Journal of Probability and Statistical Science
24(2), 1-15Aug. 2026

Lognormal distribution - Confidence Intervals for the Mean

Prasanth Sambaraju
Independent Researcher

ABSTRACT

Lognormal distributions are used in modelling continuous random variables which are greater than or equal to zero. A lognormal distribution can be converted into a normal distribution by logarithmic transformation of the random variable. Lognormal distributions are used in medicine to model latent periods of infectious disease, and in environmental sciences to estimate the distribution of particle size. The manuscript describes an approach to compute the confidence intervals for the mean of lognormal distribution using the loglikelihood ratio method. Simulations were used to evaluate the coverage probability for different sample sizes and the results obtained were compared with different reported methods.

***Keywords**:* Lognormal distribution, maximum likelihood estimation, loglikelihood ratio, simulation, confidence intervals, mean.

1. Introduction

Data obtained from biomedical research is often skewed. The statistical inferences based on the normal distribution is well known, one way to deal is to transform such data into normally distributed data. The log transformation is widely used for this purpose when the data is from a population that is positive and skewed [1].

If X denotes a variable that follows a lognormal distribution, the expected value is given by $E(X)=\theta$ and variance $Var(X)=\delta^2$. The log transformed, normally distributed variable Y = log(X), has a mean value of $E(Y)=\mu$ and variance $Var(Y)=\sigma^2$. The expected value (θ) is given by Equation (1), and the mean value of X is not equal to the antilog of mean value of Y [5].

$$\theta = exp\,(\mu + \frac{\sigma^2}{2}) \tag{1}$$

The loglikelihood (LL) estimate for samples from a lognormal distribution with parameters, mean (μ) and standard deviation (σ) is given by Equation (2)

$$LL = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}ln\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(lnx_i - \mu)^2 - \sum_{i=i}^{n} lnx_i \tag{2}$$

The LL estimate for samples $\{x_1, \ldots, x_n\}$ lognormal distribution is equal to the LL estimate for samples from $\{\ln x_1, \ldots, \ln x_n\}$ minus the last term ($\sum_{i=1}^{n} lnx_1$). This term doesn't affect the parameter value that maximizes LL, so it can be assumed that maximum LL can be achieved by using the values of μ and σ given by Equation (3) from the samples $\{\ln x_1, \ldots, \ln x_n\}$ obtained from a normal distribution [6, 2].

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

To obtain the unbiased estimator of standard deviation n is replaced by n-1 [6, 4].

- Received August 24, 2026, in final form August 2026.
- Prasanth Sambaraju (corresponding author) is an Independent Researcher
  Email: prashanth.kng1@gmail.com

1.1 Confidence intervals using loglikelihood ratio

Confidence intervals for mean of lognormal distribution based on loglikelihood ratio has not been reported in the literature. The loglikelihood confidence intervals can be obtained by estimating the values of values for μ_1 and σ_1 that satisfy the Equation (4) [3]

$$-2\,\frac{LL(\mu_1,\sigma_1)}{LL(\hat{\mu}_1,\hat{\sigma}_1)} = \chi^2{}_{1-\alpha,df} \tag{4}$$

where α is the significance level; $LL(\hat{\mu}_1,\hat{\sigma}_1)$ = loglikelihood value obtained using lognormal distribution parameters for the sample data; *df* is the degrees of freedom, in case of joint confidence intervals a value of 2 is used.

The above equation can be rearranged as shown in Equation (5). By estimating the values for μ_1 and σ_1 that satisfy this equation, the confidence intervals for mean of lognormal distribution can be obtained.

$$LL(\mu_1,\sigma_1) = LL(\hat{\mu}_1,\hat{\sigma}_1) - \frac{\chi^2{}_{1-\alpha,df}}{2} \tag{5}$$

2. Materials and Method

The use of loglikelihood ratio to compute the confidence interval for θ was demonstrated using reported sample data given in Table 1. The sample consisted of 40 observations from a lognormal distribution with parameters μ=5 and σ=1. The population mean of these sample values is $\theta = \exp(\mu+\sigma^2/2)$ = 244.69 [5].

Table 1: Sample data from lognormal distribution.

914.9	1568.3	50.5	94.1	199.5	23.8	70.5	213.1	914.9	1568.3
44.1	331.7	139.3	115.6	38.4	357.1	725.9	253.2	44.1	331.7
905.6	155.4	138.1	95.2	75.2	275	401.1	653.8	905.6	155.4
390.8	483.5	62.6	128.5	81.5	218.5	308.2	41.2	390.8	483.5

Microsoft Excel was used to compute the confidence interval using the loglikelihood ratio method. The confidence intervals were obtained as described below.

1. Sample size, mean and standard deviation of log sample values are computed
2. Using Equations (2) and (3), MLE (maximum likelihood estimation) was obtained by using mean and standard deviation values.
3. Chi-square distribution function in Microsoft Excel and MLE value were used to obtain the $LL(\mu_1,\sigma_1)$ value.
4. Using the fill series option: Home → Fill → Series → Series in Columns → Step value (0.01) and Stop Value (1.4)—numbers from 0.76 to 1.4 are entered into cells H5 to H69. These values correspond to different standard deviation values.
5. Using the fill series option: Home → Fill → Series → Series in Rows → Step value (0.01) and Stop Value (5.6)—numbers from 4.7 to 5.6 are entered into cells I4 to CU4. These values correspond to different mean values.
6. Different MLE values were obtained for different mean and standard deviation values using mixed reference formula in cells I5 to CU69.
7. Using conditional formatting, cell values greater than $LL(\mu_1,\sigma_1)$ values are highlighted. The confidence intervals of the parameters correspond to the highlighted lowest and highest values on both axes. Figure *1* shows the different formulas used to obtain the loglikelihood ratio confidence intervals, whereas Figure 2 shows the contour plot using the loglikelihood ratio method.

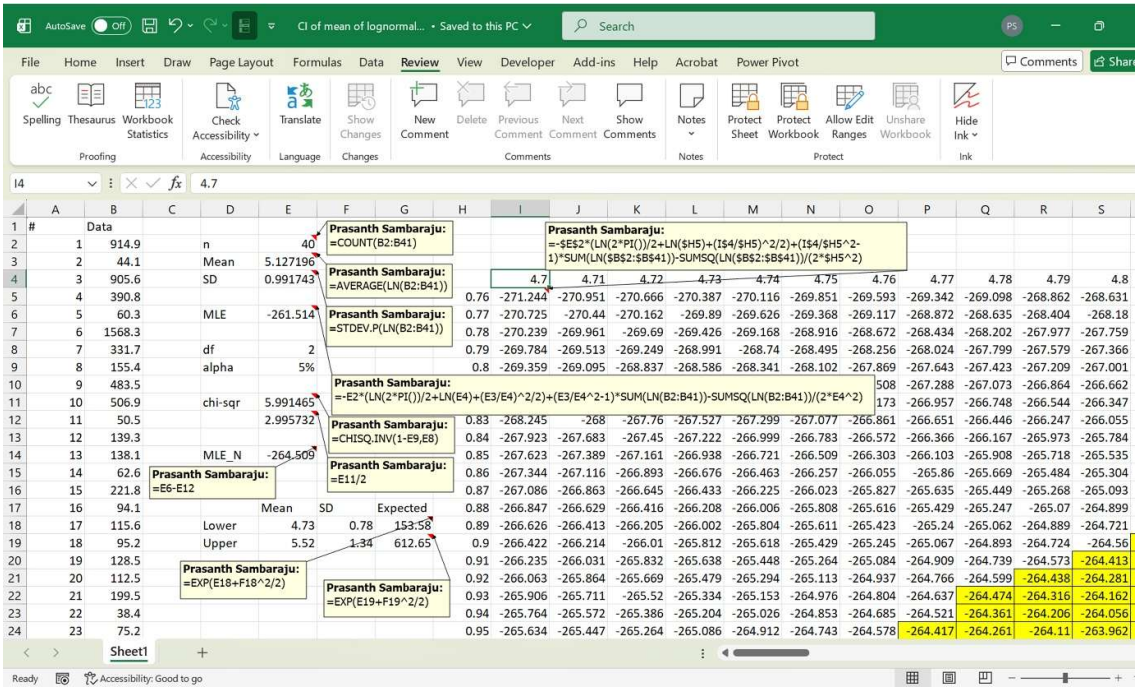


Figure 1: Different formulas used in Microsoft Excel to obtain the loglikelihood ratio confidence intervals.

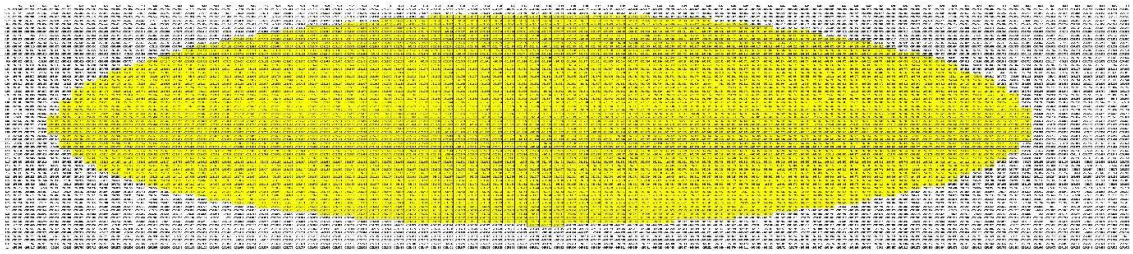


Figure 2: Confidence intervals estimation using loglikelihood ratio method contour plot.

2.1 Carbon Monoxide measurement

Error! Reference source not found. contains nine measurements of carbon monoxide levels in the air, which could be modelled using lognormal distribution are used to obtain the confidence intervals for the mean [5]. The comparison of reported values using different methods with loglikelihood ratio method is included in

Table 3.

Table 2: Carbon monoxide levels in the air at an oil refinery

CO levels	Date
12.5	9/11/90
20	10/4/90
4	12/3/91
20	12/10/91
25	5/7/92
170	8/6/92
15	9/10/92
20	9/22/92
15	3/30/93

Table 3: Lower and upper limits in confidence intervals using the different reported methods [5] and loglikelihood ratio method

Method	Lower 95%	Upper 95%
Naïve approach	9.15	40.95
Cox method	14.15	68.49
Modified Cox method	12.31	78.72
Large-sample approach	-6.11	73.11
Generalized confidence interval	16.65	153.16
Loglikelihood ratio method	9.32	270.01

2.2 Simulation Study

Different sample sizes were generated from a lognormal distribution with parameters $\mu = 5$ and $\sigma = 1$. Confidence intervals for θ using the loglikelihood ratio method were calculated using 1000 replications in Microsoft Excel. A custom written Visual Basic for Applications (VBA) code was used to automate the simulation process.

3. Results and Discussions

For the sample containing 40 observations, the confidence intervals obtained using the loglikelihood ratio method and other reported methods are given in Table 4. Based on these results, it can be concluded that the loglikelihood ratio method described in the manuscript produces the widest confidence intervals.

Table 4: Comparison of different reported methods [5] and loglikelihood ratio method for confidence intervals of mean of lognormal distribution.

S. No.	Method	95% Confidence interval
1	Naïve method	122.24 – 232.29
2	Cox method	190.24 – 409.82
3	Cox method: a modified version	188 – 414.7
4	Based on large-sample theory	178.84 – 371.16
5	Loglikelihood ratio method	153.58 – 612.65

For simulation study samples, the coverage probability obtained using loglikelihood ratio method and other reported methods [5] are given in

Table *5*.

Table 5: Results from the simulation study showing the comparison of percent intervals from different reported methods [5] and loglikelihood ratio method that contains the true parameter value.

Sample size	Naïve approach	Cox method	Modified Cox method	Large sample approach	Generalized CI	Loglikelihood ratio method
5	86.2	87.2	93.5	83	94.1	98.2
10	68.5	91.1	93.9	83.6	93.7	99.3
20	45.2	94.2	95.7	87.9	95.2	99.7
30	24.1	92.6	93.9	85.6	94.6	99.7
50	5.7	95.4	95.5	90.4	95	99.8

4. Conclusions

Based on the results from simulations, the modified Cox method and generalized confidence intervals approach produced coverage probability of 95% for all sample sizes. By using the loglikelihood ratio method described in the manuscript produced confidence intervals with coverage probability greater than 98% for all sample sizes. The major downside with the loglikelihood ratio method is that the time required to run 1000 simulations is about 45 minutes using the custom written VBA code in Microsoft Excel.

Data availability statement: The Microsoft Excel file and the macro enabled Microsoft Excel file used for simulations are available on request.

References

[1]. Chami P, Antoine R., and Sahai A. (2007). On Efficient Confidence Intervals for the Log-Normal Mean. *Journal of Applied Sciences*, 7 (13), 1790-1794.

[2]. Ginos BF. (2009). *BYU Scholars Archive*. https://scholarsarchive.byu.edu/etd/1928. Accessed July 01, 2026.

[3]. HBK (2025). *The Lognormal Distribution*. https://help.reliasoft.com/reference/life_data_analysis/lda/the_lognormal_distribution.html. Accessed July 01, 2026.

[4]. O'Connor, A. N., Modarres, M., and Mosleh, A. (2016). *Probability Distributions Used in Reliability Engineering.*

[5]. Olsson U. (2007). Confidence Intervals for the Mean of a Log-Normal Distribution. *Journal of Statistical Education*, 13 (1). https://doi.org/10.1080/10691898.2005.11910638

[6]. Zaiontz C. (2026). *Real Statistics Using Excel*. https://real-statistics.com/distribution-fitting/distribution-fitting-via-maximum-likelihood/fitting-lognormal-distribution-via-mle/. Accessed July 01, 2026.