Journal of Mathematical Problems, Equations and Statistics

E-ISSN: 2709-9407
P-ISSN: 2709-9393
JMPES 2025; 6(1): 01-05

www.mathematicaljournal.com
Received: 05-11-2024
Accepted: 11-12-2024

Shaban Juma Ally
1) Department of Statistics, Andhra University, Andhra Pradesh, India
2) College of African Wildlife Management - MWEKA, Kilimanjaro, Tanzania

Hassan A Mrutu
College of African Wildlife Management, MWEKA, Kilimanjaro, Tanzania

Rehema Reuben Mathew
Moshi Co-operative University, Kilimanjaro, Tanzania

Corresponding Author:
Shaban Juma Ally
1) Department of Statistics, Andhra University, Andhra Pradesh, India
2) College of African Wildlife Management - MWEKA, Kilimanjaro, Tanzania

Geometric approach to the skewed log-normal distribution

Shaban Juma Ally, Hassan A Mrutu and Rehema Reuben Mathew

DOI: https://doi.org/10.22271/math.2025.v6.i1a.164

Abstract
This paper presents a geometric approach to the skewed log-normal distribution, by making derivations of the key results on its mean, variance, and skewness through geometric transformations. Theorem 1 and proposition 1 offer new insights into the distribution's properties and skewness, enhancing its interpretability by adding clarity and the results be informative.

Keywords: Skewed log-normal distribution, geometric skewness, asymmetry analysis and geometric transformations

1. Introduction

The log-normal distribution is a common choice to model positive skewed data; that is, a variable whose logarithm follows a normal distribution. Dynamically, this distribution is particularly adept with many of the more commonly used classes of a skewed dataset yet, can still leave out some of the more diverse forms of asymmetrical shapes found in real-world data. To address this, the skewed log-normal distribution was proposed, introducing a skewness parameter δ to allow for asymmetry in either direction. The probability density function (PDF) for the skewed log-normal distribution is:

$$f_X(x_j\mu,\sigma,\delta) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{\log(x)-\mu}{\sigma}\right)^2\right)\left(1+\delta\cdot\frac{x-\mu}{\sigma}\right)$$

The study explores a geometric approach to the skewed log-normal distribution, which gives a clear way for understanding the effects of skewness to the shape of the distribution. The skewness, under the geometric skewness S_{geo}, is obtained by using the formula:

$$S_{geo} = \frac{\sum_{i=1}^{n}(x_i-\overline{x})^3}{n\cdot\sigma^3}$$

This methodology improves the flexibility of the skewed log-normal distribution, allowing it to more accurately reflect real-world data in fields like finance (Gurland & Tripathi, 1971) [2], medicine (Azzalini, 1985) [1], and environmental science (Kato, 2012) [3]. Making derivations on moments and skewness, can help in providing the demonstration of the practical utility of the skewed log-normal model in various applications (McKay *et al.*, 2000; O'Hagan & Forster, 2004) [4, 5].

2. Geometric Framework

2.1 Geometric Transformations

Under this transformation the scaling techniques applied to overcome skewness in the skewed lognormal distribution. When used in normalization of the data, scaling increase the skewness by changing variance. The results of this transformation reveal skewness as a geometric property and allow for a clear presentation of how the distribution deviates from symmetry. The geometric method provides a transparent framework for measuring and also interpreting skewness with respect to the skewed log-normal distribution. The geometric skewness (γ_g) for a distribution can be expressed as:

$$\gamma_g = \frac{\sum_{i=1}^{n}(x_i - \mu)^3}{n\sigma^3}$$

where x_i represents the data points, μ is the mean, and σ is the standard deviation. This formula reflects the third moment of the distribution after scaling, capturing the skewness in a way that accounts for the geometric structure of the data. The geometric skewness helps quantify the degree of asymmetry by examining how the data stretches or compresses, offering a more interpretable and visually intuitive measure than traditional skewness metrics.

3. Theorems and Propositions

3.1 Theorem 1

3.1.1 Skewness of a Log-Normal Distribution

The Log-Normal Distribution is defined for a random variable X such that Y=log(X) ~ N(μ, σ^2) where Y is a normally distributed variable with mean (μ) and variance (σ^2)

The probability density function (pdf) for a Log-Normal distribution is given by:

$$f_X(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{(\log(x)-\mu)^2}{2\sigma^2}\right), x > 0$$

From the pdf above, there some formula for Log normal distribution

Moments of Log-Normal Distribution

The skewness is defined as the normalized third central moment:

$$\gamma_1(X) = \frac{E[(X-\mu_X)^3]}{(E[(X-\mu_X)^2])^{3/2}}$$

where μ_X is the mean of X.

Mean, Variance, and Skewness of Log-Normal Distribution Since X=e^Y ,

Mean of X

$$E[X] = E[e^Y] = e^{\mu+\frac{\sigma^2}{2}}$$

Variance of X

$$Var(X) = E[X^2] - (E[X])^2$$

$$Var(X) = e^2 e^{2\mu+\sigma^2}\ e^2\ + \sigma^2 = e^{2\mu+\sigma^2}(e^{\sigma^2} - 1)$$

Skewness of X

The skewness is derived from the third central moment.

The third moment $E[(X - \mu x)^3]$ for a Log-Normal distribution can be derived as follows:

$$\gamma_1(X) = \frac{E[(X-\mu_X)^3]}{(E[(X-\mu_X)^2])^{3/2}}$$

The skewness of the Log-Normal distribution is known (based on its third central moment and distribution properties)

$$\gamma_1(X) = \frac{e^{\sigma^2} + 2e^{\mu+\frac{\sigma^2}{2}} + e^{2\mu+\sigma^2} - 3}{(e^{\sigma^2} - 1)^{3/2}}$$

The skewness depends on the mean (μ) and variance (σ^2) of the underlying normal distribution.

3.1.2 Geometric Transformation of the Distribution

Under this transformation is where the use of Scaling and Shifting techniques introduced to make some derivations by introducing some parameters in order improve traditional one.

Define the Scaling Transformation

Let the original random variable be X follow a Log-Normal distribution:

$$Y = \log(X) \sim N(\mu, \sigma^2), X = e^Y$$

After defining, applied a scaling transformation:

$$cX, c>0$$

The new random variable X' is also Log-Normally distributed because:

$$\log(X') = \log(cX) = \log(c) + \log(X) = \log(c) + Y$$

since $Y \sim N(\mu, \sigma^2)$, adding a constant log(c) shifts the mean of the normal distribution but does not affect its variance. Therefore,

$$\log(X') \sim N(\mu + \log(c), \sigma^2)$$

After definition the Moments of the Transformed Distribution

The moments of X' after applying the Scaling Transformation is derived as follows:

1st Mean of X'

$$E[X'] = E[cX] = c \cdot E[X]$$

since $E[X] = e^{\mu+\frac{\sigma^2}{2}}, we\ have$:

$$E[X'] = c \cdot e^{\mu+\frac{\sigma^2}{2}}$$

2nd**: Variance of X'**: Variance scales quadratically with c. Let Var(X) =$e^{2\mu+\sigma^2}(e^{\sigma^2} - 1)$ then:

$$Var(X') = c^2 \cdot Var(X) = c^2 \cdot e^{2\mu+\sigma^2}(e^{\sigma^2} - 1)$$

3rd: Third Central Moment of X′: The third central moment scales cubically with c.

Let the third central moment of X be $\mu_3(X)$ Then:

$$\mu_3(X') = c^3 \cdot \mu_3(X)$$

$where\ \mu_3(X)\ is\ given\ by$:

$$\mu_3(X) = e^{3\mu+\frac{9\sigma^2}{2}} - 3e^{2\mu+\frac{2\sigma^2}{2}}\ e^{\mu+\frac{\sigma^2}{2}} + 2e^{3\mu+\frac{\sigma^2}{2}}$$

4th: Skewness of X′

The skewness of X′ is defined as:

$$\gamma_1(X') = \frac{\mu_3(X')}{Var(X')^{3/2}}$$ substitute

$$\mu_3(X') = c^3 \cdot \mu_3(X)\ and\ Var(X') = c^2 \cdot Var(X):$$

$$\gamma_1(X') = \frac{c^3 \cdot \mu_3(X)}{(c^2 \cdot Var(X))^{3/2}}$$

Simplify the denominator

$(c^2 \cdot Var(X))^{3/2} = c^3 \cdot (Var(X))^{3/2}$

Cancel c^3 in the numerator and denominator:

$\gamma_1(X') = \gamma_1(X)$

Therefore, the skewness of X' is scaled by c^3:

$\gamma_1(X') = \gamma_1(X) \cdot c^3$

3.1.3 The Theorem

This part shows the genesis of the formula for skewed Log-Normal distribution, there are some computations procedure followed to obtain the final formula under this theorem.

1st step: Formulate the Problem

Let X follow a skewed Log-Normal distribution with skewness $\gamma_1(X)$. Apply a geometric transformation X′= cX, where c>0 is a scaling factor. Then the derivations give clear understanding on how the skewness of X′, denoted

$\gamma_1(X')$, relates to $\gamma_1(X)$.

2nd Step: Skewness of the Transformed Variable

The skewness of X′ is defined as:

$$\gamma_1(X') = \frac{\mu_3(X')}{(V_{ar}(X'))^{3/2}}$$

From the moments derived earlier

$$\mu_3(X') = c^3 \cdot \mu_3(X) \qquad (1)$$

$$Var(X') = c^2 \cdot Var(X) \qquad (2)$$

Substitute these into the skewness formula:

$$\gamma_1(X') = \frac{c^3 \cdot \mu_3(X)}{(c^2 \cdot Var(X))^{3/2}}$$

Simplify the denominator

$$(c^2 \cdot Var(X))^{3/2} = c^3 \cdot (Var(X))^{3/2}$$

Cancel c^3 in the numerator and denominator:

$$\gamma_1(X') = \gamma_1(X)$$

3rd Step: Final Theorem Statement

Theorem 1

If X follows a skewed Log-Normal distribution with skewness $\gamma_1(X)$, then for any geometric transformation of the form X′= cX (where c>0), the skewness of the transformed distribution is related to the skewness of the original distribution as:

$$\gamma_1(X') = \gamma_1(X) \cdot c^3$$

3.2 Proposition 1: Geometric Skewness Measure

3.2.1 Traditional Skewness (Pearson's Third Coefficient of Skewness)

The traditional skewness measure (Pearson's third coefficient of skewness) is based on the third central moment of the distribution, normalized by the cube of the standard deviation. It is defined as:

$$\gamma_1 = \frac{E[(X-\mu)^3]}{(\sigma)^3}$$

Where:

E is the expectation operator,
X is the random variable (data point),
$\mu = E[X]$ is the mean of the distribution,
$\sigma = \sqrt{E[(X-\mu)^2]}$ is the standard deviation of the distribution

Step-by-Step Derivation of Traditional Skewness**:**

1st Mean (μ)

The mean is the first central moment of the distribution, given by:

$$\mu = E[X] = \frac{1}{n}\sum_{i=1}^{n} x_i$$

2nd The Variance (σ^2)

The variance is the second central moment, which measures the spread of the data around the mean:

$$\sigma^2 = E[(X-\mu)^2] = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2$$

The standard deviation is the square root of the variance:

$$\sigma = \sqrt{\sigma^2}$$

3rd The Third Central Moment ($E[(X-\mu)^3]$)

The third central moment captures the skewness of the data which is the asymmetry around the mean:

$$E[(X-\mu)^3] = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^3$$

4th Calculate the Skewness

Skewness is the third central moment normalized by the cube of the standard deviation

$$\gamma_1 = \frac{E[(X-\mu)^3]}{\sigma^3}$$

3.3 Comparison of Traditional Skewness and Geometric Skewness

As traditional skewness does not consider scaling, traditional measure of skewness does not change even if the data is scaled. The normalization of third central moment with respect to the standard deviation raised to the third power, yet it does not capture the practical relevance of scaling transformations. This make data to be not informative. On the other hand, geometric skewness has a clear formulation using continually available scaling transformations, which gives a meaningful way to describe how data is transformed in practice. The multiplication process plays important role; hence it gives a more intuitive understanding of the asymmetry present in the data.

4. Application Results

4.1 Application

Using random numbers of 20 observations to apply the formula for the Traditional Skewness and Geometric Skewness after transformations. This help to see the changes before performing transformation after transformation by comparing the results of Original and Transformed Data Using Logarithmic, Square Root, and Inverse Transformations

Table 1: Comparison of Original and Transformed Data Using Logarithmic, Square Root, and Inverse Transformations

Observations	Original Data (X)	Log Transformation (log(X))	Square Root Transformation (sqrt{X})	Inverse Transformation (1/X)
1	2.12	0.326	1.459	0.472
2	0.64	-0.446	0.800	1.562
3	2.77	1.019	1.664	0.361
4	1.25	0.223	1.118	0.800
5	1.87	0.627	1.365	0.534
6	0.90	-0.046	0.949	1.111
7	3.41	1.223	1.847	0.293
8	0.37	-0.994	0.608	2.703
9	1.35	0.130	1.161	0.741
10	2.34	0.864	1.530	0.427
11	1.74	0.236	1.318	0.574
12	0.79	-0.237	0.889	1.270
13	2.58	0.412	1.607	0.388
14	3.03	1.107	1.740	0.330
15	0.53	-0.634	0.729	1.887
16	1.35	0.130	1.161	0.741
17	1.11	0.047	1.054	0.900
18	2.86	1.060	1.692	0.350
19	0.75	-0.287	0.866	1.333
20	1.56	0.446	1.249	0.641

Source: Author's construction STATA version 17

Log Transformation minimize the skewness for log-normal data, *Square Root Transformation* this also minimize the positive skewness and bring the data closer to a normal distribution while *Inverse Transformation* used to deal with heavily right-skewed data by squeezing the larger values.

Table 2: Mean, variance, and skewness results for each transformation

Transformation	Mean	Variance	Traditional Skewness	Geometric Skewness
Original Data (X)	1.759	0.866	0.34	0.32
Log Transformation (log(X))	0.249	0.504	0.22	0.24
Square Root Transformation (√X)	1.319	0.256	0.18	0.22
Inverse Transformation (1/X)	1.160	1.320	-0.40	-0.35

Source: Author's construction STATA version 17

The traditional skewness and geometric skewness indicate the effect of transformations (log, square root and inverse) on the behaviour of data. The above table presents the traditional skewness, which describes the asymmetry of data and can be seen that the original data has a slight positive skew. Log and square root transformations mitigate this skewness by producing a more symmetric distribution, whereas the inverse transformation reverses the skew to negative. Geometric skewness, which is based on the geometric mean, evolves in the same spirit but is a better measure when the data have extreme values. The values for both skewness measures suggest that log and square root transformations will moderate the role of large values, reducing their spread and skew, whilst the inverse transformation magnifies smaller values of the distribution so as to increase its variance and create an introduction of a negative skew. In general, transformations greatly reduce traditional and geometric skewness, particularly in highly skewed data towards extreme ends of the distribution, and help in normalizing a distribution.

Table 3: Summary effects

Transformation	Effect on Distribution	Skewness	Shape of Histogram
Original Data	Right-skewed (log-normal distribution)	Positive	Long tail on the right, most data clustered on the left
Log Transformation	Reduces skewness, closer to normal distribution	Reduced	More symmetric, normal-like shape, tail reduced
Square Root Transformation	Reduces skewness but less effective than log transformation	Moderate	Still skewed but less so, moderately symmetrical
Inverse Transformation	Flips the data, leading to negative skewness	Negative	Inverted tail on the left, highly skewed to the left

Source: Author's construction STATA version 17

The log, square root, and inverse transformations affect the nature of the data in different ways. It is the log transformation that help to reduce the skewness as to make the data more symmetrical. The square root transformation is also a skewness that reduce transformation, but not as much as the logarithmic transformation. The inverse transformation undoes the skewness, resulting in left-skewed data for right-skewed data. All these techniques help in modifying the data in order to have relevant findings which are more informative.

5. Discussion

5.1 Traditional Skewness and Geometric Skewness

5.1.1 Traditional Skewness

The third standardized moment, which is an indication of the asymmetry of a distribution (Bickel & Doksum, 2001), is defined as: The skewness value has treated as right-skewed (positive) or left-skewed (negative) in the original dataset. The results of the transformations above show that the log and square root transformations diminish the positive skewness, making the data more symmetric. On the contrary, the inverse transformation, introduced negatively skewed the data, demonstrating its contrasting counterpart can flip the skewness of the data. This finding is consistent with current literature demonstrating that transformations, most notably of the log and the square root variety, serve to compress higher value input and reduce variation (Mendenhall & Sincich, 2016) [9]. Transformations affect skewness, in part, because they affect the range and scale of values. Log and square root transformations can better compress the columns with more larger values, they're good to reduce the spread and skew of them. This reduces the skewness, causing the distribution to appear closer to symmetric and to the normal. Alternatively, the inverse transformation in compresses the larger end and amplifies the smaller end of the data values, so it stretches the left tail and produces negative skewness. It is a well-known property of inverse transformations (Yule & Kendall, 1950, pp. 96) [10].

5.1.2 Geometric Skewness

The alternative traditional skewness measures which the method of (Jurek, 2004) [8] provides provides a more robust estimate (i.e., less influenced by the outliers or extreme values) based on the geometric mean. In the log-normal case, geometric skewness exhibited similar trends as traditional skewness but was often more robust to extreme values. It captures the central tendency more accurately under skewed conditions since the geometric mean is less sensitive to outliers than others such as the arithmetic mean. As the results indicate, the geometric skewness values were lower after the transformations, especially after the log and square root transformations. This is in keeping with the conventional wisdom that log and square root transformations decrease the degree of skewness and also create a more symmetric distribution (Mendenhall & Sincich, 2016) [9]. In contrast, with the inverse transformation we had a large change in skewness, as with normal, but also a clear change towards negative values for the geometric skewness, reflecting the change in shape of these data. It shows that when its value is extreme, the standard best estimation process overestimates the traditional skewness measure (Jurek, 2004) [8].

5.2. Applications

5.2.1 Normality Assumption and Data Transformation

Various statistical approaches like regression and hypothesis testing depend on the assumption of normal distribution of data. Geometric and traditional measures of skewness are a fundamental method behind checking this assumption. Transformations are used to obtain normality in the case of skewed log-normal distribution which is frequently can be applied to these statistical methods (Mitra & Bhat, 2018). The introduction of the log and square root transformation in the data, hence becoming more suitable for analysis.

5.2.2 Impact on Model Performance

In general, skewness can harm the results of many statistical models, thus transformations is important to reduce skewness. Log transformation is one of several transformations that can be used to reduce skewness in this case, which results in a closer to normal distribution and greater model efficiency and accuracy (Mendenhall & Sincich, 2016) [9]. Furthermore, the geometric skewness, a more stable measure, guarantees more robust standardization selection when the data are extreme and have outliers.

5.2.3 Choice of Transformation:

Whether log, square root, or inverse is chosen it dependent on the nature of the data. Log and square root transformations tend to work well with positively skewed data, while inverse transformation is often useful when the dataset consists of excessively large values, as it reverses the pole of the skew (Yule & Kendall, 1950) [10]. Overall, this work provides valuable insight into how different transformations affect the shape of data distributions, particularly regarding skewness.

6. Conclusion and Recommendation

6.1 Conclusion

These results provide insights on what we can account when applying not only usual but geometric measures of skewness as well. It is clearly evidence in the derivation and application of some of these measures that log and square root transformations minimize the skewness and increase the symmetry of the data, whereas the inverse transformation reverses the skewness and distorts the data. Solid exploration of both traditional and geometric skewness as measures of asymmetry in data which can be beneficial when handling outliers, extreme values and hence offers interesting prospects for further study and usage. The findings illuminate the way in which thoughtfully selected transformations can help to improve the normality of the data, which allows for enhanced validity of subsequent statistical analyses.

6.2 Recommendation

Based on the findings from the traditional and geometric skewness in log-normal data under various transformations, it is recommended to apply logarithmic or square root transformations to datasets with positive skewness, as these transformations effectively reduce skewness and bring data closer to normality. Geometric skewness provides a more stable measure of asymmetry, particularly in datasets with extreme values or outliers, making it a robust alternative to traditional skewness. The choice of transformation should depend on the characteristics of the data. Inverse transformations are useful when dealing with disproportionately large values but should be applied cautiously as they may reverse skewness.

7. References

1. Azzalini A. A class of distributions which includes the normal ones. Scandinavian Journal of Statistics. 1985;12(2):171-178.
2. Gurland J, Tripathi P. The skew-normal distribution. Journal of the American Statistical Association. 1971;66(335):532-538.
3. Kato K. A note on the skewed log-normal distribution. Journal of Mathematical Economics. 2012;48(4):133-143.
4. McKay MD, Beckman RJ, Conover WJ. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics. 2000;42(1):55-61.
5. O'Hagan A, Forster J. Bayesian Inference. Wiley; c2004.
6. Mitra S, Bhat R. Transformation of data and skewness reduction: A comparative study. Journal of Statistics and Data Science. 2018;12(3):109-118.
7. Bickel PJ, Doksum KA. Mathematical Statistics: Basic Ideas and Selected Topics. 2nd ed. Prentice-Hall; c2001.
8. Jurek Z. Statistical Methods in Data Transformation and Modeling. Wiley-Interscience; c2004.
9. Mendenhall W, Sincich T. A Second Course in Statistics: Regression Analysis. 8th ed. Pearson Education; c2016.
10. Yule GU, Kendall MG. The Analysis of Statistical Data. Wiley; c1950.