq 0.23$ .

Bell-shaped frequency distributions will have

$$\gamma_1 = \left( \frac{\mu_3^2}{\mu_2^3} \right)^{1/2} = 0 \quad (31)$$

and

$$\gamma_2 = \frac{\mu_4}{\mu_2^2} - 3 = 0 \quad (32)$$

where

$$\mu_k = E[(Y - E(Y))^k] = \int_{-\infty}^{\infty} (y - \mu_y)^k f(y) dy \quad (33)$$

is the  $k^{\text{th}}$  moment about the mean  $\mu_y$  of  $Y = \log_e X$ . We must recall at this point that  $\gamma_1$  is a measure of skewness and  $\gamma_2$  is a measure of kurtosis. The solution of (33) cannot be solved easily with an analytic approach, we therefore resort to the Monte Carlo method.

Sample estimates of the moments about the mean are given by

$$\hat{\mu}_k = M_k = \frac{\sum_{i=1}^n (Y_i - \bar{Y})^k}{n}, \quad k = 1, 2, 3, \dots \quad (34)$$

where  $n$  is the sample size. Artificial data was generated from the  $N(1, \sigma^2)$  density function for the variable  $X$  and the  $X$  values obtained were subsequently used to generate from the model  $Y = \log_e X$ , where  $0.07 \leq \sigma \leq 0.23$ . For each configuration of  $(n = 100, 0.07 \leq \sigma \leq 0.23)$ , 1000 replications were performed. For want of space, we show the results of the first 30 replications for two configurations:  $(n = 100, \sigma \leq 0.07)$  in Table 2 and  $(n = 100, \sigma \leq 0.23)$  in Table 3.

We now consider the significance of skewness and kurtosis for samples of size  $n = 100$ . We must note that  $\gamma_1$  and  $\gamma_2$  measure nonnormality on the basis of skewness and kurtosis of the distribution, respectively. The 0.05 and 0.01 points of the sampling distributions of  $\gamma_1$  and  $\gamma_2$  are given in Duncan (1974) as Tables F and G, respectively. If a given sample has a value of  $\gamma_1$  and  $\gamma_2$  beyond the 0.05 point (a stricter test would use the 0.01 points) for that statistics the population may be deemed to be nonnormal. For samples of  $n = 100$ , the upper 0.05 point for  $\gamma_1$  is 0.389

**Table 2: Simulation results for skewness ( $\gamma_1$ ) and ( $\gamma_2$ ) kurtosis when  $\sigma = 0.07$**

X ~ N(1, $\sigma^2$ ), $\sigma = 0.07$						Y = log X, X ~ N(1, $\sigma^2$ ), $\sigma = 0.07$						% increase in $\sigma$
Mean	STD	Min.	Max.	$\gamma_1$	$\gamma_2$	Mean	STD	Min.	Max.	$\gamma_1$	$\gamma_2$	
1.00	0.07	0.8165	1.1621	0.030	-0.024	-0.0024	0.0705	-0.2028	0.1503	0.232	-0.108	0.70
1.00	0.07	0.8228	1.1981	0.118	-0.024	-0.0025	0.0707	-0.1951	0.1807	0.316	-0.143	0.97
1.00	0.07	0.8347	1.1682	0.059	-0.024	-0.0025	0.0706	-0.1807	0.1554	0.260	-0.120	0.81
1.00	0.07	0.8242	1.1822	0.089	-0.024	-0.0025	0.0707	-0.1933	0.1674	0.288	-0.132	0.93
1.00	0.07	0.8263	1.2043	0.118	-0.024	-0.0024	0.0701	-0.1908	0.1859	0.088	-0.049	0.17
1.00	0.07	0.8240	1.1913	0.030	-0.024	-0.0025	0.0705	-0.1936	0.1751	0.260	0.292	0.77
1.00	0.07	0.7995	1.1801	0.030	-0.024	-0.0025	0.0705	-0.2238	0.1656	0.232	0.305	0.70
1.00	0.07	0.8109	1.1848	0.118	-0.024	-0.0024	0.0702	-0.2096	0.1696	0.117	0.359	0.26
1.00	0.07	0.7956	1.1696	0.030	-0.024	-0.0025	0.0703	-0.2287	0.1566	0.174	-0.085	0.46
1.00	0.07	0.8444	1.1836	0.207	-0.024	-0.0024	0.0699	-0.1691	0.1686	0.004	-0.012	-0.11
1.00	0.07	0.8279	1.2081	0.118	-0.024	-0.0024	0.0702	-0.1888	0.1890	0.118	-0.049	0.23
1.00	0.07	0.8366	1.1706	0.059	-0.024	-0.0024	0.0706	-0.1784	0.1575	0.260	-0.120	0.79
1.00	0.07	0.8154	1.1736	0.059	-0.024	-0.0025	0.0705	-0.2040	0.1601	0.232	-0.108	0.76
1.00	0.07	0.8133	1.1782	0.118	-0.024	-0.0024	0.0701	-0.2067	0.1640	0.088	-0.049	0.17
1.00	0.07	0.8272	1.1780	0.030	-0.449	-0.0024	0.0704	-0.1898	0.1638	0.204	-0.096	0.63
1.00	0.07	0.8067	1.1658	0.089	0.401	-0.0025	0.0707	-0.2149	0.1534	0.317	0.278	0.96
1.00	0.07	0.8142	1.1671	0.030	-0.036	-0.0024	0.0704	-0.2056	0.1545	0.204	0.318	0.57
1.00	0.07	0.8216	1.1891	0.003	-0.024	-0.0024	0.0704	-0.1965	0.1732	0.204	-0.096	0.57
1.00	0.07	0.8071	1.1622	0.059	-0.024	-0.0025	0.0706	-0.2144	0.1503	0.260	-0.120	0.83
1.00	0.07	0.8299	1.2123	0.059	-0.024	-0.0024	0.0703	-0.1865	0.1926	0.146	-0.073	0.41
1.00	0.07	0.8345	1.1882	0.014	-0.024	-0.0024	0.0704	-0.1809	0.1725	0.204	-0.096	0.56
1.00	0.07	0.8253	1.1684	0.089	-0.024	-0.0024	0.0702	-0.1920	0.1556	0.117	-0.061	0.33
1.00	0.07	0.8211	1.1861	0.059	-0.024	-0.0024	0.0703	-0.1971	0.1707	0.117	-0.073	0.36
1.00	0.07	0.8049	1.1592	0.148	-0.024	-0.0025	0.0708	-0.2170	0.1478	0.372	0.226	1.16
1.00	0.07	0.8364	1.1957	0.089	-0.024	-0.0024	0.0702	-0.1786	0.1787	0.117	-0.061	0.27
1.00	0.07	0.8154	1.1736	0.059	-0.024	-0.0025	0.0705	-0.2040	0.1601	0.232	-0.108	0.76
1.00	0.07	0.8216	1.1891	0.003	-0.024	-0.0024	0.0704	-0.1965	0.1732	0.204	-0.096	0.57
1.00	0.07	0.8299	1.2123	0.059	-0.024	-0.0024	0.0703	-0.1865	0.1926	0.146	-0.073	0.41
1.00	0.07	0.8343	1.1882	0.014	-0.024	-0.0024	0.0704	-0.1809	0.1725	0.204	-0.096	0.56
1.00	0.07	0.8253	1.1684	0.089	-0.024	-0.0024	0.0702	-0.1920	0.1556	0.117	-0.061	0.33

**Table 3: Simulation results for skewness ( $\gamma_1$ ) and ( $\gamma_2$ ) kurtosis when  $\sigma = 0.23$**

X ~ N(1, $\sigma^2$ ), $\sigma = 0.23$						Y = log X, X ~ N(1, $\sigma^2$ ), $\sigma = 0.23$						% increase in $\sigma$
Mean	STD	Min.	Max.	$\gamma_1$	$\gamma_2$	Mean	STD	Min.	Max.	$\gamma_1$	$\gamma_2$	
1.00	0.23	0.3970	1.5327	0.032	-0.104	-0.0288	0.2495	-0.9239	0.4270	0.861	1.362	8.48
1.00	0.23	0.4177	1.6508	0.109	-0.026	-0.0291	0.2511	-0.8730	0.5012	0.874	0.996	9.17
1.00	0.23	0.4569	1.5526	0.062	-0.112	-0.0289	0.2495	-0.7834	0.4399	0.825	0.945	8.48
1.00	0.23	0.4225	1.5986	0.093	-0.166	-0.0290	0.2504	-0.8615	0.4691	0.854	1.053	8.87
1.00	0.23	0.4292	1.6714	0.119	-0.053	-0.0279	0.2433	-0.8458	0.5137	0.618	0.649	5.78
1.00	0.23	0.4218	1.6286	0.037	0.104	-0.0289	0.2494	-0.8633	0.4877	0.820	0.911	8.43
1.00	0.23	0.3413	1.5917	0.026	-0.015	-0.0290	0.2512	-1.0751	0.4648	0.986	2.299	9.22
1.00	0.23	0.3786	1.6072	0.114	0.149	-0.0283	0.2467	-0.9712	0.4745	0.824	1.750	7.26
1.00	0.23	0.3284	1.5571	0.038	-0.148	-0.0287	0.2488	-1.1134	0.4428	0.929	2.455	8.17
1.00	0.23	0.4889	1.6033	0.195	-0.048	-0.0275	0.2406	-0.7156	0.4721	0.514	0.380	4.60
1.00	0.23	0.4346	1.6837	0.116	0.172	-0.0281	0.2447	-0.8333	0.5210	0.703	0.968	6.39
1.00	0.23	0.4628	1.5606	0.065	-0.173	-0.0288	0.2492	-0.7705	0.4450	0.797	0.757	8.35
1.00	0.23	0.3936	1.5704	0.051	-0.196	-0.0287	0.2485	-0.3658	0.4514	0.789	0.917	8.04
1.00	0.23	0.3864	1.5856	0.116	-0.094	-0.0280	0.2442	-0.9508	0.4610	0.684	1.091	6.17
1.00	0.23	0.4321	1.5848	0.023	-0.254	-0.0285	0.2467	-0.8392	0.4605	0.702	0.516	7.26
1.00	0.23	0.3647	1.5449	0.081	0.205	-0.0294	0.2535	-1.0086	0.4350	1.029	1.969	10.22
1.00	0.23	0.3896	1.5486	0.028	0.045	-0.0286	0.2480	-0.9427	0.4373	0.818	1.298	7.83
1.00	0.23	0.4139	1.6212	0.003	-0.090	-0.0285	0.2469	-0.8821	0.4832	0.731	0.733	7.35
1.00	0.23	0.3661	1.5329	0.064	-0.099	-0.0290	0.2507	-1.0050	0.4272	0.909	1.578	9.00
1.00	0.23	0.4410	1.6977	0.053	0.030	-0.0283	0.2454	-0.8187	0.5293	0.674	0.541	6.70
1.00	0.23	0.4562	1.6185	0.014	0.001	-0.0285	0.2471	-0.7847	0.4815	0.742	0.743	7.43
1.00	0.23	0.4260	1.5533	0.094	0.011	-0.0282	0.2461	-0.8533	0.4404	0.765	1.337	7.00
1.00	0.23	0.4122	1.6116	0.073	-0.072	-0.0282	0.2448	-0.8863	0.4772	0.662	0.676	6.43
1.00	0.23	0.3590	1.5232	0.141	0.173	-0.0298	0.2565	-1.0244	0.4208	1.156	2.592	11.52
1.00	0.23	0.4626	1.6429	0.095	-0.008	-0.0281	0.2441	-0.7710	0.4965	0.629	0.494	6.13
1.00	0.23	0.3936	1.5704	0.051	-0.196	-0.0287	0.2485	-0.9324	0.4514	0.789	0.917	8.04
1.00	0.23	0.4139	1.6212	0.003	-0.090	-0.0285	0.2469	-0.8821	0.4832	0.731	0.733	7.35
1.00	0.23	0.4410	1.6977	0.053	0.030	-0.0283	0.2454	-0.8187	0.5293	0.674	0.541	6.70
1.00	0.23	0.4562	1.6185	0.014	0.001	-0.0285	0.2471	-0.7847	0.4815	0.742	0.743	7.43
1.00	0.23	0.4261	1.5529	0.094	-0.010	-0.0283	0.2460	-0.8530	0.4402	0.764	1.328	6.96

and the upper 0.01 point is 0.567. Equivalently, the upper 0.05 point for  $\gamma_2$  is -0.65, for  $\gamma_2 < 0$  and 0.77 for  $\gamma_2 > 0$ ; while the upper 0.01 point is -0.82 for  $\gamma_2 < 0$  and 1.39 for  $\gamma_2 > 0$ . It will be noted that a population distribution can have  $\gamma_1 = 0$  or  $\gamma_2 = 0$  without their being normal (Shewart, 1931). Sample values close to these normal values do not therefore prove normality; they merely do not lead to the rejection of the normal hypothesis.

It is clear from Table 2 that when  $\sigma = 0.07$ , both  $X$  and  $Y = \log_e X$  can be assumed to be normally distributed with  $\text{var}(X) \approx \text{var}(Y)$ . On the other hand, we infer from Table 3 that when  $\sigma = 0.23$ ,  $Y = \log_e X$ , cannot be assumed to be normal on the basis of the sample values of  $\gamma_1$  and/or  $\gamma_2$ . Our search for normality of  $Y = \log_e X$  for  $X \sim N(1, \sigma^2)$ ,  $0.07 \leq \sigma \leq 0.23$  showed that the sample value of  $\gamma_1$  indicates departure from normality from  $\sigma = 0.10$ . We conclude that when  $X \sim N(1, \sigma^2)$ , the logarithmic transform  $Y = \log_e X$  can be assumed to be normally distributed with mean zero and variance  $\sigma^2$ , provided  $\sigma < 0.10$ .

### CONCLUSION

In this research, we have examined some implications of truncating the  $N(1, \sigma^2)$  distribution to the left at zero. We observed that the truncated values are always greater or equal to the non-truncated values for all values of  $\sigma$ . However, the two stochastic variables behave alike in the interval  $\sigma < 0.30$ . It follows from our 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  $\sigma^2$ .

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  $\sigma < 0.10$ .

The results of our study suggests that the logarithmic transforms of the normal distribution are

worthy of future research. Part of this research must include the left-truncated  $N(1, \sigma^2)$  distribution which has important applications in descriptive time series studies. Further work is needed to compute analytically the moments about the mean of the logarithmic transform. However, our results suggests one should exercise extreme caution with respect to the value of  $\sigma$  when linearizing multiplicative models in descriptive time series analysis.

### REFERENCES

- Barr, D.R. and E.T. Sherill, 1999. Mean and variance of truncated normal distributions. *American Statistician*, 53: 357-361.
- Chartfield, C., 1980. *The Analysis of Time Series: An Introduction*, Chapman and Hall, London.
- Cohen, A.C., 1991. *Truncated and Censored Samples*, Marcel Dekker, New York.
- Duncan, A.J., 1974. *Quality Control and Industrial Statistics*, Richard D. Irwin; Inc, Illinois.
- Hald, A., 1952. *Statistical Theory with Engineering Applications*. John Wiley and Sons, New York.
- Johnson, N.L. and S. Kotz, 1970. *Distributions in Statistics: Continuous Univariate Distributions I*, Houghton Mifflin, New York.
- Johnson, A.C. and N.T. Thomopoulos, 2004. Characteristics and tables of the left-truncated normal distribution. *Annual Conference Proceedings of the Midwest Decision Sciences Institute*, April 2004, Cleveland.
- Rohatgi, V.K., 1976. *An Introduction to Probability Theory and Mathematical Statistics*, John Wiley and Sons, New York.
- Schneider, H., 1986. *Truncated and Censored Samples From Normal Population*, Marcel Dekker, New York.
- Shewart, W.A., 1931. *Economic Control of Quality of Manufactured Product*, D. Van Nostrand Co; Inc, New York.
- Uche, P.I., 2003. *Probability: Theory and Applications*, Longman, Lagos.