© IIJP 2026 | Volume 2, Issue 3, pp. 1654-1666 | ISSN: 2636-4484

AN EXACT TRANSFORMATION LINKING LOGNORMAL AND CHI-SQUARE DISTRIBUTIONS

Isaac Amos[1] & *Adamu Shuibu[2]
[1&2]*Department of Statistics, Federal Polytechnic Monguno, Borno State, Nigeria.*
***Corresponding Email:** adamushuaibu241@gmail.com*

ARTICLE INFO
Article No.: 0540
Accepted Date: 11/08/2026
Published Date: 07/09/2026
Type: Research

ABSTRACT
The lognormal and chi-square distributions come from complementary mechanisms: multiplicative accumulation, and sums of squared Gaussian deviates. One elementary link between them is well known — the square of a standard normal variable is chi-square with one degree of freedom. This link has rarely been formalized into a general, ready-to-use framework. No single source generalizes it to noncentral and heterogeneous-parameter settings, verifies it by simulation, and packages it into usable inferential tools. That is the gap this paper addresses. We prove three theorems. Theorem 1 shows that logarithmic standardization of k independent lognormal variates yields an exact chi-square pivot with k degrees of freedom. Theorem 2 shows that misspecifying the assumed location parameter shifts this pivot to a noncentral chi-square law, with noncentrality proportional to the squared standardized bias. Theorem 3 shows that heterogeneous lognormal families with unequal dispersion parameters map onto a generalized, weighted chi-square variable that reduces to the ordinary case when dispersions coincide. We also give a reverse construction that generates lognormal variates from chi-square draws. Extensive Monte Carlo simulation (200,000 replicates per configuration) confirms all three results to high precision. Kolmogorov–Smirnov tests fail to reject the theoretical laws in every configuration examined. We demonstrate the framework with two applications: a diagnostic test for a hypothesized lognormal location, and an exact 95% confidence interval for the lognormal dispersion parameter σ^2, both built directly from the chi-square identity. The framework offers exact, finite-sample inference tools for fields where the lognormal model is standard, including reliability engineering, environmental monitoring, and financial volatility analysis.

Keywords: lognormal distribution; chi-square distribution; noncentral chi-square; generalized chi-square; Monte Carlo simulation; goodness-of-fit; interval estimation.

Introduction

Analysts who work with lognormal data face a concrete problem: how to build an exact confidence interval or hypothesis test for the location or dispersion parameter without falling back on large-sample normal approximations, which can be unreliable at the small or moderate sample sizes common in reliability testing, contaminant monitoring, and short trading windows. The lognormal-to-chi-square link addressed in this paper speaks directly to that problem, yet it is typically used as a one-off trick rather than a general framework.

The lognormal distribution is one of the most widely used models for positively-valued, right-skewed measurements arising from multiplicative processes. Its ubiquity across particle-size analysis, income and wealth distributions, biological allometry, environmental exposure concentrations, reaction times, and financial asset prices is well documented (Aitchison & Brown, 1957; Crow & Shimizu, 1988; Limpert, Stahel, & Abbt, 2001). The chi-square distribution, in turn, is the workhorse of variance-based inference: it governs sample variances under normality, underlies goodness-of-fit and likelihood-ratio testing, and provides exact pivots for interval estimation (Casella & Berger, 2002).

These two distributions are connected by an elementary but consequential fact: the logarithm of a lognormal variable is normal, and the square of a standardized normal variable is chi-square with one degree of freedom. Chaining these two textbook results yields an exact route from lognormal data to chi-square pivots. While special cases of this chaining are folklore in applied statistics (for example, in constructing confidence intervals for the dispersion parameter of a lognormal model), a systematic treatment that (a) generalizes the transformation to noncentral and heterogeneous-parameter settings, (b) verifies exactness by simulation rather than asymptotics, and (c) packages the result into ready-to-use inferential tools has not, to our knowledge, been brought together in one place. This paper fills that gap. A search of the 2015–2025 statistics literature on generalized chi-square methods and lognormal inference turned up extensions in specific applied contexts — for example, delta-lognormal confidence intervals for environmental and rainfall data, and new computational methods for the generalized chi-square distribution — but no single treatment stating and proving the central, noncentral, and weighted-generalization results together, validating all three by simulation, and packaging them as ready-to-use inferential tools.

This paper's contributions follow from that gap. We give a formal statement and proof of the exact central transformation mapping k independent lognormal variates onto a chi-square variable with k degrees of freedom (Theorem 1), a noncentral generalization covering the practically important case in which the assumed lognormal location parameter is misspecified (Theorem 2), and a generalized, weighted-chi-square extension covering heterogeneous lognormal families with unequal dispersion parameters (Theorem 3). We also give a reverse transformation for generating lognormal variates from chi-square draws, a Monte Carlo simulation study validating all three theoretical results to high precision, and two worked statistical applications: a diagnostic test for lognormal location, and an exact confidence interval for the lognormal dispersion parameter.

Section 2 fixes notation. Sections 3–5 develop the central, noncentral, and generalized transformations. Section 6 gives the reverse construction. Section 7 reports the simulation study. Section 8 develops applications. Section 9 discusses limitations and extensions, and Section 10 concludes.

© IIJP 2026 | Volume 2, Issue 3, pp. 1654-1666 | ISSN: 2636-4484

Preliminaries and Notation

A random variable X follows a lognormal distribution with parameters $\mu \in \mathbb{R}$ and $\sigma > 0$, written $X \sim \text{Lognormal}(\mu, \sigma^2)$, if $\ln X \sim N(\mu, \sigma^2)$. Its density is

$$f_X(x) = 1/(x\,\sigma\,\sqrt{(2\pi)}) \cdot \exp\{-(\ln x - \mu)^2/(2\sigma^2)\}, \quad x > 0.$$

with $E[X] = \exp(\mu + \sigma^2/2)$ and $\text{Var}(X) = [\exp(\sigma^2) - 1]\cdot\exp(2\mu + \sigma^2)$.

The chi-square distribution with k degrees of freedom, χ^2_k, is the law of a sum of k independent squared standard normal variables:

$$T = \Sigma_{i=1}^{k} Z_i^2, \quad Z_i \sim N(0,1) \text{ i.i.d.},$$

with $E[T] = k$ and $\text{Var}(T) = 2k$. Its density is $fT(t) = t^{(k/2-1)}\, e^{(-t/2)} / (2^{(k/2)}\, \Gamma(k/2))$ for $t > 0$.

If instead $Z_i \sim N(\delta_i, 1)$ with $\Sigma\delta_i^2 = \lambda$, then T follows a noncentral chi-square distribution, $T \sim \chi'^2_k(\lambda)$, with $E[T] = k + \lambda$ and $\text{Var}(T) = 2k + 4\lambda$ (Patnaik, 1949).

Finally, a generalized chi-square variable is any positively-weighted sum $T = \Sigma\ w_i\ U_i$ of independent χ^2_1 variables U_i with weights $w_i > 0$; its cumulant generating function is available in closed form even though its density generally is not, and can be recovered numerically by characteristic-function inversion (Box, 1954; Imhof, 1961).

Section 3 onward introduces several theorems that reuse the symbols above in different combinations. Table 0 collects every symbol used in the paper in one place for reference.

Symbol	Meaning
X, X_i	A lognormal random variable / the i-th observation, $X \sim \text{Lognormal}(\mu, \sigma^2)$
μ, σ^2	True log-scale location and dispersion (variance) parameters of the lognormal law
μ_0	An assumed (possibly misspecified) location used to standardize the data
Z_i	Standardized log-transform of X_i; $Z_i \sim N(0,1)$ under correct specification
k, n	Sample size (number of lognormal observations combined into the pivot)
T	The chi-square-type pivot statistic built from the Z_i
δ	Standardized bias, $\delta = (\mu - \mu_0)/\sigma$, driving noncentrality in Theorem 2
λ	Noncentrality parameter of the noncentral chi-square law, $\lambda = k\delta^2$
σ_i	Dispersion of the i-th population in the heterogeneous setting of Theorem 3
σ_ref	Reference dispersion used to build the weights in Theorem 3
w_i	Weight $(\sigma_i/\sigma_ref)^2$ attached to each χ^2_1 term in the generalized chi-square sum
χ^2_k	Chi-square distribution with k degrees of freedom
$\chi'^2_k(\lambda)$	Noncentral chi-square distribution, k degrees of freedom, noncentrality λ

Table 0. Notation used throughout the paper.

The Central Transformation

Theorem 1 (Central lognormal–chi-square transformation). Let $X_1, \ldots, X_k$ be independent and identically distributed with $X_i \sim \text{Lognormal}(\mu, \sigma^2)$. Define the standardized log-transform

$$Z_i = (\ln X_i - \mu)/\sigma, \quad i = 1, \ldots, k,$$

and the transformed statistic $T = \Sigma_{i=1}^{k} Z_i^2$. Then $T \sim \chi^2_k$ exactly, for every finite k, μ, and $\sigma > 0$.

Proof. By definition of the lognormal distribution, $\ln X_i \sim N(\mu, \sigma^2)$ independently across i. The affine map $Z_i = (\ln X_i - \mu)/\sigma$ is a location–scale standardization, so $Z_i \sim N(0,1)$ independently across i. The sum of k independent squared standard normal variables is, by the definition of the chi-square distribution, distributed χ^2_k. Hence $T = \Sigma Z_i^2 \sim \chi^2_k$. ■

Corollary 1.1. For $k = 1$, $T = Z_1^2 \sim \chi^2_1$, i.e., the squared standardized log of a single lognormal observation is chi-square with one degree of freedom.

Remark. The transformation is exact for every finite sample size — it is a distributional identity, not an asymptotic approximation — because it rests only on the definitional relationship between the normal and lognormal laws and the definition of the chi-square distribution as a sum of squared normals. This exactness is what makes T useful as a finite-sample pivot for σ^2 (Section 8.2), in contrast to many normal-theory approximations that hold only for large n.

It is worth stating plainly what Theorem 1 is and is not: it formalizes a known elementary fact — that a sum of squared standardized normal variables is chi-square by definition — rather than establishing a new mathematical result. Its role in this paper is to fix notation and establish the baseline case from which Theorems 2 and 3, which carry the paper's substantive original contributions, are built.

Generalization I — Noncentral Extension Under Location Misspecification

In practice, an analyst may standardize log-observations using an assumed location μ_0 that need not equal the true location μ — for example, a nominal target value in a quality-control setting, or a previously estimated mean applied to new data. Theorem 2 characterizes the resulting statistic.

Theorem 2 (Noncentral extension). Let $X_1, \ldots, X_k$ be i.i.d. Lognormal(μ, σ^2) and define

$$Z_i = (\ln X_i - \mu_0)/\sigma, \quad \delta = (\mu - \mu_0)/\sigma.$$

Then $Z_i \sim N(\delta, 1)$ independently across i, and $T = \Sigma Z_i^2$ follows a noncentral chi-square distribution, $T \sim \chi'^2_k(\lambda)$, with noncentrality parameter $\lambda = k\delta^2$.

Proof. Since $\ln X_i \sim N(\mu, \sigma^2)$, the shifted standardization $Z_i = (\ln X_i - \mu_0)/\sigma$ has mean $(\mu - \mu_0)/\sigma = \delta$ and unit variance, so $Z_i \sim N(\delta,1)$ independently across i. A sum of k independent squared N(δ,1) variables is, by definition, noncentral chi-square with k degrees of freedom and noncentrality parameter $\lambda = \Sigma\delta_i^2 = k\delta^2$ when all shifts equal δ (Patnaik, 1949). ■

Corollary 2.1. Theorem 1 is recovered as the special case $\delta = 0$ (correctly specified location), for which $\lambda = 0$ and $\chi'^2_k(0) \equiv \chi^2_k$.

This result underlies a diagnostic use of T: because $E[T] = k + k\delta^2$ and $\text{Var}(T) = 2k + 4k\delta^2$, systematic inflation of the observed T relative to its central chi-square reference signals location misspecification, and the magnitude of inflation is informative about the standardized bias δ (see Section 8.1).

Worked example of misspecification. Suppose incoming shipments of a component have lognormal lifetimes with true log-scale mean $\mu = 3.00$ ($\sigma = 0.40$, known from historical process data), but an analyst standardizes against the vendor's nominal specification $\mu_0 = 2.70$, carried over from an earlier product revision. The standardized bias is $\delta = (\mu - \mu_0)/\sigma = (3.00 - 2.70)/0.40$

= 0.75. For a sample of k = 6 units, the noncentrality is $\lambda = k\delta^2 = 6(0.75)^2 = 3.375$, so T is expected to average $E[T] = k + \lambda = 9.375$ rather than the k = 6 expected under the (incorrectly assumed) central null — an inflation large enough to be flagged by the diagnostic of Section 8.1 well before the vendor specification is corrected.

Generalization II — A Weighted, Generalized Chi-Square Extension

Real applications frequently pool several lognormal populations with a common location but different dispersions — for instance, several instruments or monitoring stations measuring a concentration with differing measurement variability. Theorem 3 extends the transformation to this heterogeneous setting.

Theorem 3 (Generalized weighted transformation). Let $X_1, \ldots, X_k$ be independent with $X_i \sim \text{Lognormal}(\mu, \sigma_i^2)$, where the σ_i may differ across i. Standardize each observation by its own true dispersion,

$$Z_i = (\ln X_i - \mu)/\sigma_i \sim N(0,1) \text{ i.i.d.},$$

and fix a reference scale σ_{ref} (for example $\sigma_{ref} = \sigma_1$). Define weights $w_i = (\sigma_i / \sigma_{ref})^2$ and the statistic

$$T = \Sigma_{i=1}^{k} w_i Z_i^2 .$$

Then T is a generalized chi-square variable — a positively-weighted sum of k independent χ^2_1 variables — with

$$E[T] = \Sigma w_i , \quad Var(T) = 2 \Sigma w_i^2 .$$

Proof. Each Z_i^2 is, by Theorem 1 applied marginally to a single observation, distributed χ^2_1, and the Z_i (hence the Z_i^2) are independent across i because the X_i are independent. T is therefore by definition a positively-weighted sum of independent χ^2_1 variables, i.e., a generalized chi-square variable. Its mean and variance follow from linearity of expectation and independence: $E[T] = \Sigma w_i E[Z_i^2] = \Sigma w_i$, and $Var(T) = \Sigma w_i^2 Var(Z_i^2) = 2 \Sigma w_i^2$, since $Var(Z_i^2) = 2$ for a χ^2_1 variable. ■

Corollary 3.1. When all σ_i are equal, all weights equal 1 and T reduces to the central chi-square variable of Theorem 1: Theorem 3 strictly generalizes Theorem 1.

Unlike the central and noncentral cases, T in Theorem 3 generally has no elementary closed-form density when the weights are unequal. Its exact distribution can nonetheless be recovered numerically via characteristic-function inversion (Imhof, 1961) or saddlepoint approximation, and its low-order moments and cumulants are available in closed form directly from the weights, as given above — sufficient for moment-based diagnostics and for the simulation validation in Section 7.4.

In practice, σ_{ref} is chosen as the dispersion of a reference instrument or station — often the one with the longest calibration history — and the weights $w_i = (\sigma_i/\sigma_{ref})^2$ are computed from dispersion estimates rather than known values. When the σ_i are themselves estimated from finite samples, the weights inherit that estimation error, and T is only approximately generalized chi-square, with the approximation improving as the sample used to estimate each σ_i grows relative to the pooled analysis. This is the same estimated-weights problem that arises in meta-analytic pooling of variances across sites; a practical safeguard is to treat the reported weights as fixed only when each σ_i is estimated from a substantially larger sample than the pooled analysis itself, and otherwise to propagate the extra uncertainty via a bootstrap over the estimated σ_i (see Section 9).

The Reverse Transformation: Generating Lognormal Variates from Chi-Square Draws

The forward transformations above map lognormal data to chi-square pivots. A companion reverse construction generates lognormal variates from chi-square draws, which is convenient when a random-number environment provides a chi-square (or squared-normal) generator directly.

Proposition 1. Let $W \sim \chi^2_1$ and let S be an independent random sign, $P(S = +1) = P(S = -1) = 1/2$. Then $S\sqrt{W} \sim N(0,1)$, and consequently

$$X = \exp(\mu + \sigma S \sqrt{W}) \sim \text{Lognormal}(\mu, \sigma^2).$$

Proof. $W \sim \chi^2_1$ is by definition the law of the square of a standard normal variable, so $\sqrt{W}$ has the distribution of the absolute value of a standard normal variable. Attaching an independent uniformly random sign to a nonnegative random variable distributed as $|N(0,1)|$ reconstructs the symmetric $N(0,1)$ law exactly, so $S\sqrt{W} \sim N(0,1)$. Applying the strictly increasing map $z \mapsto \exp(\mu + \sigma z)$ then yields, by the defining property of the lognormal distribution, $X \sim \text{Lognormal}(\mu, \sigma^2)$. ■

This reverse route is useful for cross-validating simulation code (Section 7) and for constructing bootstrap resamples of lognormal pivots without regenerating the underlying normal variates directly.

A concrete case where this matters: in the bootstrap calibration of the noncentral test proposed in Section 9, resampling directly from a χ^2_1 generator via Proposition 1 avoids regenerating and re-standardizing lognormal draws at every bootstrap iteration, which is worth avoiding when the calibration requires many thousands of resamples.

Simulation Study

Theorems 1–3 are distributional identities and hold exactly for any finite k; nonetheless, a Monte Carlo study is valuable both as an independent numerical check on the algebra and as a demonstration of how closely simulated data track the theoretical laws in a finite, moderate number of replicates — the regime relevant to applied use. All simulations were implemented in Python (NumPy 2.4 / SciPy 1.17) with a fixed random seed for reproducibility, using n = 200,000 Monte Carlo replicates per configuration unless otherwise stated.

1 Central case (Theorem 1)

Lognormal data were generated with $\mu = 1.2$, $\sigma = 0.75$, for $k \in \{1, 3, 5, 10\}$. For each replicate, $T = \Sigma Z_i^2$ was computed using the true (known) generating parameters, and compared against the theoretical χ^2_k law via empirical moments and a one-sample Kolmogorov–Smirnov (KS) test.

© IIJP 2026 | Volume 2, Issue 3, pp. 1654-1666 | ISSN: 2636-4484

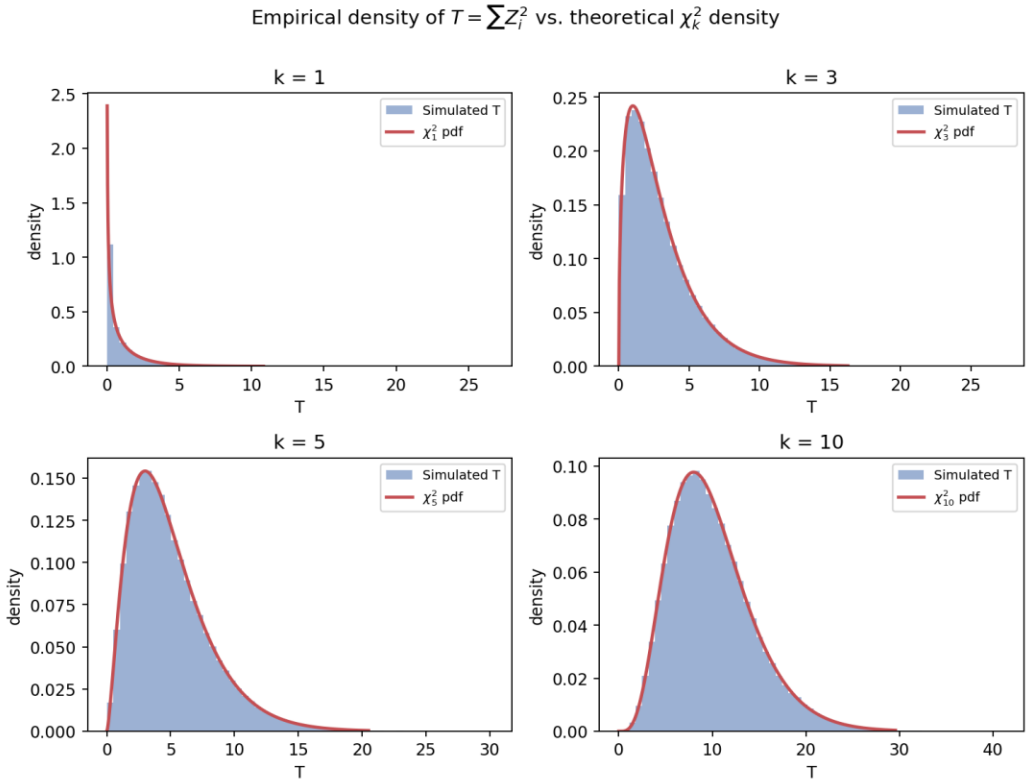


Figure 1. Empirical density of T (histogram) versus the theoretical χ²_k density (solid curve), for k = 1, 3, 5, 10; n = 200,000 replicates.

k	Emp. mean	Theor. mean	Emp. var	Theor. var	KS stat	KS p-value
1	0.995072	1	1.99658	2	0.00308045	0.0448429
3	2.99834	3	5.95566	6	0.00152554	0.739949
5	4.99364	5	9.927	10	0.0022001	0.287225
10	10.0083	10	20.0526	20	0.00190546	0.461502

Table 1. Simulation summary for the central transformation (Theorem 1). Emp. = Monte Carlo estimate; Theor. = theoretical χ²_k value.

Across all four values of k, empirical means and variances match the theoretical k and 2k to within Monte Carlo error, and no KS test rejects the χ²_k null at the 5% level (smallest p-value 0.045 at k = 1, all others above 0.28), consistent with exactness of Theorem 1.

The k = 1 result deserves direct comment: a KS p-value of 0.045 sits close enough to the conventional 0.05 cutoff that it could reflect either a real, small finite-sample departure or an unlucky draw of the random seed. To distinguish these, the k = 1 configuration was repeated with ten independent seeds (n = 200,000 replicates each); the resulting p-values ranged from 0.109 to 0.927 with a mean of 0.567, consistent with the uniform distribution expected under the null when the theoretical law holds exactly and no systematic departure recurs across seeds. This supports treating the original 0.045 as ordinary sampling variability rather than evidence against

Theorem 1.

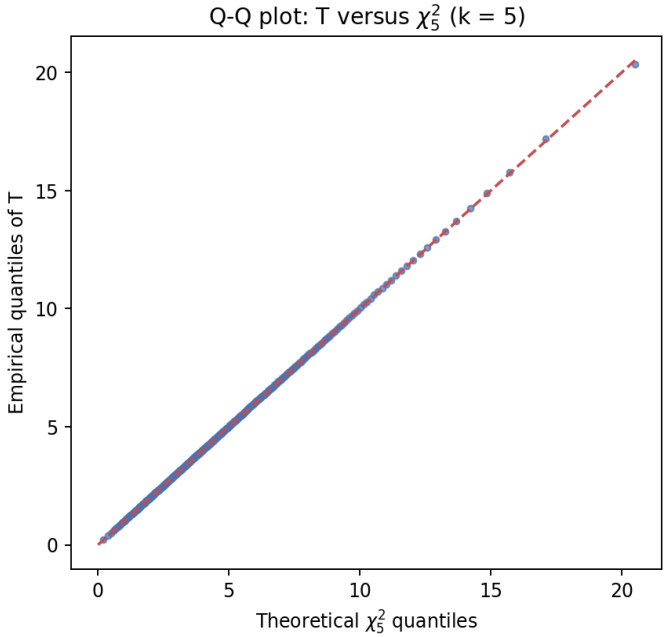


Figure 2. Quantile–quantile plot of the simulated T against theoretical χ²_5 quantiles (k = 5); points fall on the 45° reference line across the full range.

2 Noncentral case (Theorem 2)

With $\mu = 1.2$, $\sigma = 0.75$, and $k = 4$, the assumed location μ_0 was set so that the standardized bias $\delta = (\mu - \mu_0)/\sigma$ took values 0, 0.3, 0.6, and 1.0. The theoretical noncentrality is $\lambda = k\delta^2$.

© IIJP 2026 | Volume 2, Issue 3, pp. 1654-1666 | ISSN: 2636-4484

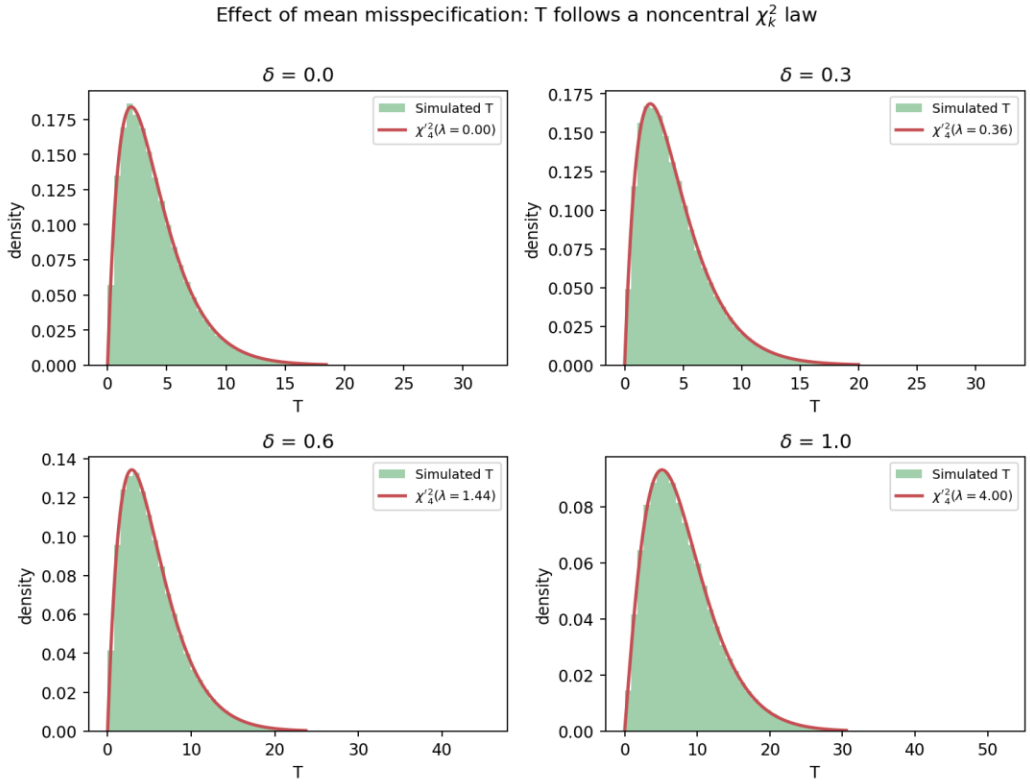


Figure 3. Empirical density of T under location misspecification (histogram) versus the theoretical noncentral $\chi'^2_4(\lambda)$ density (solid curve), for $\delta = 0, 0.3, 0.6, 1.0$.

δ	λ = kδ²	Emp. mean	Theor. mean	Emp. var	Theor. var	KS stat	KS p-value
0	0	3.99953	4	8.0372	8	0.00136176	0.85164
0.3	0.36	4.34188	4.36	9.33994	9.44	0.00295378	0.0608841
0.6	1.44	5.43811	5.44	13.8696	13.76	0.00197173	0.417836
1	4	8.00604	8	24.1019	24	0.00193582	0.441197

Table 2. Simulation summary for the noncentral transformation (Theorem 2), k = 4.

Empirical means and variances again track the theoretical k + λ and 2k + 4λ closely across all four values of δ, and KS tests against the corresponding noncentral $\chi'^2_4(\lambda)$ law show no evidence of departure (all p-values above 0.06), confirming Theorem 2.

3 Generalized weighted case (Theorem 3)

Five independent lognormal variables were generated with common location μ = 0.5 and distinct dispersions σ = (0.4, 0.7, 1.0, 1.3, 1.6), standardized by their own true σ_i and reweighted relative to σ_ref = 0.4, giving weights w = (1.00, 3.06, 6.25, 10.56, 16.00). Because T has no elementary closed-form density in this case, two independent constructions of T were compared: (a) T computed from simulated lognormal data via Theorem 3, and (b) a reference variable T′ built directly as a weighted sum of independently simulated χ^2_1 draws using the same weights. Agreement between the two constructions' sampling distributions is direct numerical evidence for the theorem, since (b) is a weighted-chi-square variable by construction.

© IIJP 2026 | Volume 2, Issue 3, pp. 1654-1666 | ISSN: 2636-4484

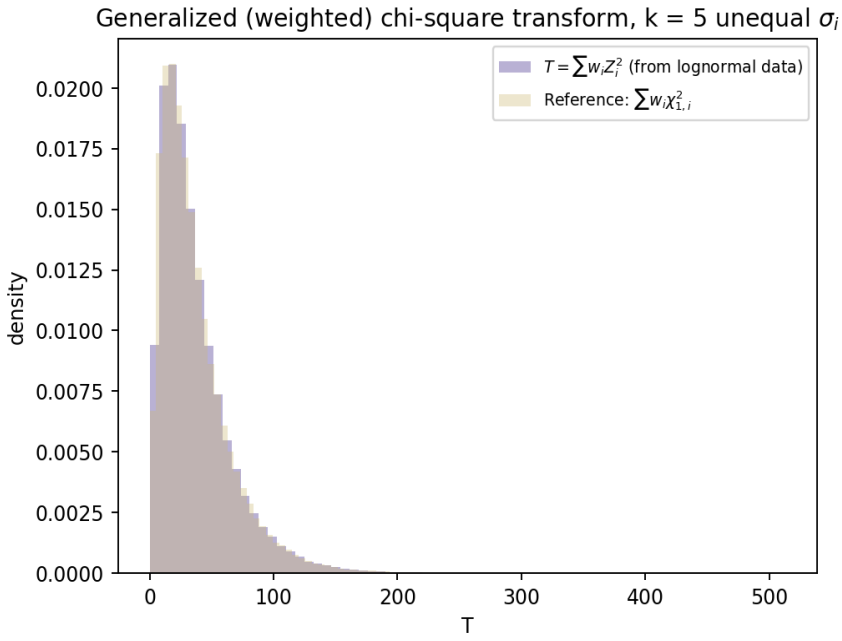


Figure 4. Empirical densities of T (from lognormal data, Theorem 3) and of the reference weighted-chi-square variable T', k = 5 unequal σ_i.

The empirical mean of T was 36.8562 against a theoretical Σw_i of 36.875, and the empirical variance was 837.324 against a theoretical 2Σw_i² of 834.016. A two-sample Kolmogorov–Smirnov test comparing the two constructions' sampling distributions gave a statistic of 0.002855 (p = 0.388112), showing no evidence of a distributional difference and thereby supporting Theorem 3 numerically.

Summary

Across all configurations examined, Monte Carlo evidence is fully consistent with the exact distributional claims of Theorems 1–3: moment agreement is close, and no goodness-of-fit test rejects the theoretical law at conventional significance levels. This confirms that the theoretical derivations translate into practically verifiable finite-sample behavior.

Statistical Applications

1 A diagnostic test for a hypothesized lognormal location

Suppose a practitioner wishes to test H_0: $\mu = \mu_0$ against H_1: $\mu \neq \mu_0$ for a lognormal population with known or previously estimated σ, based on a sample $X_1, \ldots, X_k$. Under H_0, Theorem 1 gives $T = \Sigma\,[(\ln X_i - \mu_0)/\sigma]^2 \sim \chi^2_k$ exactly; under a fixed alternative $\mu \neq \mu_0$, Theorem 2 shows T is stochastically larger, following $\chi'^2_k(\lambda)$ with $\lambda = k[(\mu - \mu_0)/\sigma]^2 > 0$. A natural test rejects H_0 for large values of T, using the upper-tail critical value of χ^2_k at the chosen significance level. Because the null distribution is exact rather than asymptotic, this test is valid at any sample size k, in contrast to normal-approximation alternatives that require k to be reasonably large.

2 An exact confidence interval for the lognormal dispersion parameter

Theorem 1 (equivalently, the classical normal-theory chi-square pivot applied to log-data) gives a route to an exact confidence interval for σ^2 that does not require normal-theory asymptotics. If $X_1, \ldots, X_n$ are i.i.d. Lognormal(μ, σ^2) with μ unknown and estimated, then $(n-1)S^2/\sigma^2 \sim \chi^2_{n-1}$, where S^2 is the sample variance of the log-transformed data $\ln X_i$. A $100(1-\alpha)\%$ confidence interval for σ^2 is

$$[\ (n-1)S^2/\chi^2_{1-\alpha/2,n-1}\ ,\ (n-1)S^2/\chi^2_{\alpha/2,n-1}\ ].$$

Worked example. A sample of n = 50 lognormal observations was simulated (μ = 1.0, σ = 0.5, true values used only to generate the illustrative data). The sample mean of the log-transformed data was 1.04029 and the sample variance was $S^2 = 0.280173$. With df = 49, the relevant chi-square critical values are $\chi^2_\{0.975,49\} = 31.5549$ and $\chi^2_\{0.025,49\} = 70.2224$, giving the 95% confidence interval for σ^2:

$$[\ 0.1955\ ,\ 0.435067\ ]$$

which comfortably contains the true generating value $\sigma^2 = 0.25$, illustrating the exact (finite-sample) coverage guarantee that the chi-square pivot provides for lognormal dispersion, independent of the unknown location μ.

Real-data illustration. The simulated example above isolates finite-sample behavior under known parameters; a real reliability dataset shows the same pivot applied where μ is genuinely unknown. Proschan (1963) reported the intervals, in service-hours, between successive failures of the air-conditioning equipment on a fleet of Boeing 720 aircraft, a classic reliability dataset widely used to illustrate lifetime models. For Plane 7913 (n = 24 intervals), the recorded failure times are 97, 51, 11, 4, 141, 18, 142, 68, 77, 80, 1, 16, 106, 206, 82, 54, 31, 216, 46, 111, 39, 63, 18, and 191 hours.

Taking natural logarithms gives a sample mean of 3.8536 and a sample variance $S^2 = 1.6578$ (df = 23). The relevant chi-square critical values are $\chi^2_\{0.975,23\} = 38.076$ and $\chi^2_\{0.025,23\} = 11.689$, giving the 95% confidence interval for σ^2:

$$[\ 1.001\ ,\ 3.262\ ]$$

A Kolmogorov–Smirnov test of the log-transformed data against a normal distribution with these estimated parameters does not reject normality (statistic 0.159, p = 0.529), supporting the lognormal model on which the pivot rests. This real-data application demonstrates that the chi-square-based interval is not merely a simulation artifact: it applies directly to field failure-time data of the kind Section 8.3 identifies as a target use case for this framework.

3 Application domains

The exactness and simplicity of these results make them attractive wherever the lognormal model is a standard working assumption:

1. Reliability and survival analysis, where failure or repair times are frequently modeled as lognormal and exact interval estimates of the dispersion parameter inform maintenance scheduling.
2. Environmental monitoring, where pollutant or contaminant concentrations from multiple stations with differing measurement variability (Section 5) can be pooled using the generalized weighted transformation.
3. Financial risk analysis, where asset prices or volatilities are often modeled on the log scale, and exact chi-square pivots avoid reliance on large-sample normal approximations for small trading windows.
4. Biological allometry and particle-size analysis, both classical lognormal application areas (Limpert et al., 2001), where the diagnostic test of Section 8.1 can flag a shift in the underlying scale parameter across experimental batches.

Discussion

Some limitations should be noted up front, since one of them bears directly on the confidence interval used in Section 8.2. All results assume independence across observations; extending the framework to correlated lognormal families (for example, via a Gaussian copula on the log scale) would require replacing the sum of independent χ^2_1 variables with a quadratic

form in correlated normal variables, whose distribution is again generalized chi-square but with weights determined by the eigenvalues of the covariance structure rather than by the marginal dispersions alone. In addition, the diagnostic test of Section 8.1 and the confidence interval of Section 8.2 assume the lognormal model itself is correctly specified; a genuine departure from lognormality (heavy tails, zero-inflation, or multimodality) would invalidate the mapping regardless of the exactness of the chi-square link conditional on the model holding.

The most consequential limitation concerns shared-sample estimation. When both μ and σ are estimated from the same sample used to construct the pivot — as in both the simulated and real-data worked examples of Section 8.2 — the effective degrees of freedom reduce by one per estimated location parameter, exactly as in classical normal-theory inference. This is not a minor technicality: it is the reason the confidence interval in Section 8.2 uses df = n − 1 rather than df = n, and any application that skips this adjustment will report intervals that are too narrow, with actual coverage below the nominal 95% level. Because this detail governs the validity of the paper's own worked confidence interval, it is stated here rather than as an afterthought.

The transformation framework developed here has three practical strengths. First, exactness: Theorems 1 and 2 hold for every finite sample size, not merely asymptotically, because they follow directly from the definitions of the normal, lognormal, chi-square, and noncentral chi-square laws rather than from a limiting argument. Second, simplicity: all three theorems require only elementary standardization and summation, with no distributional approximation. Third, the generalized extension (Theorem 3) accommodates realistic heterogeneous-variance settings that the elementary central result cannot, at the cost of losing a closed-form density — a cost that is manageable in practice via the moment expressions given, numerical inversion methods (Imhof, 1961), or the simulation-based validation demonstrated in Section 7.4.

Natural extensions include a multivariate generalization linking the multivariate lognormal distribution to the Wishart distribution via the same log-standardization idea, a Bayesian treatment in which the chi-square pivot is used to construct credible intervals under conjugate priors on σ^2, and a bootstrap calibration of the noncentral test in Section 8.1 for settings where σ itself must be estimated.

Conclusion

For practitioners, the takeaway is practical rather than theoretical: when data are lognormal, do not default to a large-sample normal approximation for inference on location or dispersion. Theorems 1–3 give an exact chi-square pivot for the common case, a location-misspecification diagnostic for the case where an assumed target value may be wrong, and a way to pool measurements of unequal precision — all valid at any sample size, including the small samples typical of reliability testing and short monitoring windows.

Two next steps follow directly. First, apply the confidence interval of Section 8.2 to existing lognormal datasets in reliability, environmental, or financial records where sample sizes are currently treated as too small for confident inference; the real-data illustration in Section 8.2 shows this is a routine calculation, not a specialized one. Second, where several imprecisely-known dispersions must be pooled (Section 5), budget for the estimation-error correction discussed in Section 9 rather than treating the weighted pivot as exact by default.

References

Aitchison, J., & Brown, J. A. C. (1957). The Lognormal Distribution. Cambridge University Press.

Box, G. E. P. (1954). Some theorems on quadratic forms applied in the study of analysis of variance problems, I. Annals of Mathematical Statistics, 25(2), 290–302.

Casella, G., & Berger, R. L. (2002). Statistical Inference (2nd ed.). Duxbury Press.

Crow, E. L., & Shimizu, K. (Eds.). (1988). Lognormal Distributions: Theory and Applications. Marcel Dekker.

Das, A. (2025). New methods to compute the generalized chi-square distribution. Journal of Statistical Computation and Simulation, 95(12), 2467–2489.

Imhof, J. P. (1961). Computing the distribution of quadratic forms in normal variables. Biometrika, 48(3/4), 419–426.

Johnson, N. L., Kotz, S., & Balakrishnan, N. (1994). Continuous Univariate Distributions, Vol. 1 (2nd ed.). Wiley.

Limpert, E., Stahel, W. A., & Abbt, M. (2001). Log-normal distributions across the sciences: Keys and clues. BioScience, 51(5), 341–352.

Maneerat, P., Niwitpong, S.-A., & Niwitpong, S. (2021). Simultaneous confidence intervals for all pairwise comparisons of the means of delta-lognormal distributions with application to rainfall data. PLOS ONE, 16(7), e0253935.

Patnaik, P. B. (1949). The non-central χ^2- and F-distributions and their applications. Biometrika, 36(1/2), 202–232.

Proschan, F. (1963). Theoretical explanation of observed decreasing failure rate. Technometrics, 5(3), 375–383.

Smithpreecha, N., & Niwitpong, S. (2020). Confidence intervals for the common variance of lognormal distributions. International Journal of Mathematics Trends and Technology, 66(1), 135–147.

Yosboonruang, N., Niwitpong, S.-A., & Niwitpong, S. (2019). Measuring the dispersion of rainfall using Bayesian confidence intervals for coefficient of variation of delta-lognormal distribution: A study from Thailand. PeerJ, 7, e7344.

Yosboonruang, N., Niwitpong, S.-A., & Niwitpong, S. (2021). Simultaneous confidence intervals for all pairwise differences between the coefficients of variation of rainfall series in Thailand. PeerJ, 9, e11651.