

ASYMPTOTIC CONFIDENCE ELLIPSE FOR LOG-NORMAL DISTRIBUTION WITH APPLICATIONS IN ACTUARIAL PRICING

Nassamon Bootwisas¹, Uparittha Intarasat^{1,*} and Pasin Marupanthorn²

Received: September 10, 2024; Revised: January 12, 2025; Accepted: May 30, 2025

Abstract

This study develops and validates asymptotic confidence ellipses for the parameters of the log-normal distribution by leveraging the asymptotic properties of Maximum Likelihood Estimators (MLEs). Monte Carlo simulations across various parameter settings are used to evaluate the performance of these ellipses. The results indicate that they yield accurate parameter estimates with empirical coverage probabilities closely matching the nominal 95% level, ranging from 94.27% to 95.13%, even in small samples. A theoretical lower bound on the sample size required to attain a specified level of precision is also derived and confirmed via simulation. Furthermore, the methodology is applied to health insurance pricing to visualize parameter uncertainty and quantify its impact on premium estimates under a log-normal loss model. The framework allows the derivation of confidence intervals for premiums and estimation of the number of policyholders needed to meet target pricing accuracy.

Keywords: Actuarial Pricing; Asymptotic Statistics; Confidence Ellipse; Log-normal Distribution

¹ Division of Computational Science, Rajamangala University of Technology Suvarnabhumi, Phra Nakhon Si Ayutthaya, 13000.
E-mail: nassamon.b@rmutsb.ac.th; Uparittha.i@rmutsb.ac.th

² QuantCornor Research Laboratory, Department of Mathematics, Faculty of Science, Maejo University, Chiang Mai, 50290.
E-mail: quantfilab@gmail.com

* Corresponding author

DOI: <https://doi.org/10.55766/sujst6597>

Introduction

The confidence ellipse is a type of confidence region generated by the covariance matrix in two dimensions. It can be derived as shown in the following progression. Hotelling (1931) derived the generalized means test, $T^2 = \frac{p(n-1)}{n-p} \times F(p, n-p)$, where p is the number of parameters and F stands for the F distribution. He concluded that T^2 can be used to define an ellipse with a specified confidence level. As the sample size n becomes large, the distribution $2 \times F(2, n-2)$ approximates the χ^2 distribution with 2 degrees of freedom. The empirical probability, P , that a random point falls within this ellipse is determined by the χ^2 distribution with 2 degrees of freedom. This concept inspired the development of confidence ellipses (Fisher, 1925; Hotelling, 1931; Durand, 1954), which for a given confidence level $1-\alpha$ are defined as the set of points (x_1, x_2) that satisfy the equation:

$$(X - \mu)^T \Sigma^{-1} (X - \mu) \leq \chi_{2,\alpha}^2. \quad (1)$$

Here $X = (x_1, x_2)^T$ is a point in the parameter space, $\mu = (\mu_1, \mu_2)^T$ is the mean vector, Σ^{-1} is the inverse of the covariance matrix, and $\chi_{2,\alpha}^2$ is the critical value from the chi-squared distribution with 2 degrees of freedom at the significance level α . Since this confidence ellipse relies on the asymptotic properties of the parameters, it is referred to as the asymptotic confidence ellipse, which will be used throughout the remainder of this paper.

Confidence ellipses are particularly advantageous for several reasons: First, they provide a visual tool for assessing the precision of parameter estimates. Second, they account for the correlation between parameters, which is especially relevant in multivariate settings. This correlation makes the confidence ellipse tighter than the confidence intervals for individual parameters, as discussed in Durand (1954). Finally, confidence ellipses can be precisely derived from the asymptotic distribution of the Maximum Likelihood Estimators (MLEs), which are known to be efficient and unbiased under regular conditions, particularly when Fisher's information matrix is known.

The confidence ellipse is typically used to determine and visualize the uncertainty associated with parameter estimates. Numerous studies, including the work by Friendly *et al.* (2013), provide valuable insights into its meaning and interpretation. Their work explores how ellipses can be utilized to gain visual insights into various statistical methods, highlighting the strong relationships

among statistical methods, matrix algebra, and geometry. Geometrically, the confidence ellipse is centered at the estimated parameter values, with its shape and orientation determined by the covariance matrix of the estimates. The axes of the ellipse correspond to the principal components of the covariance matrix, indicating the directions of maximum and minimum variability.

The confidence ellipse has a long history of development, particularly in statistical modeling. A well-known success case is its use in determining a tighter confidence region for the coefficients of a simple linear regression model, as discussed in Durand (1954) and Adkins and Hill (1990). This concept has been generalized to the confidence ellipsoid for multiple linear regression models. More broadly, it can be applied to determine and visualize the confidence region of parameters in general statistical models such as AR(1) (Sriram, 2001), Distributed Lag Model (Wahba, 1969), Time-Sequential Model (Aras and Gardiner, 1995), and Logistic Regression Model (Chang and Martinsek, 1992). In addition, recent works from 2024-2025, such as Linear Regression with Quasi-Associated Random Errors (Moumen and Bouzidi, 2024) and Linear Regression Models with Deterministic Regressors (Wang and Weyer, 2024), have further advanced the study of confidence ellipses in statistical modeling.

In addition to statistics, the applications of confidence ellipses are found in various research areas. For instance, confidence ellipses are used for similarity quantification by applying them to soil parametric data and observing the intersections of the ellipses (Han *et al.*, 2022). They are also employed to identify the relationship between catchment area and slope in the context of building large dams. The confidence ellipse approach establishes a quantitative relationship between these two variables, which is then used to evaluate site selection for large dams based on topographic characteristics (Liu *et al.*, 2023). Additionally, confidence ellipses are constructed using the stochastic sensitivity function technique to describe the configurational arrangement of stochastic coexistence equilibria, and to estimate the tipping threshold in a stochastic dynamics model of Alzheimer's disease (Gao *et al.*, 2023). The medical research (Zhang *et al.*, 2024) utilized the confidence ellipse for benefit-risk assessment in clinical trials. The statistical test was conducted by separating the zones of the confidence ellipse within the clinical trial data. A survey conducted between 2022 and 2024 revealed that several studies employed

the confidence ellipse to visualize and assess the uncertainty in parameter estimation and data analysis, highlighting its continued use in research practices.

The definition and application of the confidence ellipse discussed earlier are based on the general probability density of the experimental data. The accuracy of the confidence ellipse can be enhanced if the distribution of the samples is known, making it particularly beneficial for constructing confidence ellipses for specific distributions. The (asymptotic) confidence ellipse for specific distributions has been extensively explored throughout history. Examples include the Beta-Poisson dose-response model (Angkanavisal *et al.*, 2012), the Inverse Gaussian distribution (Duangchana and Budsaba, 2014), the Birnbaum-Saunders distribution (Thonglim *et al.*, 2014; Hansoongnern *et al.*, 2018), the Length-biased Inverse Gaussian distribution (Phaphan and Pongsart, 2019), and the Re-parameterized Inverse Gaussian distribution (Phaphan, 2020). The work by Phaphan (2020) represents the most recent development in model-based asymptotic confidence ellipses, published in 2020. The research gap identified is the lack of focus on commonly known distributions, such as the log-normal distribution, which has been widely studied but remains underexplored in this context.

The log-normal distribution (Sen *et al.*, 2010) is a continuous probability distribution of a random variable whose logarithm is normally distributed. This distribution is widely applicable across various fields due to its capability to model skewed data. Formally, a random variable X is said to be log-normally distributed if $\log(X)$ follows a normal distribution with mean μ and variance σ^2 . Mathematically, if X is a log-normal random variable, then:

$$\log(X) \sim N(\mu, \sigma^2), \quad (2)$$

where $N(\mu, \sigma^2)$ denotes a normal distribution with mean μ and variance σ^2 . The probability density function (PDF) of the log-normal distribution is given by:

$$f(x; \mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{(\log x - \mu)^2}{2\sigma^2}\right) \quad (3)$$

for $x > 0$, $\mu \in \mathbb{R}$, $\sigma > 0$.

The log-normal distribution is extensively utilized in various domains. In finance, it models stock prices and financial indices, as asset prices are often assumed to follow a geometric Brownian motion, resulting in log-normal returns (Black and Scholes, 1973; Merton, 1973). In environmental

studies, it describes the distribution of pollutant concentrations (Ott, 1990) and species population sizes (Edden, 1971), accommodating the natural skewness in such data. In reliability engineering, it is employed to model the lifetimes of products and materials (Herrera and Jovani, 2010), providing insights into failure rates and durability. It is extensively applied in insurance for modeling claim sizes due to its ability to capture the positively skewed and heavy-tailed nature of insurance data (Jobe, 1989; Kleiber and Kotz, 2003; Dickson *et al.*, 2020). When claim amounts exhibit significant variability and are bounded below by zero, the log-normal model provides an excellent fit. This model is particularly effective in scenarios where claim sizes span multiple magnitudes, such as in health insurance, property damage, or catastrophe claims. These applications highlight the versatility and importance of the log-normal distribution in modeling real-world phenomena characterized by positive skewness.

The objectives of this research are threefold. First, it aims to construct asymptotic confidence ellipses for the parameters of the log-normal distribution and validate these ellipses through numerical simulations. This study addresses a gap in the existing literature, as previous research has not extensively explored the application of asymptotic confidence ellipses to the log-normal distribution. Second, the research seeks to propose a lower bound for the sample size required to achieve the desired accuracy in parameter estimation. While this work focuses on the log-normal distribution, the methodology developed can be generalized to other distributions, making it a contribution to this field. Third, given the prevalence of the log-normal distribution as a common loss distribution in insurance, this study aims to apply the log-normal asymptotic confidence ellipse to actuarial pricing. This approach provides a foundational framework for assessing uncertainty in model-based actuarial premium pricing and offers insights for future research in the field.

The paper is structured as follows: The section titled "Concept of Constructing the Asymptotic Confidence Ellipse" provides an overview of the concept of asymptotic confidence ellipses, drawing on insights from previous research. The section "Constructing Asymptotic Confidence Ellipses" details the methodology for constructing confidence ellipses specifically for the parameters of the log-normal distribution. The section "Sample Size Determined by the Asymptotic Confidence Ellipse" derives the lower bound for the sample size required under the log-normal distribution, using the properties of the asymptotic confidence ellipse. The section "Numerical Experiment" presents

the results of Monte Carlo simulations conducted to evaluate the accuracy of the constructed confidence ellipses and validate the derived lower bound for the sample size. The section “Applications in Actuarial Pricing” demonstrates the application of the asymptotic confidence ellipse in actuarial pricing, using real-world health insurance data. This section includes an analysis of the uncertainty in the estimation of log-normal parameters and the premium. Furthermore, it derives the confidence interval for the premium projected from the asymptotic confidence ellipse. Finally, the paper concludes with a discussion of the findings and provides suggestions for future research directions in the concluding section.

Concept of Constructing the Asymptotic Confidence Ellipse

This section theoretically discusses the construction of the asymptotic confidence ellipse, along with the steps involved.

Theoretical Outline for Asymptotic Confidence Ellipse

To discuss the construction of an asymptotic confidence ellipse for an arbitrary parameter estimator $\hat{\theta}$ we begin with the general framework of confidence regions in multivariate statistics. Here p is a positive number representing the number of parameter in the model. If $p = 2$, it correspond to the derivation of the confidence ellipse. Given a parameter vector $\theta \in \mathbb{R}^p$ and its estimator $\hat{\theta}$ assume that $\hat{\theta} \sim N(\theta, \Sigma)$, where Σ is the covariance matrix of $\hat{\theta}$ either known or estimated. The general form of the confidence region for θ at confidence level $1-\alpha$ is:

$$(\hat{\theta} - \theta)^T \Sigma^{-1} (\hat{\theta} - \theta) \leq \chi_{p,\alpha}^2, \quad (4)$$

where $\chi_{p,\alpha}^2$ denotes the critical value from the chi-squared distribution with p degrees of freedom at the significance level α .

The brief proof is discussed as follows. Assume $\hat{\theta} \sim N(\theta, \Sigma)$. Consider the transformation

$$z = \Sigma^{-\frac{1}{2}} (\hat{\theta} - \theta), \quad (5)$$

where $\Sigma^{-\frac{1}{2}}$ is the matrix square root of the inverse of Σ . This transformation normalizes $\hat{\theta}$ such that z follows a standard multivariate normal distribution:

$$Z \sim N(0, I), \quad (6)$$

where I_p is the $p \times p$ identity matrix. The squared Mahalanobis distance, which is the left-hand side of Equation (5), can be expressed as:

$$(\hat{\theta} - \theta)^T \Sigma^{-1} (\hat{\theta} - \theta) = z^T z. \quad (7)$$

Since z follows a standard normal distribution, $z^T z$ follows a chi-squared distribution with p degrees of freedom. Therefore,

$$z^T z \sim \chi_p^2. \quad (8)$$

The cumulative distribution function (CDF) of the chi-squared distribution χ_p^2 is denoted as $P(\chi_p^2 \leq \chi_{p,\alpha}^2) = 1 - \alpha$, where $\chi_{p,\alpha}^2$ is the critical value corresponding to the confidence level $1-\alpha$. Consequently, the confidence region for the parameter vector θ is given by:

$$P\left((\hat{\theta} - \theta)^T \Sigma^{-1} (\hat{\theta} - \theta) \leq \chi_{p,\alpha}^2\right) = 1 - \alpha. \quad (9)$$

The resulting ellipse represents the region within which the true parameter values (θ_1, θ_2) are expected to lie with a specified level of confidence. The orientation and shape of the ellipse provide insights into the correlation and variability of the parameter estimates. A larger ellipse indicates greater uncertainty, while the tilt of the ellipse reveals the direction of correlation between the parameters. In addition, the orientation of the ellipse indicates the correlation between parameters. If the ellipse is elongated along certain axes, it suggests a stronger correlation between those parameters.

Steps for Constructing the Asymptotic Confidence Ellipse

Following methodologies detailed in (Angkanavisal *et al.*, 2012; Duangchana and Budsaba, 2014; Thonglim *et al.*, 2014; Hansoongnern *et al.*, 2018; Phaphan and Pongsart, 2019; Phaphan, 2020), after selecting the distribution, the construction of the confidence ellipse involves the following steps:

Estimate the Parameter Vector: Compute the point estimates $\hat{\theta}_1$ and $\hat{\theta}_2$ using the chosen estimation method (e.g., Maximum Likelihood Estimator (MLE) or Method of Moments (MOM)), ensuring the estimator adheres to asymptotic normality assumptions necessary for deriving the asymptotic covariance matrix.

Estimate the Asymptotic Covariance

Matrix: Compute the asymptotic covariance matrix Σ_{asym} of the estimators. This is commonly done using the multivariate Delta method, which approximates the covariance of a transformed estimator based on the gradient (Jacobian) of the transformation function and the covariance of the original estimator.

Under suitable regularity conditions, if $\hat{\theta}_n \xrightarrow{d} \theta$ and g is differentiable at θ , then

$$\sqrt{n}(g(\hat{\theta}_n) - g(\theta)) \xrightarrow{d} N(0, \nabla g(\theta)^T \Sigma \nabla g(\theta)), \quad (10)$$

where $\nabla g(\theta)$ is the gradient vector of g evaluated at θ , and Σ is the covariance matrix of θ . The asymptotic covariance matrix of the transformed variable is denoted by $\Sigma_{asym} = \nabla g(\theta)^T \Sigma \nabla g(\theta)$. For some special classes of estimators, theoretical proofs for the asymptotic properties have been established.

In the case of the MLE, this typically involves computing the inverse of the Fisher information matrix $I(\theta)$. Under the Cramer-Rao-Frechet Regularity conditions and the uniform continuity condition, the MLE $\hat{\theta}_{MLE} = [\hat{\mu}_{MLE}, \hat{\sigma}_{MLE}^2]^T$ is asymptotically normal as $n \rightarrow \infty$ (Fisher, 1922):

$$\sqrt{n}(\hat{\theta}_{MLE} - \theta) \xrightarrow{d} N(0, \Sigma_{MLE}), \quad (11)$$

where $\Sigma_{MLE} = I(\theta)^{-1}$ and $I(\theta)$ is the Fisher information matrix, defined as:

$$I(\theta)_{i,j} = -E \left[\frac{\partial^2}{\partial \theta_i \partial \theta_j} \log f(X; \theta) | \theta \right]. \quad (12)$$

Compute the Squared Mahalanobis

Distance: Evaluate the Mahalanobis distance:

$$D(\theta, \hat{\theta}) = (\hat{\theta} - \theta)^T \Sigma_{asym}^{-1} (\hat{\theta} - \theta). \quad (13)$$

Here, $\Sigma_{asym} = \frac{1}{n} \Sigma_{MLE}$, where n is derived from the multiplication of $\sqrt{n}$ by the asymptotic terms of $(\hat{\theta} - \theta)^T$ and $(\hat{\theta} - \theta)$.

Determine the Confidence Region:

Construct the confidence region by comparing the computed squared Mahalanobis distance to

$\chi_{2,\alpha}^2$, which defines the boundary of the confidence ellipse. The confidence region is the set,

$$\left\{ \hat{\theta} \left| (\hat{\theta} - \theta)^T \Sigma_{asym}^{-1} (\hat{\theta} - \theta) \leq \chi_{2,\alpha}^2 \right. \right\}.$$

Constructing Asymptotic Confidence Ellipses

Construction of asymptotic confidence ellipses requires the point estimators and their asymptotic normality. This section delves into the methods of parameter estimation for the log-normal distribution, with a particular focus on MLE, MOM, and Method of Percentiles (MOP). The asymptotic properties of these estimators are explored, emphasizing their asymptotic normality.

In this work, the asymptotic confidence ellipse of the MLE is considered, as it satisfies the regularity conditions outlined in the Appendix. Unlike some distributions (Thonglim *et al.*, 2014; Hansoongnern *et al.*, 2018), where the covariance matrix of the limiting distribution of the parameters derived from the MOM can be simplified, the covariance matrix of the log-normal distribution yields a complex analytical formula. However, this complexity poses significant challenges in interpreting the geometry of the confidence ellipse. Similarly, in the MOP, where the efficiency of the estimator depends on specific parameters such as quantiles, numerical comparisons become difficult. Furthermore, this work does not focus on a comparison between estimation methods. Instead, this study emphasizes the geometric interpretation of the confidence ellipse with varying parameters of the MLE, rather than comparing different types of estimators.

MLE is a widely used method for estimating the parameters of a probability distribution by maximizing the likelihood function. For a sample $X_1, X_2, \dots, X_n$ drawn from a log-normal distribution, the log-likelihood function is given by:

$$L(\mu, \sigma^2) = -\frac{n}{2} \log(2\pi\sigma^2) - \sum_{i=2}^n \log X_i - \frac{1}{2\sigma^2} \sum_{i=1}^n (\log X_i - \mu)^2. \quad (14)$$

Maximizing this log-likelihood function with respect to μ and σ^2 yields the following estimators:

$$\hat{\mu}_{MLE} = \frac{1}{n} \sum_{i=1}^n \log X_i. \quad (15)$$

$$\hat{\sigma}_{MLE}^2 = \frac{1}{n} \sum_{i=1}^n (\log X_i - \hat{\mu}_{MLE})^2. \quad (16)$$

The details of the calculation can be found in literary textbooks such as in (Johnson *et al.*, 1994). Next, considering the asymptotic distribution of the estimator using the Lindeberg-Levy Central Limit Theorem (CLT), Explicit calculation reveals $I_{\mu\mu} = \frac{1}{\sigma^2}$, $I_{\sigma\sigma} = \frac{2}{\sigma^2}$, and $I_{\mu\sigma} = I_{\sigma\mu} = 0$. Therefore, the Fisher information matrix for a sample $X_1, X_2, \dots, X_n$ from the log-normal distribution is:

$$I(\mu, \sigma) = \begin{pmatrix} \frac{1}{\sigma^2} & 0 \\ 0 & \frac{2}{\sigma^2} \end{pmatrix}. \quad (17)$$

More details on the calculation can be found in (Kleiber and Kotz, 2003). The asymptotic covariance matrix can be calculated as the inverse of the Fisher information matrix:

$$\Sigma_{MLE} = I^{-1}(\mu, \sigma) = \begin{pmatrix} \sigma^2 & 0 \\ 0 & \frac{\sigma^2}{2} \end{pmatrix}. \quad (18)$$

Next, we calculate the squared Mahalanobis distance, which is given by:

$$D_{MLE}(\theta, \hat{\theta}_{MLE}) = \frac{n}{\sigma^2} \left((\mu - \hat{\mu}_{MLE})^2 + 2(\sigma - \hat{\sigma}_{MLE})^2 \right). \quad (19)$$

Finally, the asymptotic confidence ellipse with the confidence level of α can be written in the form of:

$$\frac{n}{\sigma^2} \left((\mu - \hat{\mu}_{MLE})^2 + 2(\sigma - \hat{\sigma}_{MLE})^2 \right) \leq \chi_{2,\alpha}^2. \quad (20)$$

The size of the asymptotic confidence ellipse is influenced by both the sample size n and the variance σ^2 of the underlying distribution. Specifically, the ellipse's size is inversely proportional to the sample size, meaning that as n increases, the confidence ellipse shrinks. This reduction in size reflects the increased precision of the parameter estimates, as larger samples provide more information about the underlying distribution, leading to less uncertainty. On the other hand, the size of the ellipse is directly proportional to the variance σ^2 . When the variance is higher, the ellipse expands, indicating greater uncertainty in the estimates of the parameters. This is because higher variance implies more variability

in the data, making it more challenging to precisely estimate the true values of the parameters. Therefore, a larger sample size decreases the uncertainty (shrinking the ellipse), while higher variance increases the uncertainty (expanding the ellipse).

Sample Size Determined by the Asymptotic Confidence Ellipse

According to the confidence ellipse in Equation (20), the farthest pair of (μ, σ) is located at the vertices of the main axis with the distance of

$$E \leq \max_{\theta} D(\theta, \hat{\theta}) = \sqrt{\frac{\chi_{2,\alpha}^2}{n}} \sigma. \quad (21)$$

This represents the maximum error in estimating the true parameters. Here, D represents the squared Mahalanobis distance, reflecting the distance between the parameter θ and its estimator $\hat{\theta}$. To ensure that this error does not exceed a specified value E the sample size n must satisfy the condition

$$n \geq \frac{\chi_{2,\alpha}^2 \sigma^2}{E^2}. \quad (22)$$

This inequality shows that the required sample size is directly proportional to the variance σ^2 and the critical value $\chi_{2,\alpha}^2$, and inversely proportional to the square of the allowable error E . Consequently, achieving a smaller maximum estimation error epsilon necessitates a larger sample size n , especially in the presence of high underlying variability σ^2 . This relationship underscores the trade-off between data quantity and estimation precision inherent in asymptotic inference.

Numerical Experiment

To validate the confidence ellipse obtained in the previous section, a Monte Carlo simulation was applied. The simulation was conducted in R (R Core Team, 2022), version 4.2.2, across nine different cases of the log-normal parameters from the set $M \times S$, where $M = \{0, 1, 2\}$ for shape parameters and $S = \{0.2, 0.5, 1\}$ for scale parameters. Here, $M \times S$ represents the Cartesian product of sets, resulting in ordered pairs of parameters. These parameters result in variations in the shape of the log-normal distributions. The number of samples was set to $n = 10, 100, 1,000$. In each iteration, the sampling

was repeated 1,000 times. The significance level, α , was set to 0.05.

In the case of a low mean and low variability, the distribution has a peak close to 1 and will be relatively narrow. When the mean is high but variability remains low, the distribution's peak shifts significantly to the right compared to the low mean case but stays similarly narrow. In a scenario with moderate mean and high variability, the distribution's peak resembles that of the high mean, low variability case, but the distribution is more spread out, exhibiting a heavier tail. These different scenarios illustrate how changes in the mean and variability affect the shape and spread of the log-normal distribution, providing insight into the robustness of the confidence ellipse under varying conditions.

Figure 1 displays the confidence ellipses for the log-normal distributions generated by varying the shape and scale parameters, with a sample size of $n = 10$ and $n = 1,000$. The shape of the confidence ellipse consistently aligns along the horizontal axis because the scaling factor of the scale parameter term is always smaller than that of the shape parameter, as shown in the ellipse Equation (19). The scale and shape parameters are independent when considering the geometry of the ellipse due to the non-rotation of the ellipse. Most of the log-normal parameters estimated by maximum likelihood lie within the confidence ellipse.

Since the behavior of the simulation for $n = 10$, 100, and 1,000 did not show significant differences, the case of $n = 100$ was not visualized. However, the R code used to generate any case of the log-normal parameters is available in the GitHub

repository:

<https://github.com/QuantFILab/Confidence-Region>

Table 1 shows the simulation results for various parameters and cases. The values μ_{true} and σ_{true} represent the true parameters of the log-normal distribution, while μ_{est} and σ_{est} are the average values over 1,000 samples of the scale and shape parameters estimated by the maximum likelihood method. The coverage represents the percentage of samples that lie inside the ellipse (N_{in}) relative to the total number of samples (N), which is 1,000 in our study, i.e., Coverage = $100 \times \frac{N_{in}}{N}$. The coverage represents the percentage of sample points that lie within the confidence ellipse. This ratio serves as an empirical estimate of the confidence level α in Equation (20) and correspondingly reflects the significance level $1-\alpha$. The last column presents the percentage relative error, computed based on the exact confidence level of 0.95 (corresponding to 95% coverage).

The accuracy of the MLEs for the log-normal distribution increases as the number of samples increases, following the asymptotic property of the MLE. For the log-normal distribution, the MLE demonstrates very good convergence properties for its confidence ellipse, as indicated by the simulation results. This can be observed from the coverage, which approximates the significance level of 0.05 even when the number of samples n is as small as 10. This can be seen from the coverage values,

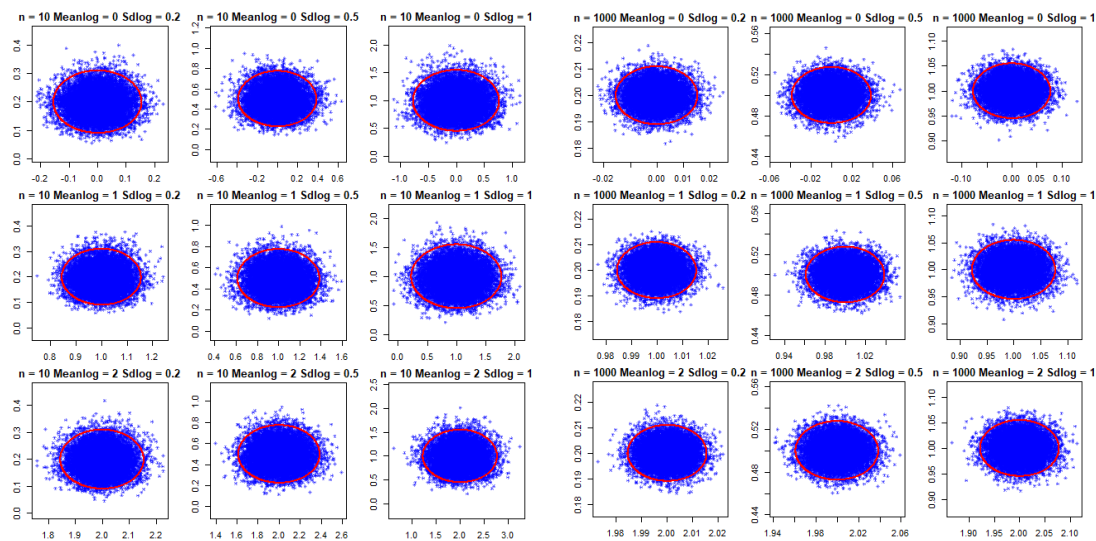


Figure 1. Confidence ellipse for the log-normal distributions generated by varying the shape and scale parameters $\mu \in \{0, 1, 2\}$ and $\theta \in \{0.2, 0.5, 1\}$, with a sample size of $n = 10$ and $n = 1,000$

Table 1. Comparison of estimated log-normal parameters and empirical coverage via Monte Carlo simulation

n	μ_{true}	σ_{true}	μ_{est}	σ_{est}	Coverage	Error
10.0000	0.0000	0.2000	0.0002	0.1938	94.5300	-0.4947
		0.5000	0.0012	0.4859	94.6200	-0.4000
		1.0000	0.0023	0.9750	94.9400	-0.0632
	1.0000	0.2000	1.0004	0.1943	94.2700	-0.7684
		0.5000	1.0016	0.4848	94.6000	-0.4211
		1.0000	0.9995	0.9726	94.7300	-0.2842
	2.0000	0.2000	1.9997	0.1947	94.3900	-0.6421
		0.5000	2.0019	0.4861	94.9300	-0.0737
		1.0000	1.9947	0.9718	94.5600	-0.4632
		2.0000	2.0000	2.0000	94.9800	-0.0211
100.0000	0.0000	0.2000	-0.0001	0.1992	94.7300	-0.2842
		0.5000	-0.0005	0.4990	94.9100	-0.0947
		1.0000	0.0019	0.9971	94.8500	-0.1579
	1.0000	0.2000	1.0001	0.1994	94.9500	-0.0526
		0.5000	0.9994	0.4985	94.9700	-0.0316
		1.0000	0.9997	0.9976	94.8700	-0.1368
	2.0000	0.2000	2.0002	0.1996	95.1300	0.1368
		0.5000	2.0000	0.4980	94.6000	-0.4211
		1.0000	2.0000	0.9978	94.9800	-0.0211
		2.0000	2.0000	2.0000	94.9800	-0.0211
1,000.0000	0.0000	0.2000	0.0000	0.1999	94.7600	-0.2526
		0.5000	0.0001	0.4999	95.2700	0.2842
		1.0000	0.0004	0.9997	95.3300	0.3474
	1.0000	0.2000	1.0000	0.2000	95.0400	0.0421
		0.5000	0.9997	0.4999	94.8200	-0.1895
		1.0000	1.0000	0.9995	95.2500	0.2632
	2.0000	0.2000	1.9999	0.1998	95.6100	0.6421
		0.5000	2.0002	0.4999	95.2300	0.2421
		1.0000	2.0000	0.9998	94.9900	-0.0105
		2.0000	2.0000	2.0000	94.9900	-0.0105

which closely approximate the nominal confidence level of 95.00% (corresponding to a significance level of 0.05), with overall relative errors ranging from -0.7684% to 0.6421%. Notably, even when the sample size is as small as $n = 10$, the relative errors remain within the range of -0.7684% to -0.0737%.

In addition to the visualization of the asymptotic confidence ellipse, the error bound with a given number of samples in Equation (22) is verified for two cases: $\mu = 0$, $\sigma = 0.2$, and $\mu = 0$, $\sigma = 1$. The simulation is configured as follows: The value of n ranges from 100 to 10,000, increasing in steps of 100. For each value of $n = 1,000$ sample groups are generated, where each sample group contains n samples. After fitting the samples using the maximum likelihood method to the log-normal density, the estimators are averaged, and their standard deviation is computed. The results are presented in Figure 2.

Figure 2 illustrates the error bound, $E \max_{\hat{\theta}} D(\theta, \hat{\theta}) \sqrt{\frac{\chi^2_{2,\alpha}}{n \max}}$ as a blue line. The average of the error, E , for each n is visualized with red dots, referred to as the actual error, along with its standard deviation band. As observed, the average error and its standard deviation do not exceed the bound for $n \geq 100$. These results support the theoretical bound provided in Equation (22) for both cases. As expected, the standard deviation width becomes

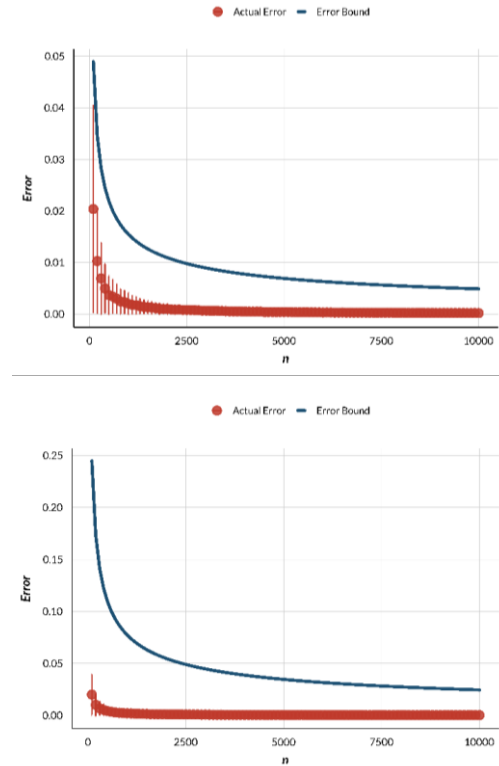


Figure 2. Comparison of actual error and error bound across sample sizes, where $\mu = 0$, $\sigma = 0.2$ (Upper) and $\mu = 0$, $\sigma = 1$ (Lower)

smaller as n increases, reflecting the statistical justification of the CLT.

Applications in Actuarial Pricing

The asymptotic confidence ellipse is commonly used to define the uncertainty of dependent estimators, providing both a visual representation and asymptotic accuracy as the sample size increases. This case study focuses on health insurance pricing, where the medical costs billed to the health insurance provider are known.

The confidence ellipse can be utilized to identify and visualize the uncertainty of estimators derived from the health insurance dataset. This visualization aids in decision-making and risk assessment in actuarial pricing. Furthermore, the confidence ellipse facilitates the mapping of tighter confidence intervals for premiums calculated using the Standard Deviation Principle (SDP) under a log-normal loss model. Unlike the naive method, this approach considers the dependency between parameters. The length of the confidence interval reflects the uncertainty of the premium, and reducing this length can lower the reserves required for managing uncertain events.

Data Description

The analysis in this study is based on the publicly available Medical Cost Personal Datasets, provided by Miri Choi on Kaggle. This dataset contains information on 1,338 individuals, detailing their demographic, lifestyle, and healthcare-related features, as well as the medical costs billed to their health insurance.

This dataset is particularly relevant for predictive modeling and statistical analysis in healthcare. It enables researchers and practitioners to explore the relationships between individual factors, such as age, BMI, smoking habits, and medical costs. By analyzing this dataset, valuable insights can be derived to inform healthcare cost patterns, assess risk, and develop pricing strategies for health insurance policies. The dataset is publicly available for download from Kaggle at <https://www.kaggle.com/datasets/mirichoi0218/insurance>. The actuarial pricing of the health insurance policy can be determined by analyzing the distribution, expectation, and standard deviation of the data in the last column, which represents the medical costs billed to health insurance.

Health Insurance Policy Pricing

In this study, community rating (Pauly and Herring, 2002) is considered. It is a pricing approach in insurance where the insurer charges the same

premium to all policyholders, regardless of individual risk factors such as age, gender, or health conditions. This method is commonly applied in social insurance or group health insurance plans, with an emphasis on promoting fairness and accessibility rather than risk-based pricing.

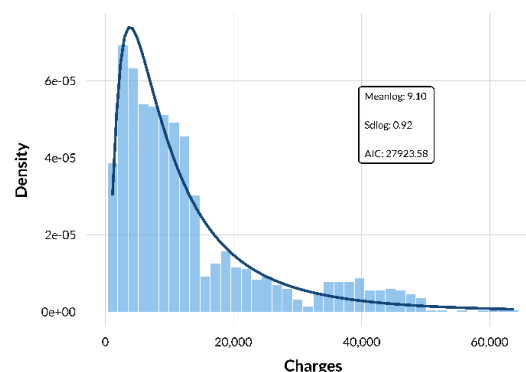


Figure 3. Histogram of charges with log-normal density fit

For this approach, the community's charges are analyzed under the following assumptions. First, the charges are assumed to be independent and follow a log-normal distribution with a single set of parameters. The frequency and density of all charges are treated as a single group. The distribution of the charges, along with the fitted log-normal density, is shown in Figure 3. The MLEs for the log-scale parameters are obtained as meanlog = 9.10 and sdlog = 0.92 which correspond to the original scale parameters $\mu = \$13,644.72$ and $\sigma = \$15,722.49$. The mean and variance of a log-normal distribution on the original scale are derived from its properties, where the mean is given by $\mu = e^{mean\ log + \frac{sd\ log^2}{2}}$ and the variance by $\sigma^2 = (e^{sd\ log^2} - 1)e^{2mean\ log + \frac{sd\ log^2}{2}}$. These formulas enable the calculation of the original scale parameters using the log-scale parameters meanlog and sdlog. The Akaike Information Criterion (AIC) is employed for model selection. The data are fitted to several continuous distributions whose supports lie in the interval $(0, \infty)$: Log-normal, Gamma, Weibull, Exponential, and Log-logistic distributions. The corresponding AIC values are 27,923.58, 27,996.18, 28,027.32, 28,082.05, and 27,986.69, respectively. Among these, the Log-normal distribution yields the lowest AIC, indicating the best statistical fit under the principle of parsimony, and suggesting that it provides the most appropriate model for the data.

The second step is to calculate the community rating premium. For general insurance, one of the standard principles of pricing is derived from the premium principle. The SDP (Deprez and Gerber, 1985; Kreps, 1990; Young, 2004), a widely recognized premium principle, is a method for calculating insurance premiums by incorporating a risk margin proportional to the standard deviation of the claim distribution. This approach ensures that the premium accounts for both the average claim cost and the uncertainty associated with claims. The premium is calculated as:

$$P = E[X] + a\sqrt{\text{Var}(X)}, \quad (23)$$

where $E[X]$ is the expected p value of the claim amount, $\text{Var}(X)$ is the variance of the claim amount, and $\sqrt{\text{Var}(X)}$ is its standard deviation. The parameter $a > 0$ is the safety loading factor that reflects the insurer's tolerance for risk. A higher a corresponds to a more conservative premium, while a lower a results in a more competitive premium. Note that this premium does not account for the time discount factor, which can be incorporated after determining the duration of the payment.

Using the SDP in Equation (23), the point estimator of the actuarially fair price is calculated as

$$P_{com}(a) = 13,644.72 + 15,722.49a. \quad (24)$$

Assessment of Premium Pricing Uncertainty

This step is to assess the uncertainty, which can be considered a risk, using the concept of the asymptotic confidence ellipse. Figure 4 displays the asymptotic confidence ellipses with $\alpha = 0.05$, $a = 0.1$, $\mu \log = 9.10$ and $\sigma \log = 0.92$, for $n \in \{1,388, 5,000, 10,000, 20,000\}$. The gradient color represents the value of the SDP. The estimators have a theoretical probability of 5% of lying outside the confidence ellipse. For the baseline case, $n = 1,388$, the estimator of the SDP can vary from approximately \$16,000 to \$17,000, resulting in a price relative error of about 5.14% to 11.72%. Conceptually, the uncertainty in the premium estimator may lead to increased uncertainty in the actuarial reserves required (Kreps, 1990; Rehman and Klugman, 2010), which are the financial provisions insurers allocate to cover future claims. As shown in Figure 4, increasing the sample size reduces the uncertainty in the SDP pricing, leading to more precise and reliable premium estimates.

The uncertainty of the SDP can be assessed numerically using the confidence interval of the estimator. The confidence intervals for the parameters μ and σ of the log-normal

distribution are derived using the asymptotic normality of the MLEs. For a given confidence level $(1-\alpha)$, the intervals are constructed as follows, using standard asymptotic normal theory following Newcombe (2012): $\hat{\mu} \pm z_{\alpha/2}/\sqrt{n}\sigma$ and $\hat{\sigma} \pm z_{\alpha/2}/\sqrt{n}\frac{\sigma}{\sqrt{2}}$, where $\hat{\mu}$ is the MLE of μ , σ is the standard deviation of the underlying normal distribution, and $z_{\alpha/2}$ is the critical value of the standard normal distribution corresponding to the desired confidence level. To obtain the confidence interval for the premium calculated using the SDP under the log-normal loss distribution, these MLEs are substituted into Equation (23). The resulting naive confidence interval is given by:

$$\begin{aligned} & \left(\hat{\mu} \pm \frac{z_{\alpha/2}}{\sqrt{n}} \cdot \sigma \right) + a \left(\hat{\sigma} \pm \frac{z_{\alpha/2}}{\sqrt{n}} \cdot \frac{\sigma}{\sqrt{2}} \right) \\ & = \hat{P} \pm \frac{z_{\alpha/2}\sigma}{\sqrt{n}} \left(1 + \frac{a}{\sqrt{2}} \right), \end{aligned} \quad (25)$$

Where $\hat{P}$ is the estimated premium under the SDP. However, this approach does not account for the dependency between the parameters μ and σ , which may lead to an underestimation or overestimation of the confidence interval width.

Another approach for constructing the confidence interval involves accounting for the joint distribution of the log-normal estimators, where the confidence ellipse represents the confidence region of these estimators. The Delta method, as detailed in the Appendix, is employed to derive the limiting distribution of the SDP. The limiting distribution is expressed as:

$$P \sim N \left(\mu + a\sigma, \sigma^2 \left(1 + \frac{a^2}{2} \right) \right). \quad (26)$$

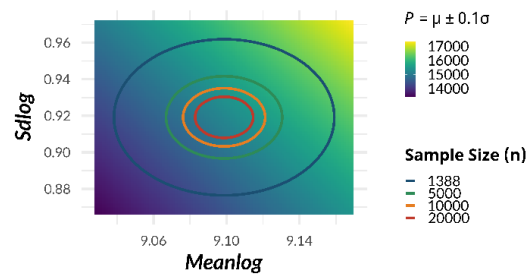


Figure 4. Asymptotic confidence ellipses with $\alpha = 0.05$, $a = 0.1$, $\mu \log = 9.10$ and $\sigma \log = 0.92$, with the gradient color representing the value of the SDP

Using this result, the projected confidence interval for $\hat{P}$ is given by:

$$\hat{P} \pm \frac{z_{\alpha/2}\sigma}{\sqrt{n}} \left(1 + \frac{a^2}{\sqrt{2}} \right) \quad (27)$$

which is tighter than the naive approach, thereby reducing the uncertainty described in Equation (25) for $a > 0$. The proof of this result is provided in the Appendix. For the given data, the projected confidence interval is:

$$\begin{aligned} & \hat{P}_{com} \pm \frac{z_{\alpha/2}\sigma}{\sqrt{n}} \left(1 + \frac{a^2}{\sqrt{2}} \right) \\ &= (13,644.72 + 15,722.49a) \pm 827.13 \sqrt{1 + \frac{a^2}{2}}. \end{aligned} \quad (28)$$

which results in a confidence interval of [\$14,387.77, \$16,046.16] when $a = 0.1$, with the estimated premium $\hat{P}_{com} = \$15,216.97$. This level of uncertainty may not be negligible for businesses with smaller capital, depending on the company's characteristics, as it could affect the insurer's ability to set accurate premiums and allocate adequate reserves.

Reducing Premium Pricing Uncertainty

Reducing uncertainty through larger sample sizes (e.g., increased policy sales) or more refined modeling techniques can enhance pricing accuracy and minimize the risk of financial shortfalls.

The projected confidence interval in Equation (27) establishes the relationship between the number of samples and the length of the confidence interval, given by:

$$n \geq \left(\frac{2z_{\alpha/2}\sigma \sqrt{1 + \frac{a^2}{2}}}{L} \right)^2, \quad (29)$$

where L denotes the length of the confidence interval. This relationship is grounded in the Law of Large Numbers (LLN), which is fundamental to insurance pricing. The logarithmic transformation of this equation is expressed as:

$$\log_{10}(n) \geq -2\log_{10}(L) + C, \quad (30)$$

where $C = 2 \log_{10} \left(2z_{\alpha/2}\sigma \sqrt{1 + \frac{a^2}{2}} \right)$. The slope of the equation, -2 , indicates that improving the precision of L by one digit requires the number of samples n to increase by two orders of magnitude.

If the insurer seeks to ensure that the confidence interval for the estimated SDP has a width less than 0.001, the required number of policies must be significantly increased to achieve this level of precision. For a precision of one decimal places in the confidence interval length L , the interval must be reduced by a factor of 10^1 . According to the relationship in Equation (30), the number of samples n must increase by $2 \times 1 = 2$ orders of magnitude. For instance, if the initial sample size is $n_0 = 1,000$, the required sample size becomes:

$$n = n_0 \times 10^2 = 1,000 \times 10^2 = 100,000$$

Therefore, to achieve the desired accuracy of three decimal places in the SDP estimate, the insurer would need to sell approximately 100,000 policies. Note that the sufficient conditions for the formula to hold require the loss distribution to be independent and identically distributed (i.i.d.), as stipulated by the LLN (Kolmogorov, 1950).

Conclusions

The log-normal distribution is used in various fields to model skewed data, such as financial returns, pollutant concentrations, and product lifetimes. In this paper, we focus on the construction and validation of asymptotic confidence ellipses for the parameters of the log-normal distribution. We begin by deriving the theoretical framework for confidence ellipses based on the asymptotic properties of the MLEs. The derived confidence ellipses provide a visual and quantitative method for assessing the precision of parameter estimates while accounting for their correlation.

To evaluate the performance of the proposed confidence ellipses, we conduct a series of Monte Carlo simulations across different parameter settings of the log-normal distribution, varying the shape and scale parameters. The simulations were implemented in R (version 4.2.2) and covered a range of sample sizes ($n = 10, 100$, and $1,000$). The results indicate that the MLEs for the log-normal distribution exhibit strong convergence properties, with the coverage of the confidence ellipses closely approximating the nominal

significance level of 0.05, even for small sample sizes.

The lower bound of the sample size required to achieve the desired accuracy for a given error in the squared Mahalanobis distance under the log-normal distribution is also provided. The formula is derived from the maximum distance between the center of the asymptotic confidence ellipse (i.e., the MLEs) and the farthest point on the ellipse. A simulation study is conducted to verify the validity of this lower bound.

In the context of actuarial pricing for health insurance, the asymptotic confidence ellipse proves to be a valuable tool, providing a means to visualize and quantify the joint uncertainty of the log-normal parameters. This study highlights its effectiveness in constructing a tighter confidence intervals for premiums calculated using the SDP, significantly reducing uncertainty compared to naive methods. Furthermore, the relationship between sample size n and confidence interval precision is elucidated, emphasizing the substantial increase in sample size required to achieve higher levels of precision. This approach not only improves the accuracy of premium estimates but also facilitates better reserve planning, supporting robust and reliable actuarial decision-making.

The asymptotic confidence ellipse studied in this work relies on the assumption that the data follows an i.i.d. log-normal distribution, as required by the Lindeberg-Levy CLT. Nonetheless, this framework may be limited in dynamic environments where the underlying distribution or its parameters evolve over time, underscoring the need for future research on adaptive or non-stationary modeling techniques.

Appendix

Regularity Conditions for the Estimator

The regularity conditions (Lehmann and Casella, 1998; Casella and Berger, 2002) required for the efficiency and unbiasedness of an estimator are as follows. First, the true parameter a must lie in the interior of the parameter space, ensuring that the estimation process does not occur at the boundary where constraints may limit the estimator's performance. Second, the likelihood function $L(a; X)$ must be sufficiently smooth with respect to a , allowing for differentiation. This smoothness ensures that the necessary mathematical operations, such as calculating derivatives, are well-defined. Third, the parameter a must be identifiable from the data, which is guaranteed by the Fisher information $I(a)$. The Fisher information is defined as

$$I(a) = -E \left[\frac{\partial^2}{\partial a^2} \log L(a; X) \right] \quad (31)$$

and it must be finite and positive to ensure that the parameter is uniquely determined by the observed data. Fourth, the variance of the score function, represented as the first derivative of the log-likelihood, must be finite:

$$E \left[\left(\frac{\partial}{\partial a} \log L(a; X) \right)^2 \right] < \infty \quad (32)$$

Finally, the sample size n must be sufficiently large, allowing the asymptotic properties of the estimator to hold. These conditions collectively ensure that the estimator achieves desirable statistical properties, such as efficiency and unbiasedness, under the specified model.

Lindeberg-Levy Central Limit Theorem

The Lindeberg-Levy CLT (Feller, 1971; Wackerly *et al.*, 2014; Larsen and Marx, 2017; Hogg *et al.*, 2019) provides a foundational result in probability theory and statistics, stating that the sum (or average) of a large number of i.i.d. random variables with finite mean and variance approaches a normal distribution, regardless of the original distribution of the variables. Formally, suppose $X_1, X_2, \dots, X_n$ are i.i.d. random variables with mean μ and variance σ^2 . The standardized sum is given by:

$$Z_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{X_i - \mu}{\sigma} \quad (33)$$

As $n \rightarrow \infty$, Z_n converges in distribution to a standard normal random variable, i.e.,

$$Z_n \xrightarrow{d} N(0,1) \quad (34)$$

This result explains why the normal distribution appears so frequently in real-world data and underpins statistical methods such as hypothesis testing and confidence intervals.

Delta Method

The Delta Method (van der Vaart, 2000; Casella and Berger, 2002) is a foundational technique in asymptotic statistics, used to approximate the distribution of a smooth transformation of an asymptotically normal estimator. In multivariate contexts, it is particularly useful for deriving

the variance-covariance structure of functions involving multiple parameters. Suppose $\hat{a}_n$ is a vector of estimators for the parameter vector a , such that:

$$\sqrt{n}(\hat{a}_n - a) \xrightarrow{d} N(0, \Sigma_a), \quad (35)$$

where Σ_a is the asymptotic covariance matrix of $\hat{a}_n$. Let $g(a)$ be a differentiable function of a . Using a first-order Taylor expansion around a , $g(\hat{a}_n)$ can be approximated as:

$$g(\hat{a}_n) \approx g(a) + J_g(a)(\hat{a}_n - a), \quad (36)$$

where $J_g(a)$ is the Jacobian matrix of $g(a)$, evaluated at a , i.e., $J_g(a) = \nabla g(a)$. Substituting the asymptotic normality of $\hat{a}_n$, the scaled deviation of $g(\hat{a}_n)$ satisfies:

$$\sqrt{n}(g(\hat{a}_n) - g(a)) \xrightarrow{d} N(0, J_g(a) \Sigma_a J_g(a)^T). \quad (37)$$

Thus, the asymptotic covariance matrix of $g(\hat{a}_n)$ is given by:

$$\Sigma_g = J_g(a) \Sigma_a J_g(a)^T. \quad (38)$$

The multivariate Delta Method is widely used in statistical applications, such as deriving the asymptotic variance of transformations of parameter vectors or constructing confidence regions for multivariate functions of estimators. Its flexibility and computational efficiency make it an indispensable tool in large sample inference.

Monte Carlo Simulation

Monte Carlo Simulation is a powerful numerical technique widely used in parameter estimation, particularly when analytical solutions are difficult or impossible to derive. This method relies on repeated random sampling and statistical inference to estimate model parameters. It is especially effective in scenarios where closed-form solutions are unavailable, likelihood functions are computationally challenging to evaluate, or simulation-based methods such as bootstrap or Bayesian inference are required.

The effectiveness of Monte Carlo Simulation is grounded in the LLN (Feller, 1971; Hogg *et al.*, 2019), a fundamental theorem in probability theory. The LLN states that as the number of independent samples N increases, the sample mean converges to the true mean of the population. Formally, if $X_1, X_2, \dots, X_N$ are i.i.d. random variables with mean μ , then the sample mean $\bar{X}_N = \frac{1}{N} \sum_{i=1}^N X_i$ satisfies:

$$\lim_{N \rightarrow \infty} \bar{X}_N = \mu \quad \text{almost surely.} \quad (39)$$

This property ensures that Monte Carlo estimators, which are essentially sample averages, become increasingly accurate as the number of simulations N grows. Consequently, the LLN provides the theoretical justification for the convergence of Monte Carlo estimates to the true parameter values.

The process of Monte Carlo Simulation for parameter estimation begins with defining the statistical model and the parameters to be estimated, denoted as a . Next, synthetic data is generated from the assumed model using known parameter values. This often involves specifying the probability distribution of the data. Subsequently, the chosen estimator, such as the or the MOM estimator, is applied to compute parameter estimates based on the simulated data. This process is repeated N times, generating new datasets and corresponding parameter estimates for each dataset.

Monte Carlo Simulation enables the evaluation of the performance of an estimator by calculating key metrics, including bias, variance, and mean squared error (MSE). The bias of an estimator $\hat{a}$ is defined as $Bias(\hat{a}) = E[\hat{a}] - a$, while the variance is $Var(\hat{a})$. The mean squared error is given by $MSE(\hat{a}) = Bias^2(\hat{a}) + Var(\hat{a})$. These metrics provide insights into the accuracy, consistency, and efficiency of the estimator.

Comprehensive details on Monte Carlo methods in parameter estimation can be found in standard references such as Robert and Casella (Robert and Casella, 2004; Gentle, 2009).

Delta Method for Projecting the Limiting Distribution of the SDP Under a Log-normal Distribution

To find the limiting distribution of $g(\mu, \sigma) = \mu + a\sigma$, where a is a constant, we use the Delta method. The Fisher Information Matrix for the log-normal distribution is:

$$I(\mu, \sigma) = \begin{pmatrix} \frac{1}{\sigma^2} & 0 \\ 0 & \frac{2}{\sigma^2} \end{pmatrix}.$$

The covariance matrix, which is the inverse of the Fisher Information Matrix, is given by:

$$I^{-1}(\mu, \sigma) = \begin{pmatrix} \sigma^2 & 0 \\ 0 & \frac{\sigma^2}{2} \end{pmatrix}.$$

Thus, the asymptotic distribution of $(\hat{\mu}, \hat{\sigma})$ is:

$$\begin{pmatrix} \hat{\mu} \\ \hat{\sigma} \end{pmatrix} \sim N \left(\begin{pmatrix} \mu \\ \sigma \end{pmatrix}, \begin{pmatrix} \sigma^2 & 0 \\ 0 & \frac{\sigma^2}{2} \end{pmatrix} \right).$$

Let $g(\mu, \sigma) = \mu + a\sigma$. The gradient of g with respect to (μ, σ) is:

$$\nabla g(\mu, \sigma) = \begin{pmatrix} \frac{\partial g}{\partial \mu} \\ \frac{\partial g}{\partial \sigma} \end{pmatrix} = \begin{pmatrix} 1 \\ a \end{pmatrix}.$$

Using the Delta Method, the variance of $g(\hat{\mu}, \hat{\sigma})$ is given by:

$$\text{Var}(g(\hat{\mu}, \hat{\sigma})) = \nabla g^T I^{-1} \nabla g.$$

Substituting the gradient $\nabla g = \begin{pmatrix} 1 \\ a \end{pmatrix}$ and the covariance matrix I^{-1} :

$$\text{Var}(g(\hat{\mu}, \hat{\sigma})) = \begin{pmatrix} 1 & a \end{pmatrix} \begin{pmatrix} \sigma^2 & 0 \\ 0 & \frac{\sigma^2}{2} \end{pmatrix} \begin{pmatrix} 1 \\ a \end{pmatrix}.$$

Perform the matrix multiplication:

$$\text{Var}(g(\hat{\mu}, \hat{\sigma})) = \begin{pmatrix} 1 & a \end{pmatrix} \begin{pmatrix} \sigma^2 \\ \frac{a\sigma^2}{2} \end{pmatrix}.$$

Simplify:

$$\text{Var}(g(\hat{\mu}, \hat{\sigma})) = \sigma^2 + a^2 \frac{\sigma^2}{2}.$$

Factorize:

$$\text{Var}(g(\hat{\mu}, \hat{\sigma})) = \sigma^2 \left(1 + \frac{a^2}{2} \right).$$

The limiting distribution of $g(\hat{\mu}, \hat{\sigma}) = \hat{\mu} + a\hat{\sigma}$ is:

$$g(\hat{\mu}, \hat{\sigma}) \square N \left(\mu + a\sigma, \sigma^2 \left(1 + \frac{a^2}{2} \right) \right).$$

The Proof of the Tighter Confidence Interval

To prove the inequality $\sqrt{1 + \frac{a^2}{2}} < \left(1 + \frac{a}{\sqrt{2}} \right)$ for $a > 0$, we use the ratio test. Let the ratio $R(a)$ be defined as:

$$R(a) = \frac{\sqrt{1 + \frac{a^2}{2}}}{1 + \frac{a}{\sqrt{2}}}.$$

Expand the denominator:

$$\left(1 + \frac{a}{\sqrt{2}} \right)^2 = 1 + \sqrt{2} \cdot a + \frac{a^2}{2}.$$

Thus:

$$R(a)^2 = \frac{1 + \frac{a^2}{2}}{1 + \sqrt{2} \cdot a + \frac{a^2}{2}}.$$

The numerator is $1 + \frac{a^2}{2}$, and the denominator is $1 + \sqrt{2} a + \frac{a^2}{2}$. Observe that for $a > 0$, the denominator grows faster than the numerator because of the $\sqrt{2} a$ term. This implies that:

$$R(a)^2 < 1.$$

Taking the square root (both sides remain positive for $a > 0$): $R(a) < 1$.

Since $R(a) < 1$, we conclude that:

$$\sqrt{1 + \frac{a^2}{2}} < \left(1 + \frac{a}{\sqrt{2}} \right), \quad \forall a > 0.$$

References

- Adkins, L.C. and Hill, R.C. (1990). An improved confidence ellipsoid for the linear regression model. *Journal of Statistical Computation and Simulation*, 36(1):9-18. <https://doi.org/10.1080/00949659008811249>
- Angkanavisal, N., Budsaba, K., and Volodin, A. (2012). Asymptotic confidence ellipses of parameters for the beta-poisson dose-response model. *Thailand Statistician*, 10(1):15-39.

- Aras, G. and Gardiner, J. (1995). Fixed diameter confidence ellipsoid in time-sequential models. *Sequential Analysis*, 14(1):29-39. <https://doi.org/10.1080/07474949508836318>
- Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637-654. <https://doi.org/10.1086/260062>
- Casella, G. and Berger, R.L. (2002). *Statistical Inference*. 2nd ed. Thomson Learning, UK, 660p.
- Chang, Y.-C.I. and Martinsek, A.T. (1992). Fixed Size Confidence Regions for Parameters of a Logistic Regression Model. *The Annals of Statistics*, 20(4):1953-1969. <https://doi.org/10.1214/aos/1176348897>
- Deprez, O. and Gerber, H.U. (1985). On convex principles of premium calculation. *Insurance: Mathematics and Economics*, 4(3):179-189. [https://doi.org/10.1016/0167-6687\(85\)90014-9](https://doi.org/10.1016/0167-6687(85)90014-9)
- Dickson, D.C.M., Hardy, M.R., and Waters, H.R. (2020). *Actuarial Mathematics for Life Contingent Risks*. 3rd ed. Cambridge University Press, 493p. <https://doi.org/10.1017/9781108784184>
- Duangchana, N. and Budsaba, K. (2014). Asymptotic confidence ellipses of parameters for the inverse gaussian distribution. *Thammasat International Journal of Science and Technology*, 19(2):22-29.
- Durand, D. (1954). Joint confidence regions for multiple regression coefficients. *Journal of the American Statistical Association*, 49(265):130-146. <https://doi.org/10.1080/01621459.1954.10501220>
- Edden, A.C. (1971). A measure of species diversity related to the log-normal distribution of individuals among species. *Journal of Experimental Marine Biology and Ecology*, 6(3):199-209. [https://doi.org/10.1016/0022-0981\(71\)90019-0](https://doi.org/10.1016/0022-0981(71)90019-0)
- Feller, W. (1971). *An Introduction to Probability Theory and Its Applications*. 2nd ed. John Wiley & Sons, New York, 669p.
- Fisher, R.A. (1922). On the mathematical foundations of theoretical statistics. *Philosophical Transactions of the Royal Society A Mathematical, Physical and Engineering Sciences*. 222(594-604):309-368. <https://doi.org/10.1098/rsta.1922.0009>
- Fisher, R.A. (1925). *Theory of statistical estimation*. In: *Mathematical Proceedings of the Cambridge Philosophical Society*, 22(5):700-725. <https://doi.org/10.1017/S0305004100009580>
- Friendly, M., Monette, G., and Fox, J. (2013). Elliptical insights: Understanding statistical methods through elliptical geometry. *Statistical Science*, 28(1):1-39. <https://doi.org/10.1214/12-STS402>
- Gao, J., Ji, J., Liu, Y., Zhang, F., Wang, R., and Wang, H. (2023). A confidence ellipse analysis for stochastic dynamics model of Alzheimer's disease. *Nonlinear Dynamics*, 111(14):13513-13533. <https://doi.org/10.1007/s11071-023-08547-y>
- Gentle, J.E. (2009). *Computational Statistics*. Springer, New York, 726p. <https://doi.org/10.1007/978-0-387-98144-4>
- Han, L., Wang, L., Ding, X., Wen, H., Yuan, X., and Zhang, W. (2022). Similarity quantification of soil parametric data and sites using confidence ellipses. *Geoscience Frontiers*, 13(1):101280. <https://doi.org/10.1016/j.gsf.2021.101280>
- Hansoongnern, K., Sakonthawichot, B., Sittisareesmer, A., and Phaphan, W. (2018). Estimation of asymptotic confidence ellipses for birnbaum-saunders distribution. *Burapha Science Journal*, 23(2):1029-1043.
- Herrera, C.M. and Jovani, R. (2010). Lognormal distribution of individual lifetime fecundity: insights from a 23-year study. *Ecology*, 91(2):422-430. <https://doi.org/10.1890/09-0849.1>
- Hogg, R.V., McKean, J.W. and Craig, A.T. (2019). *Introduction to Mathematical Statistics*. 8th ed. London, UK, 760p.
- Hotelling, H. (1931). The generalization of student's ratio. *Annals of Mathematical Statistics*, 2(3):360-378. <https://doi.org/10.1214/aoms/1177732979>
- Jobe, J.M. (1989). Log-normal distributions: Theory and applications. *Technometrics*, 31(3):392. <https://doi.org/10.1080/00401706.1989.10488576>
- Johnson, N.L., Kotz, S., and Balakrishnan, N. (1994). *Continuous Univariate Distributions*. 2nd ed. John Wiley & Sons, 784p.
- Kleiber, C. and Kotz, S. (2003). *Statistical size distributions in economics and actuarial sciences*. John Wiley & Sons, 352p. <https://doi.org/10.1002/0471457175>
- Kolmogorov, A.N. (1950). *Foundations of the Theory of Probability*. Chelsea Publishing Company, New York, 71p.
- Kreps, R.E. (1990). Reinsurer risk loads from marginal surplus requirements. *Proceedings of the Casualty Actuarial Society*, 77:196-203.
- Larsen, R.J. and Marx, M.L. (2017). *An Introduction to Mathematical Statistics and Its Applications*. 5th ed. Pearson, London, UK, 757p.
- Lehmann, E.L. and Casella, G. (1998). *Theory of Point Estimation*. 2nd ed. Springer, NY, 590p.
- Liu, S., Chen, J., Li, J., Li, T., Shi, H., and Sivakumar, B. (2023). Building a large dam: identifying the relationship between catchment area and slope using the confidence ellipse approach. *Geoscience Letters*, 10(1):4. <https://doi.org/10.1186/s40562-022-00260-9>
- Merton, R.C. (1973). Theory of rational option pricing. *Bell Journal of Economics and Management Science*, 4(1):141-183. <https://doi.org/10.2307/3003143>
- Moumen, A.E. and Bouzidi, M. (2024). Confidence ellipses for simple regression parameters with quasi-associated random errors. *Studies in Engineering and Exact Sciences*, 5(2):e9951. <https://doi.org/10.54021/seesv5n2-417>
- Newcombe, R.G. (2012). *Confidence Intervals for Proportions and Related Measures of Effect Size*. 1st ed. CRC Press, UK, 293p. <https://doi.org/10.1201/b12670>
- Ott, W.R. (1990). A physical explanation of the log-normality of pollutant concentrations. *Journal of the Air & Waste Management Association*, 40(10):1378-1383. <https://doi.org/10.1080/10473289.1990.10466789>
- Pauly, M.V. and Herring, B. (2002). Community rating and the market for private non-group health insurance. *Journal of Health Politics, Policy and Law*, 27(1):6-21.
- Phaphan, W. (2020). Asymptotic confidence ellipses for the re-parameterized inverse gaussian distribution. *The Journal of Applied Science*, 19(1):36-50. <https://doi.org/10.14416/j.appsci.2020.01.004>
- Phaphan, W. and Pongsart, T. (2019). Asymptotic confidence ellipses for length-biased inverse gaussian distribution. *The Journal of King Mongkut's University of Technology North Bangkok*, 29(2):332-341. <https://doi.org/10.14416/j.kmutnb.2018.12.08>
- R Core Team. (2022). *A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Rehman, Z. and Klugman, S. (2010). Quantifying uncertainty in reserve estimates. *Casualty Actuarial Society*, 4(1):30-46.
- Robert, C.P. and Casella, G. (2004). *Monte Carlo Statistical Methods*. 2nd ed. Springer, New York, 649p. <https://doi.org/10.1007/978-1-4757-4145-2>
- Sen, P.K., Singer, J.M., and de Lima, A.C.P. (2010). *From finite sample to asymptotic methods in statistics*. Cambridge University Press, 398p. <https://doi.org/10.1017/CBO9780511806957>
- Sriram, T.N. (2001). Fixed size confidence regions for parameters of threshold AR(1) models. *Journal of statistical planning and inference*, 97(2):293-304. [https://doi.org/10.1016/S0378-3758\(00\)00246-9](https://doi.org/10.1016/S0378-3758(00)00246-9)
- Thonglim, P., Budsaba, K., and Volodin, A.I. (2014). Asymptotic confidence ellipses of parameters for the birnbaum-saunders distribution. *Thailand Statistician*, 12(2):207-222.

- van der Vaart, A.W. (2000). Asymptotic statistics. 1st ed. Cambridge University Press, NY, 443p.
- Wackerly, D., Mendenhall, W. and Scheaffer, R.L. (2014). Mathematical Statistics with Applications. 7th ed. Cengage Learning, Boston, MA, 912p.
- Wahba, G. (1969). Estimation of the coefficients in a multidimensional distributed lag model. *Econometrica*, 37(3):398-407. <https://doi.org/10.2307/1912789>
- Wang, X. and Weyer, E. (2024). A sample based algorithm for constructing guaranteed confidence ellipsoids for linear regression models with deterministic regressor. *IFAC-PapersOnLine*, 58(15):193-198. <https://doi.org/10.1016/j.ifacol.2024.08.527>
- Young, V.R. (2004). Premium principles. *Encyclopedia of Actuarial Science*, 1(1):1343-1349. <https://doi.org/10.1002/9780470012505.tap027>
- Zhang, Y., Zhang, X., Wang, P., Wu, Y. and Chow, S.C. (2024). A proposed confidence ellipse approach for benefit-risk assessment in clinical trials. *Research Square*, p. 1-22. <https://doi.org/10.21203/rs.3.rs-5285025/v1>