

An alternative method for estimating lognormal means

Yeil Kwon^{1,a}

^aDepartment of Mathematics, University of Central Arkansas, USA

Abstract

For a probabilistic model with positively skewed data, a lognormal distribution is one of the key distributions that play a critical role. Several lognormal models can be found in various areas, such as medical science, engineering, and finance. In this paper, we propose a new estimator for a lognormal mean and depict the performance of the proposed estimator in terms of the relative mean squared error (RMSE) compared with Shen's estimator (Shen *et al.*, 2006), which is considered the best estimator among the existing methods. The proposed estimator includes a tuning parameter. By finding the optimal value of the tuning parameter, we can improve the average performance of the proposed estimator over the typical range of σ^2 . The bias reduction of the proposed estimator tends to exceed the increased variance, and it results in a smaller RMSE than Shen's estimator. A numerical study reveals that the proposed estimator has performance comparable with Shen's estimator when σ^2 is small and exhibits a meaningful decrease in the RMSE under moderate and large σ^2 values.

Keywords: lognormal distribution, relative mean squared error, variance approximation, tuning parameter, consistent estimator

1. Introduction

Lognormal distribution has a central role to play in probabilistic modeling for the phenomena described by a skewed distribution of positive random variables in nature. One can find numerous examples of lognormal models, such as clinical pharmacokinetic studies (Lacey *et al.*, 1997), failure rates in engineering (Ohring and Kasprzak, 1998), and stock prices in finance (Hull, 2018). Therefore, finding a precise and stable estimation method for the lognormal mean is crucial in lognormal modeling problems. In particular, an accurate estimation becomes much more critical when σ^2 is large because the estimation error is inclined to escalate rapidly when σ^2 becomes larger.

This paper focuses on the estimation method for the mean of the lognormal distribution. For decades, a few estimators for the lognormal mean have been proposed. The sample mean is one of the most popular estimators because of its simplicity and lack of bias. Although the maximum likelihood estimator (MLE) is not unbiased, it has desirable asymptotic properties, such as asymptotic normality and asymptotic efficiency. Finney (1941) proposed the uniformly minimum variance unbiased estimator (UMVUE) with a form of infinite series of the sample variance, and it has the minimum variance among all unbiased estimators. Evans and Shaban (1974), and Zhou (1998) proposed estimators based on Zellner's conditional minimal mean squared error (MSE) estimator (Zellner, 1971). These are not unbiased but have markedly smaller MSEs than Finney's UMVUE with small sample sizes. Shen *et al.* (2006) proposed an efficient estimator using second-order asymptotics, which improves the conditional minimal MSE estimators for large σ^2 values. Longford (2009) proposed an estimator with a

¹ Department of Mathematics, University of Central Arkansas, 201 Donaghey Avenue, Conway, AR, 72035, USA.
E-mail: ykwon1@uca.edu

form of $\exp(\bar{X} + b\hat{\sigma}^2)$ and directly found b , achieving the lower bound of the MSE. Because b is a function of unknown σ^2 , \hat{b} is determined by replacing σ^2 with the unbiased estimator for σ^2 .

In this research, we propose a new estimator of a specific form containing a tuning parameter k in it. This parameter k is determined by a value that minimizes the average relative mean squared error (RMSE) of the proposed estimator over the given range of σ^2 . The proposed estimator's performance is compared with that of the existing estimators in terms of the RMSE. The numerical study reveals that the proposed estimator has comparable performance with the existing estimators when σ^2 is small. However, it improves adequately the existing estimators when σ^2 is moderate or large, regardless of the sample size.

Section 2 includes the literature review for the existing estimators using a lognormal mean. In Section 3, we propose a new estimator and address the properties of the new estimator. We report the numerical study results to describe the performance of the proposed estimator based on the RMSE in Section 4.

2. Existing estimators

Let X be a normal random variable with mean μ and variance σ^2 : $X \sim N(\mu, \sigma^2)$. Then, $Y = e^X$ is lognormally distributed with parameters μ and σ^2 : $Y \sim LN(\mu, \sigma^2)$. The mean of Y , θ , is a function of both μ and σ^2 ,

$$\theta = E(Y) = e^{\mu + \frac{\sigma^2}{2}}. \quad (2.1)$$

Suppose X_1, \dots, X_n is a random sample from $N(\mu, \sigma^2)$. Then, $Y_i = e^{X_i} \stackrel{\text{iid}}{\sim} LN(\mu, \sigma^2)$ for $i = 1, \dots, n$. We define three statistics,

$$\bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i, \quad \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i, \quad \text{and} \quad S^2 = \sum_{i=1}^n (X_i - \bar{X})^2. \quad (2.2)$$

Let $R(\hat{\theta}, \theta)$ denote the relative mean squared error (RMSE), which is the risk of an estimator $\hat{\theta}$ for θ under the squared loss function $L(\hat{\theta}, \theta) = (\hat{\theta}/\theta - 1)^2$,

$$R(\hat{\theta}, \theta) = E\left(\frac{\hat{\theta}}{\theta} - 1\right)^2. \quad (2.3)$$

In this paper, the performance of each estimator is compared with the others in terms of the RMSE.

2.1. Sample mean and maximum likelihood estimator

In the research based on the lognormal distribution, the sample mean, $\hat{\theta}_{\text{sm}} = \bar{Y}$, is one of the most widely used estimators. The sample mean is popular because it is obtainable and unbiased. However, in a vast amount of the literature, the sample mean is significantly inefficient compared with other biased estimators with both small and large samples (Zhou, 1998; Shen *et al.*, 2006; Longford, 2009).

However, the MLEs for μ and σ^2 of the normal distribution $N(\mu, \sigma^2)$ are given by \bar{X} and S^2/n using the statistics defined in (2.2). By the invariant property of MLE, one can attain the MLE for θ : $\hat{\theta}_{\text{mle}} = \exp(\bar{X} + S^2/2n)$. The MLE has the preferable asymptotic properties of MLE, such as consistency, normality, and efficiency. In particular, when the sample size is large, the MLE has a comparable RMSE with the other existing estimators and is very easy to calculate. Zhou recommended using the MLE for large samples, such as $n \geq 200$ (Zhou, 1998). However, $\hat{\theta}_{\text{mle}}$ performs worse than the sample mean when the sample size is small (see Figure 1).

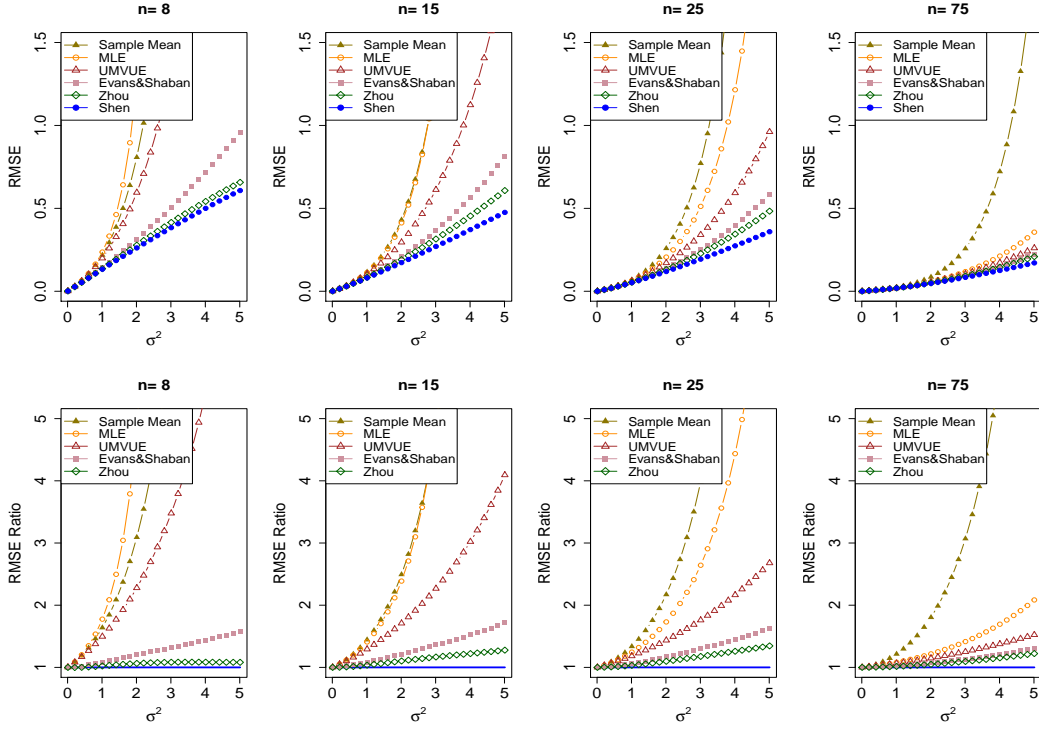


Figure 1: Risk comparison of the existing estimators with various sample sizes. [Upper] Relative mean squared error (RMSE) of six estimators. [Lower] Ratio of RMSE, $R(\hat{\theta}_*, \theta) / R(\hat{\theta}_{Shen}, \theta)$, where $\hat{\theta}_*$ denotes each of $\hat{\theta}_{sm}$, $\hat{\theta}_{mle}$, $\hat{\theta}_{umvue}$, $\hat{\theta}_{ES}$, and $\hat{\theta}_{Zhou}$. Overall, $\hat{\theta}_{Shen}$ has the smallest RMSE compared with the other five estimators.

2.2. Estimators based on the infinite series

There are several estimators for θ based on the infinite series of S^2 . Finney (1941) derived UMVUE with the joint complete and sufficient statistics (\bar{X}, S^2) for (μ, σ^2) as follows,

$$\hat{\theta}_{umvue} = \exp(\bar{X}) g\left(\frac{S^2}{2}\right),$$

where

$$g(t) = \sum_{j=0}^{\infty} \frac{\Gamma\left(\frac{n-1}{2}\right)}{j! \Gamma\left(\frac{n-1}{2} + j\right)} \left(\frac{n-1}{2n} t\right)^j.$$

Although $\hat{\theta}_{umvue}$ has the smallest risk under the squared error loss among all unbiased estimators, an estimator with a lower MSE than the UMVUE of the lognormal mean exists (Rukhin, 1986). Let $\hat{\theta}_r$ denote an estimator of the form $\exp(\bar{X})r(\sigma^2)$. Zellner (1971) found that conditioning on σ^2 , the MSE of $\hat{\theta}_r$ is minimized when $r(\sigma^2) = \exp((n-3)\sigma^2/2n)$, and it results in a conditionally minimal MSE estimator of $\exp(\bar{X} + (n-3)\sigma^2/2n)$. Evans and Shaban (1974) and Zhou (1998) proposed the following estimators,

$$\hat{\theta}_{ES} = \exp(\bar{X}) g\left(\frac{n-3}{2(n-1)} S^2\right) \quad \text{and} \quad \hat{\theta}_{Zhou} = \exp(\bar{X}) g\left(\frac{n-4}{2(n-1)} S^2\right),$$

using unbiased estimators for $\exp((n - 3)\sigma^2/2n)$ and $\exp((n - 4)\sigma^2/2n)$, respectively. Figure 1 illustrates that $\hat{\theta}_{ES}$ and $\hat{\theta}_{Zhou}$ significantly improve $\hat{\theta}_{umvue}$ for small samples. With small sample size, $\hat{\theta}_{Zhou}$ has less risk than $\hat{\theta}_{ES}$, but the two estimators are almost equivalent to each other when the sample size is large.

2.3. Shen's efficient estimator

Shen *et al.* (2006) proposed a class of estimators for θ of the form $\hat{\theta}_c = \exp(\bar{X} + cS^2/2)$ with $c = 1/(n + d)$ and $d > -n$. Let $M_U(t)$ denote the moment generate function of a random variable U . From the definition in (2.2), we have $M_{X_i}(t) = E(e^{tX_i}) = e^{\mu + \sigma^2 t^2/2}$, and $S^2/\sigma^2 \sim \chi^2_{(n-1)}$ with $M_{\chi^2_{(n-1)}}(t) = E(e^{t\chi^2_{(n-1)}}) = (1 - 2t)^{-(n-1)/2}$. Thus,

$$E(e^{\bar{X}}) = e^{\mu + \frac{\sigma^2}{2n}}, \quad E(e^{2\bar{X}}) = e^{2\mu + \frac{2\sigma^2}{n}}, \quad \text{and} \quad E(e^{cS^2}) = (1 - 2c\sigma^2)^{-\frac{n-1}{2}}. \tag{2.4}$$

Using the results in (2.4), the squared error risk of $\hat{\theta}_c$, $R(\hat{\theta}_c, \theta) = E(\hat{\theta}_c/\theta - 1)^2$, can be shown as follows,

$$\begin{aligned} R(\hat{\theta}_c, \theta) &= \frac{1}{\theta^2} E \left[\exp\left(\bar{X} + \frac{cS^2}{2}\right) - \theta \right]^2 \\ &= \exp\left\{\left(\frac{2}{n} - 1\right)\sigma^2\right\} (1 - 2c\sigma^2)^{-\frac{n-1}{2}} - 2 \exp\left\{\left(\frac{1}{n} - 1\right)\frac{\sigma^2}{2}\right\} (1 - c\sigma^2)^{-\frac{n-1}{2}} + 1, \quad c < \frac{1}{2\sigma^2}. \end{aligned}$$

With the standard expansion for $c = 1/(n + d) = 1/n - d/n^2 + o(1/n^2)$, $R(\hat{\theta}_c, \theta)$ can be written as a function of d ,

$$R(\hat{\theta}_c, \theta) = \frac{\sigma^2}{n} \left\{ 1 + \frac{\sigma^2}{2} + \frac{\sigma^2}{4n} \left[d^2 - (8 + 3\sigma^2)d + 8\sigma^2 + \frac{7}{4}\sigma^4 \right] \right\} + o\left(\frac{1}{n^2}\right),$$

which is minimized at $d = 4 + 3\sigma^2/2$. Replacing the unknown σ^2 with its unbiased estimator $S^2/(n - 1)$, we obtain the following,

$$\hat{\theta}_{Shen} = \exp\left(\bar{X} + \frac{(n - 1)S^2}{2(n + 4)(n - 1) + 3S^2}\right). \tag{2.5}$$

With small sample sizes and small values of σ^2 , the RMSE of θ_{Zhou} is slightly lower than that of θ_{Shen} Shen *et al.* (2006). Except for such a case, $\hat{\theta}_{Shen}$ has a uniformly smaller RMSE than $\hat{\theta}_{sm}$, $\hat{\theta}_{mle}$, $\hat{\theta}_{umvue}$, $\hat{\theta}_{ES}$, and $\hat{\theta}_{Zhou}$ (see Figure 1).

2.4. Longford's estimator

Longford (2009) proposed a family of estimators of the form $\hat{\theta}_b = \exp(\bar{X} + bS^2/(n - 1))$ to estimate $\theta_a = \exp(\mu + a\sigma^2)$ with a given constant a . The mean of the lognormal distribution, θ , in (2.1) is a special case of θ_a where $a = 1/2$. (The mode of the lognormal distribution is $\exp(\mu - \sigma)$, which is a special case of θ_a where $a = -1$.) The relative mean squared error of $\hat{\theta}_b$, is given by

$$R(\hat{\theta}_b, \theta) = \exp\left\{\left(\frac{2}{n} - 1\right)\sigma^2\right\} \left(1 - \frac{4b\sigma^2}{n - 1}\right)^{-\frac{n-1}{2}} - 2 \exp\left\{\left(\frac{1}{n} - 1\right)\frac{\sigma^2}{2}\right\} \left(1 - \frac{2b\sigma^2}{n - 1}\right)^{-\frac{n-1}{2}} + 1, \quad b < \frac{n - 1}{4\sigma^2}.$$

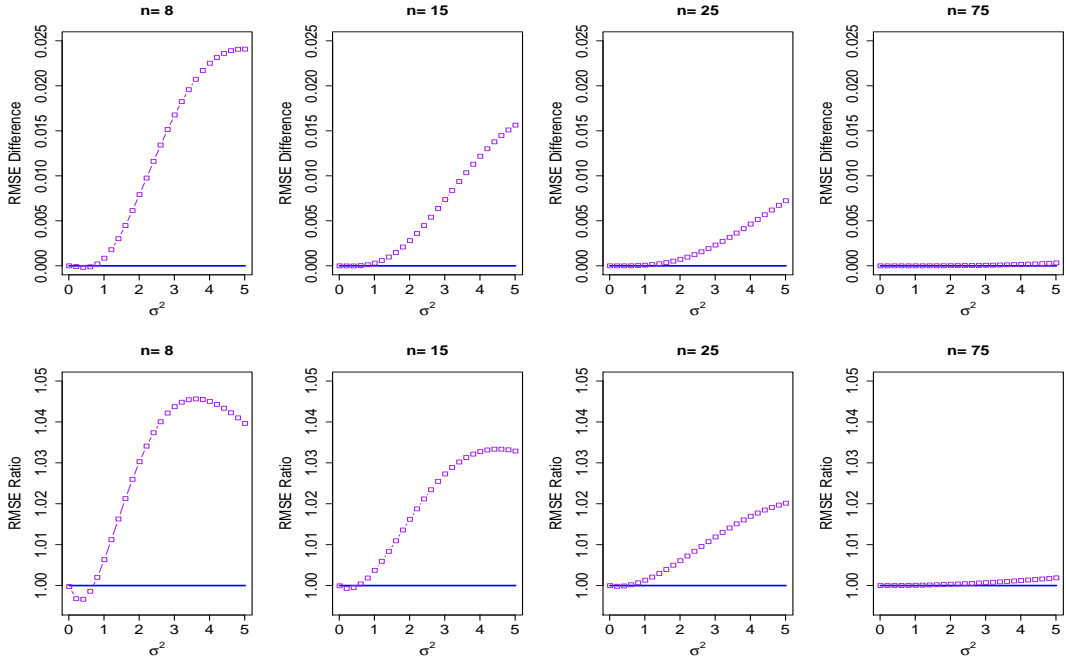


Figure 2: Risk comparison between the estimators of Longford and Shen. [Upper] Difference in the Relative mean squared error (RMSE) $R(\hat{\theta}_{Lf}, \theta) - R(\hat{\theta}_{Shen}, \theta)$. [Lower] Ratio of the RMSE, $R(\hat{\theta}_{Lf}, \theta)/R(\hat{\theta}_{Shen}, \theta)$.

The minimum of $R(\hat{\theta}_b, \theta)$ can be determined by solving the following equation,

$$\frac{\partial R(\hat{\theta}_b)}{\partial b} = e^{\sigma^2(\frac{2}{n}-1)} \left(\frac{n-1}{2}\right) \left(\frac{4\sigma^2}{n-1}\right) \left(1 - \frac{4b\sigma^2}{n-1}\right)^{-\frac{n+1}{2}} - 2e^{\frac{\sigma^2}{2}(\frac{1}{n}-1)} \left(\frac{n-1}{2}\right) \left(\frac{2\sigma^2}{n-1}\right) \left(1 - \frac{2b\sigma^2}{n-1}\right)^{-\frac{n+1}{2}} = 0,$$

and the solution is obtained as follows,

$$b = \frac{n-1}{2\sigma^2} \cdot \frac{D-1}{2D-1} \quad \text{with} \quad D = \exp\left(\frac{n-3}{n(n+1)}\sigma^2\right).$$

By substituting σ^2 with $S^2/(n-1)$, we have the Longford estimator as

$$\hat{\theta}_{Lf} = \exp\left\{\bar{X} + \frac{n-1}{2} \cdot \frac{\exp\left(\frac{n-3}{n(n^2-1)}S^2\right) - 1}{2 \exp\left(\frac{n-3}{n(n^2-1)}S^2\right) - 1}\right\}. \tag{2.6}$$

Using the Taylor expansion, $\hat{\theta}_{Lf}$ can be approximated as follows,

$$\hat{\theta}_{Lf} \approx \exp\left\{\bar{X} + \frac{(n-1)S^2}{2n(n+1)\left(1 + \frac{2}{n-3}\right) + 4S^2}\right\}.$$

Figure 2 presents the risk comparison of $\hat{\theta}_{Shen}$ and $\hat{\theta}_{Lf}$. First, two estimators are almost identical over $\sigma^2 \in (0, 5)$ with a large sample. Next, with small samples and small values of σ^2 , $\hat{\theta}_{Lf}$ performs a little better than $\hat{\theta}_{Shen}$. However, for all other settings of n and σ^2 , $\hat{\theta}_{Shen}$ outperforms $\hat{\theta}_{Lf}$. Thus, from Figures 1 and 2, we can consider $\hat{\theta}_{Shen}$ to be the best estimator of the existing methods in terms of the RMSE. Therefore, in this paper, we use $\hat{\theta}_{Shen}$ as a benchmark for evaluating the performance of the proposed estimator.

3. Proposed estimator

3.1. Analytic properties

We consider an estimator of the form, for $k \geq 1$,

$$\hat{\theta}_k = \exp\left(\bar{X} + \frac{(n-1)S^2}{2n(n+1) + kS^2}\right), \tag{3.1}$$

with the squared error risk,

$$R(\hat{\theta}_k, \theta) = e^{\sigma^2(\frac{2}{n}-1)}Q_2(k) - 2e^{\frac{\sigma^2}{2}(\frac{1}{n}-1)}Q_1(k) + 1, \tag{3.2}$$

where,

$$Q_i(k) = E\left[\exp\left(\frac{i(n-1)S^2}{2n(n+1) + kS^2}\right)\right], \quad i = 1, 2. \tag{3.3}$$

Proposition 1. $R(\hat{\theta}_k, \theta) = E\left(\frac{\hat{\theta}_k}{\theta} - 1\right)^2 \rightarrow 0$, for $0 < k < \infty$.

Proof: Let $\hat{\theta}_c = \exp(\bar{X} + cS^2/2)$. Shen *et al.* (2006) showed $R(\hat{\theta}_c, \theta) \rightarrow 0$ when $cn \rightarrow 1$. $\hat{\theta}_k$ can be written as the same form as $\hat{\theta}_c$ with $c = (n-1)/(n(n+1) + kS^2/2)$. Since $cn = n(n-1)/(n(n+1) + kS^2/2) \rightarrow 1$, we have $R(\hat{\theta}_k, \theta) \rightarrow 0$. □

Proposition 1 states the relative risk of $\hat{\theta}_k$ converges to zero as the sample size increases implying $\hat{\theta}_k$ is a consistent estimator for θ . With proposition 1, one can conclude that the bias and the variance of $\hat{\theta}_k$ converges to 0 as $n \rightarrow \infty$. Proposition 2 shows both $\hat{\theta}_k$ and $\hat{\theta}_{Shen}$ underestimate θ and, for $k \leq 3$, the squared bias of $\hat{\theta}_k$ is uniformly smaller than that of $\hat{\theta}_{Shen}$.

Proposition 2.

- (1) Both $\hat{\theta}_k$ and $\hat{\theta}_{Shen}$ are negatively biased: $E(\hat{\theta}_k) - \theta < 0$ and $E(\hat{\theta}_{Shen}) - \theta < 0$.
- (2) $Bias^2(\hat{\theta}_k) < Bias^2(\hat{\theta}_{Shen})$, for $k \leq 3$.
- (3) $Bias^2(\hat{\theta}_k)$ converges to zero as $n \rightarrow \infty$.

Proposition 3. The variance of $\hat{\theta}_k$ is given by

$$V(\hat{\theta}_k) = \exp\left(2\mu + \frac{\sigma^2}{n}\right)\left[\exp\left(\frac{\sigma^2}{n}\right)Q_2(k) - Q_1^2(k)\right],$$

where, for $i = 1, 2$,

$$Q_i(k) = E \left[\exp \left(\frac{i(n-1)S^2}{2n(n+1) + kS^2} \right) \right] = \int_0^\infty \exp \left(\frac{i\sigma^2(n-1)w}{2n(n+1) + k\sigma^2w} \right) g(w) dw,$$

and $g(w)$ is a p.d.f of chi-squared distribution with degree of freedom $n-1$. Moreover, $V(\hat{\theta}_k)$ converges to zero as $n \rightarrow \infty$.

Proposition 4. For a large sample, the variance of $\hat{\theta}_k$ can be approximated by,

$$V(\hat{\theta}_k) \approx \exp \left(2\mu + \frac{2(n-1)^2\sigma^2}{2n(n+1) + k(n-1)\sigma^2} \right) \left[\frac{\sigma^2}{n} + \frac{8\sigma^4n^2(n+1)^2(n-1)^3}{(2n(n+1) + k\sigma^2(n-1))^4} \right], \quad (3.4)$$

and the right hand side of (3.4) converges to zero as $n \rightarrow \infty$.

For a large sample, the approximated $V(\hat{\theta})$ can be estimated by substituting \bar{X} and $S^2/(n-1)$ for μ and σ^2 , respectively as follows,

$$\widehat{V}(\hat{\theta}_k) = \exp \left(2\bar{X} + \frac{2(n-1)S^2}{2n(n+1) + kS^2} \right) \left[\frac{S^2}{n(n-1)} + \frac{8n^2(n+1)^2(n-1)S^4}{(2n(n+1) + kS^2)^4} \right].$$

Theorem 1. A positive value $k = k^*$ exists, at which $R(\hat{\theta}_{k^*}, \theta)$ is the unique minimum of $R(\hat{\theta}_k, \theta)$ in (3.2), for $n \geq 3$.

3.2. Choosing a tuning parameter k

Theorem 1 states that a unique value $k = k^*$ exists that minimizes the RMSE of the $\hat{\theta}_k$ in (3.1). Evidently, the k^* value depends on σ^2 and the sample size n ; thus, the optimal k^* can be obtained numerically, given σ^2 . However, because σ^2 is unknown, we suggest the following method of determining k ,

- (1) Let v_0, v_1, \dots, v_I be values such that $\sigma_{LB}^2 = v_0 < v_1 < v_2 < \dots < v_i < \dots < v_I = \sigma_{UB}^2$, where σ_{LB}^2 and σ_{UB}^2 are the upper and lower bounds of σ^2 in which we are interested.
- (2) Let w_1, \dots, w_J be positive integer values such that $n_{LB} = w_1 < w_2 < \dots < w_j < \dots < w_J = n_{UB}$, where n_{LB} and n_{UB} are the upper and lower bounds of n in which we are interested.
- (3) Compute $R_{ij}(\hat{\theta}_k, \theta)$ for each discretized value of k over $(1, K)$, where $R_{ij}(\hat{\theta}_k, \theta)$ is the RMSE of $\hat{\theta}_k$ in (3.2) with $\sigma^2 = v_i$ and $n = w_j$, for $i = 1, \dots, I$ and $j = 1, \dots, J$. Here, K is a number large enough for finding the optimal value of k . Note that $R_{ij}(\hat{\theta}_k, \theta)$ includes the expected values $Q_i(k)$, $i = 1, 2$, in (3.3) and they will be calculated by using the numerical integration with (G.1) and (G.2) in Appendix G.
- (4) Determine $k = \tilde{k}$ for the proposed estimator $\hat{\theta}_{Prop}$. \tilde{k} is the value of k at which the average of R_{ij} is minimized,

$$\tilde{k} = \arg \min_{k>0} \frac{1}{IJ} \sum_{i=1}^I \sum_{j=1}^J R_{ij}(\hat{\theta}_k, \theta). \quad (3.5)$$

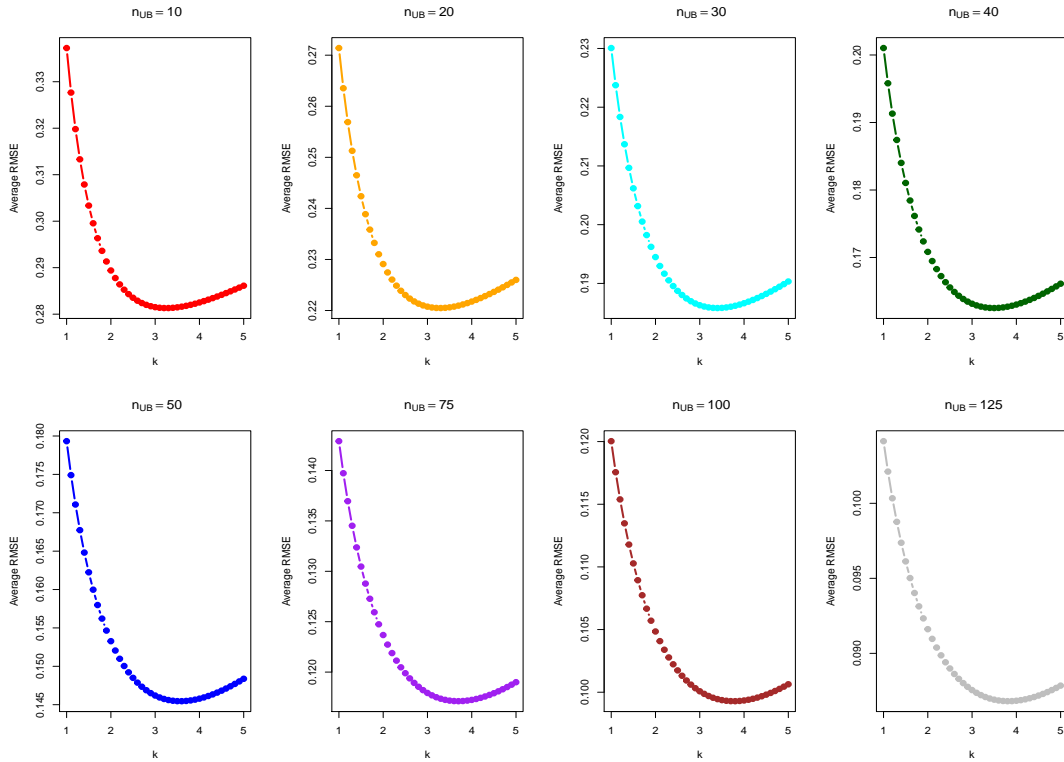


Figure 3: Average relative mean squared error (RMSE) where $\sigma_{LB}^2 = 0$, $\sigma_{UB}^2 = 5$, and $n_{LB} = 5$. Each graph presents the average RMSE with a different setting of n_{UB} . The optimal k values are 3.3, 3.3, 3.4, 3.5, 3.6, 3.7, 3.8, and 3.8 from the left-top to the right-bottom.

Therefore, \tilde{k} can be considered as the value of k that minimizes the average of RMSEs over $\sigma_{LB}^2 < \sigma^2 < \sigma_{UB}^2$ and $n_{LB} < n < n_{UB}$.

It is worth to point out that we set $\sigma_{UB}^2 = 5$. A number of researches such as Land (1972), Chelmos *et al.* (1995), Zhou (1997, 1998), and Shen *et al.* (2006) provide empirical evidence of $\sigma^2 < 5$ for the majority of the lognormal data in practice. For this reason, many literature concerning the lognormal data presented the numerical study over $\sigma^2 \in (0, 5)$: Shen *et al.* (2006), Zhou (1997, 1998). The value of the skewness coefficient with $\sigma^2 = 5$ corresponds to 1826.2, which is not commonly observed in real data (Zhou, 1998). Therefore, the proposed estimator based on \tilde{k} obtained by using $\sigma_{UB}^2 = 5$ can be applied for the majority of the lognormal data except for the extreme cases of σ^2 . Moreover, since all of the σ_{LB}^2 , σ_{UB}^2 , n_{LB} , and n_{UB} are predetermined as follows, \tilde{k} is a generic value that does not depend on data,

- (a) Set $\sigma_{LB}^2 = 0$, $\sigma_{UB}^2 = 5$. The size of increment of σ^2 is chosen as 0.1. Thus, $v_0 = 0$ and $v_i = v_{i-1} + 0.1$ for $i = 1, \dots, I$, with $I = 50$. The smaller increment size of σ^2 enable us to have a fine grid in this optimizing process. However, we found that, from the numerical study, it could not make significant difference for determining \tilde{k} .
- (b) Set $n_{LB} = 5$, $w_0 = 0$ and $w_j = w_{j-1} + 5$, for $j = 1, \dots, J$ with $J = n_{UB}/5$. To investigate the

effect of n_{UB} for $R_{ij}(\hat{\theta}_k, \theta)$, we examined the cases of $n_{UB} = 10, 20, 30, 40, 50, 75, 100$, and 125 . When n_{UB} is larger than 100 , the effect of n_{UB} for \tilde{k} diminishes remarkably; thus, $n_{UB} = 100$ ($J = 20$), is selected for finding \tilde{k} . Moreover, with the same reason as (a), the increment of sample size is chosen as 5 . The increment smaller than 5 does not show notable difference in $R_{ij}(\hat{\theta}_k, \theta)$ calculation.

- (c) As for K , we set $k = 1.0, 1.1, 1.2, \dots, K$, with $K = 20$. Figure 3 displays the graphs of k versus the average RMSE for each n_{UB} considered in (b).
- (d) As a result, the k value for the proposed estimator is determined by $\tilde{k} = 3.5$; thus, the proposed estimator is given as follows,

$$\hat{\theta}_{prop} = \exp\left(\bar{X} + \frac{(n-1)S^2}{2n(n+1) + 3.5S^2}\right). \tag{3.6}$$

On the other hand, one can use the confidence interval of σ^2 to check whether $\sigma^2 < 5$ or not. Consider $100(1 - \alpha)\%$ one-sided confidence interval for σ^2 . Since $S^2/\sigma^2 \sim \chi^2_{(n-1)}$, the upper bound of the interval is given by $S^2/\chi^2_{1-\alpha}$, where $\chi^2_{1-\alpha}$ is the 100α -th percentile of the chi-squared distribution with degree of freedom $n - 1$. If the upper bound of the confidence interval $S^2/\chi^2_{1-\alpha}$ is less than 5 , we can be highly confident that $\sigma^2 < 5$. Note that $S^2/\chi^2_{1-\alpha} < 5$ is equivalent to (sample variance) $< 5 \cdot \chi^2_{1-\alpha}/(n - 1)$. With 95% confidence interval, $5 \cdot \chi^2_{1-\alpha}/(n - 1)$ values are $1.97, 2.71, 3.08, 3.31$, and 3.48 under the sample sizes $n = 11, 21, 31, 41$, and 51 , respectively. Thus, for example, if [$11 \leq n \leq 30$ & the sample variance < 2], or [$n > 30$ & the sample variance < 3], it is recommendable to use the proposed estimator. All of the lognormal data used in the researches mentioned in section 3.2 have the sample variances less than 1.5 .

4. Numerical study for risk comparison

In this section, we compare the performance of the proposed estimator with the existing estimators, $\hat{\theta}_{extLf}$ and $\hat{\theta}_{Shen}$. The RMSEs of these are given by the following,

$$\begin{aligned} R(\hat{\theta}_{Lf}, \theta) &= e^{(\frac{2}{n}-1)\sigma^2} L_2 - 2e^{\frac{\sigma^2}{2}(\frac{1}{n}-1)} L_1 + 1, \\ R(\hat{\theta}_{Shen}, \theta) &= e^{(\frac{2}{n}-1)\sigma^2} F_2 - 2e^{\frac{\sigma^2}{2}(\frac{1}{n}-1)} F_1 + 1, \\ R(\hat{\theta}_{Prop}, \theta) &= e^{(\frac{2}{n}-1)\sigma^2} G_2 - 2e^{\frac{\sigma^2}{2}(\frac{1}{n}-1)} G_1 + 1, \end{aligned} \tag{4.1}$$

where, for $i = 1, 2$,

$$\begin{aligned} F_i &= E\left[\exp\left(\frac{i(n-1)S^2}{2(n+4)(n-1) + 3S^2}\right)\right], \\ L_i &= E\left[\exp\left(\frac{i(n-1)}{2} \cdot \frac{\exp((n-3)S^2/n(n^2-1)) - 1}{2\exp((n-3)S^2/n(n^2-1)) - 1}\right)\right], \\ G_i &= E\left[\exp\left(\frac{i(n-1)S^2}{2n(n+1) + 3.5S^2}\right)\right], \end{aligned}$$

respectively. In the research based on lognormal data, σ^2 are typically observed in $(0, 5)$ (Shen *et al.*, 2006; Zhou, 1998). Thus, in this numerical study, the values of σ^2 are selected as $0.1, 0.2, \dots, 4.9, 5.0$.

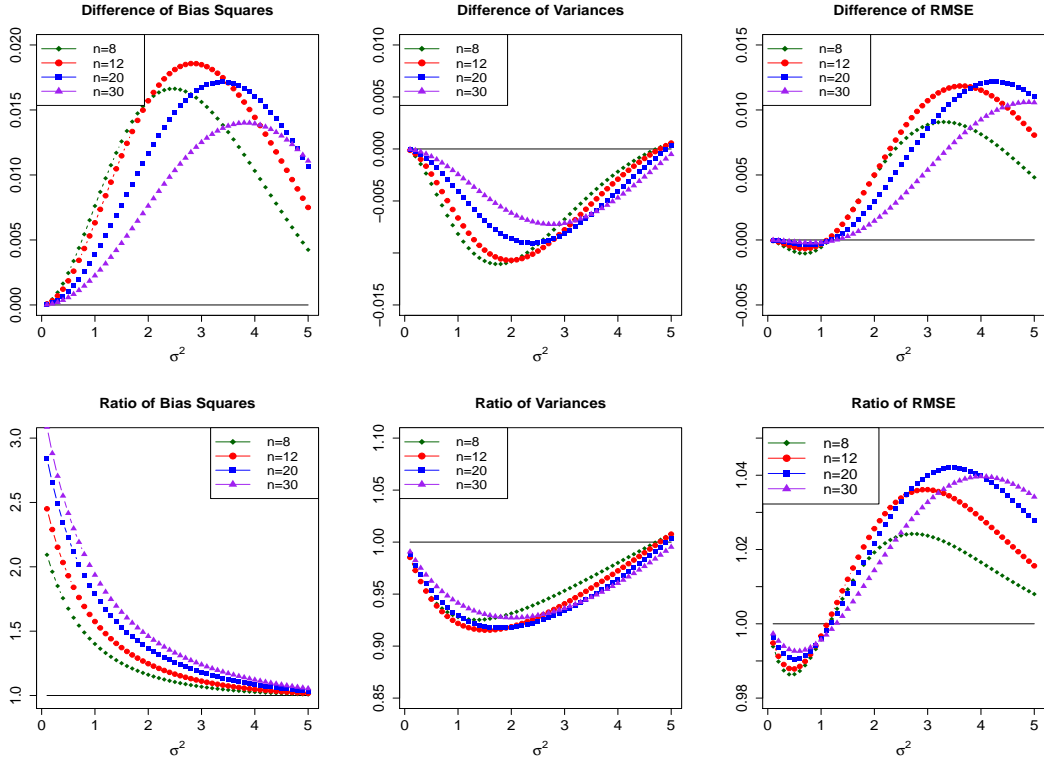


Figure 4: Risk comparison between $\hat{\theta}_{Shen}$ and $\hat{\theta}_{Prop}$ with small sample sizes, $n = 8, 12, 20$ and 30 .
 Upper: $Bias^2(\hat{\theta}_{Shen}) - Bias^2(\hat{\theta}_{Prop})$, $V(\hat{\theta}_{Shen}) - V(\hat{\theta}_{Prop})$, and $RMS E(\hat{\theta}_{Shen}, \theta) - RMS E(\hat{\theta}_{Prop}, \theta)$.
 Lower: $Bias^2(\hat{\theta}_{Shen})/Bias^2(\hat{\theta}_{Prop})$, $V(\hat{\theta}_{Shen})/V(\hat{\theta}_{Prop})$, and $RMS E(\hat{\theta}_{Shen})/RMS E(\hat{\theta}_{Prop})$.

We also take the sample size n to be 8, 12, 20, 30, 50, 75, 100, and 120. In addition, μ is set as $\mu = -\sigma^2/2$ throughout the numerical illustration.

First, $R(\hat{\theta}_{Shen}, \theta)$ and $R(\hat{\theta}_{Prop}, \theta)$ are compared in terms of the squared bias, variance, and RMSE. With $\mu = -\sigma^2/2$, the RMSE can be written as the sum of the squared bias and variance. The RMSE of the estimators in (4.1) can be calculated using numerical integration because S^2/σ^2 is a χ^2 random variable with the degrees of freedom $n - 1$. The upper panel of Figure 4 displays

$$Bias^2(\hat{\theta}_{Shen}) - Bias^2(\hat{\theta}_{Prop}), \quad V(\hat{\theta}_{Shen}) - V(\hat{\theta}_{Prop}), \quad R(\hat{\theta}_{Shen}, \theta) - R(\hat{\theta}_{Prop}, \theta),$$

from the left to the right. The lower panel of Figure 4 displays,

$$\frac{Bias^2(\hat{\theta}_{Shen})}{Bias^2(\hat{\theta}_{Prop})}, \quad \frac{V(\hat{\theta}_{Shen})}{V(\hat{\theta}_{Prop})}, \quad \frac{R(\hat{\theta}_{Shen}, \theta)}{R(\hat{\theta}_{Prop}, \theta)},$$

from the left to the right, respectively. Figure 4 displays the squared bias of $\hat{\theta}_{Prop}$, which is uniformly smaller than that of $\hat{\theta}_{Shen}$, and the variance of $\hat{\theta}_{Prop}$ is larger than that of $\hat{\theta}_{Shen}$ over $\sigma^2 \in (0, 5)$. Compared to $\hat{\theta}_{Shen}$, the proposed estimator makes a greater bias reduction than the variance increase and

Table 1: Relative mean squared error (RMSE) of $\hat{\theta}_{Lf}$, $\hat{\theta}_{Shen}$, and $\hat{\theta}_{Prop}$ under various distributions of σ^2

Distribution of σ^2	Est.	$n = 8$	$n = 12$	$n = 20$	$n = 30$	$n = 50$	$n = 100$
Uniform (0, 2)	$\hat{\theta}_{Lf}$	0.13208	0.09833	0.06546	0.04628	0.02954	0.01562
	$\hat{\theta}_{Shen}$	0.13005	0.09730	0.06512	0.04618	0.02952	0.01562
	$\hat{\theta}_{Prop}$	0.12934	0.09661	0.06467	0.04604	0.02947	0.01562
Gamma (1, 1)	$\hat{\theta}_{Lf}$	0.13216	0.09848	0.06878	0.04978	0.03319	0.01790
	$\hat{\theta}_{Shen}$	0.12944	0.09657	0.06801	0.04951	0.03315	0.01790
	$\hat{\theta}_{Prop}$	0.12851	0.09517	0.06680	0.04882	0.03289	0.01783
Gamma (2, 2)	$\hat{\theta}_{Lf}$	0.13388	0.09922	0.06566	0.04730	0.03066	0.01622
	$\hat{\theta}_{Shen}$	0.13204	0.09785	0.06519	0.04713	0.03064	0.01621
	$\hat{\theta}_{Prop}$	0.13156	0.09690	0.06451	0.04676	0.03053	0.01620
Inv. Gam. (3, 2)	$\hat{\theta}_{Lf}$	0.13091	0.09693	0.06640	0.04757	0.03059	0.01657
	$\hat{\theta}_{Shen}$	0.12894	0.09570	0.06594	0.04738	0.03054	0.01656
	$\hat{\theta}_{Prop}$	0.12867	0.09510	0.06550	0.04709	0.03041	0.01652
Inv. Gam. (6, 5)	$\hat{\theta}_{Lf}$	0.13223	0.09707	0.06432	0.04518	0.02874	0.01529
	$\hat{\theta}_{Shen}$	0.13071	0.09608	0.06402	0.04509	0.02871	0.01528
	$\hat{\theta}_{Prop}$	0.13071	0.09569	0.06384	0.04504	0.02867	0.01528
Gamma Mixture 0.5*G(4,1/5)+0.5*G(16,1/10)	$\hat{\theta}_{Lf}$	0.13306	0.09802	0.06679	0.04691	0.03045	0.01633
	$\hat{\theta}_{Shen}$	0.13080	0.09656	0.06641	0.04676	0.03044	0.01633
	$\hat{\theta}_{Prop}$	0.12985	0.09532	0.06576	0.04638	0.03039	0.01633

* Inv.Gam = Inverse gamma.

results in a decrease in the RMSE. Specifically, $R(\hat{\theta}_{Prop}, \theta)$ is smaller than $R(\hat{\theta}_{Shen}, \theta)$ when $\sigma^2 > 1$, and $R(\hat{\theta}_{Shen}, \theta)$ is smaller than $R(\hat{\theta}_{Prop}, \theta)$ when $0 < \sigma^2 < 1$. However, we can observe that the improvement in the RMSE caused by the proposed estimator over $\sigma^2 > 1$ is much greater than the deterioration of the RMSE over $0 < \sigma^2 < 1$.

Next, the RMSEs of $\hat{\theta}_{Lf}$, $\hat{\theta}_{Shen}$, and $\hat{\theta}_{Prop}$ are compared under various distributions of σ^2 including uniform, gamma, inverse gamma, and mixture gamma distributions. All of the distributions have the same mean, $E(\sigma^2) = 1$, and $N = 10,000$ values of σ^2 are randomly generated from each distribution. For each value of σ^2 and $\mu = -\sigma^2/2$, random numbers y_1, y_2, \dots, y_n are generated from the lognormal distribution with sample sizes 8, 12, 20, 30, 50, and 100. Based on the selected sample, we calculated $\hat{\theta}_{Lf}$, $\hat{\theta}_{Shen}$, $\hat{\theta}_{Prop}$, and their RMSEs. Table 1 lists the RMSE values for all the scenarios. For all cases, $\hat{\theta}_{Prop}$ uniformly outperforms the other two estimators regardless of the sample size n .

5. Conclusion

In this research, we proposed a new estimator, which is obtained by finding a tuning parameter to make the smallest average RMSE over a specific range of σ^2 values. The new estimator is primarily compared with Shen's estimator because it has been proven to almost uniformly outperform the other existing estimators, such as the sample mean, MLE, UMVUE, and estimators based on the conditional minimal MSE estimators. The proposed estimator exhibits comparable performance with Shen's estimator when σ^2 is small ($\sigma^2 < 1$) and demonstrates purposeful improvement with moderate and large σ^2 values in terms of the RMSE. Moreover, because the proposed estimator has a simpler form than the competing estimators, it is easy to compute in a real data analysis. Therefore, we suggest using the proposed estimator for the lognormal mean, unless one has strong evidence that σ^2 of the given lognormal data is very large.

Appendix A: Lemma 1

Lemma 1. For $i = 1, 2$,

$$\lim_{n \rightarrow \infty} Q_i(k) = \exp\left(\frac{i\sigma^2}{2}\right).$$

Proof: Let $U_n = iS^2/(n-1)$, $T_n = 2n(n+1)/(n-1)^2 + kS^2/(n-1)^2$ and $Z_n = U_n/T_n$. Since $S^2/(n-1) \rightarrow \sigma^2$, we have $U_n \xrightarrow{p} i\sigma^2$, $T_n \xrightarrow{p} 2$, and $Z_n \xrightarrow{p} i\sigma^2/2$ by Slutsky's theorem. Note that $Q_i(k) = E[\exp(Z_n)]$ and $\exp(\cdot)$ is continuous. Furthermore, since Z_n is a convergent sequence, it is bounded, and so is $\exp(Z_n)$. Now, since Z_n is bounded and converges to $i\sigma^2/2$, by Portmanteau lemma,

$$\lim_{n \rightarrow \infty} Q_i(k) = \lim_{n \rightarrow \infty} E[\exp(Z_n)] = E\left[\exp\left(\lim_{n \rightarrow \infty} Z_n\right)\right] = \exp\left(\frac{i\sigma^2}{2}\right).$$

□

Appendix B: Proof of Proposition 1

Proof: The proof of Proposition 1 is provided in Section 3.1. In Appendix B, we present alternative proof of Proposition 1 using Lemma 1. Recall, from (3.2), $R(\hat{\theta}_k, \theta) = e^{\sigma^2(2/n-1)} Q_2(k) - 2e^{\sigma^2(1/n-1)/2} Q_1(k) + 1$. Since $\lim_{n \rightarrow \infty} Q_2(k) = e^{\sigma^2}$ and $\lim_{n \rightarrow \infty} Q_1(k) = e^{\sigma^2/2}$,

$$\begin{aligned} \lim_{n \rightarrow \infty} R(\hat{\theta}_k, \theta) &= \lim_{n \rightarrow \infty} \left[e^{\sigma^2(\frac{2}{n}-1)} Q_2(k) - 2e^{\frac{\sigma^2}{2}(\frac{1}{n}-1)} Q_1(k) + 1 \right] \\ &= \lim_{n \rightarrow \infty} e^{\sigma^2(\frac{2}{n}-1)} \lim_{n \rightarrow \infty} Q_2(k) - 2 \lim_{n \rightarrow \infty} e^{\frac{\sigma^2}{2}(\frac{1}{n}-1)} \lim_{n \rightarrow \infty} Q_1(k) + 1 \\ &= e^{\frac{2\sigma^2}{n} - \sigma^2} e^{\sigma^2} - 2e^{\frac{\sigma^2}{2n} - \frac{\sigma^2}{2}} e^{\frac{\sigma^2}{2}} + 1 = 0. \end{aligned}$$

□

Appendix C: Proof of Proposition 2

Proof: From (2.2), we have $E(e^{\bar{X}}) = \theta e^{\sigma^2(1-n)/2n}$, and note that $(n-1)S^2/(2n(n+1) + kS^2) = (n-1)/k - 2(n-1)n(n+1)/k(2n(n+1) + kS^2)$. Thus, for $k \geq 1$,

$$\begin{aligned} E(\hat{\theta}_k) - \theta &= E\left[\exp\left(\bar{X} + \frac{(n-1)S^2}{2n(n+1) + kS^2}\right)\right] - \theta \\ &= \theta \left(\exp\left(\frac{\sigma^2(1-n)}{2n}\right) \exp\left(\frac{n-1}{k}\right) E\left[\exp\left(-\frac{2(n-1)n(n+1)}{k(2n(n+1) + kS^2)}\right)\right] - 1 \right) \\ &\leq \theta \left(\exp\left(\frac{\sigma^2(1-n)}{2n}\right) \exp\left(\frac{n-1}{k}\right) \exp\left(-\frac{2(n-1)n(n+1)}{k(2n(n+1) + kE(S^2))}\right) - 1 \right) \quad (\text{C.1}) \\ &= \theta \left(\exp\left(\frac{\sigma^2(1-n)}{2n}\right) \exp\left(\frac{n-1}{k}\right) \exp\left(-\frac{2(n-1)n(n+1)}{k(2n(n+1) + k\sigma^2(n-1))}\right) - 1 \right) \\ &= \theta \left(\exp\left(\frac{\sigma^2(1-n)}{2n}\right) \exp\left(\frac{n-1}{k}\right) \left(1 - \frac{2n(n+1)}{2n(n+1) + k\sigma^2(n-1)}\right) - 1 \right) \\ &= \theta \left(\exp\left(-\frac{\sigma^2(n-1)^2(4n/(n-1) + k\sigma^2)}{2n(2n(n+1) + k\sigma^2(n-1))}\right) - 1 \right) < 0. \end{aligned}$$

The inequality of (C.1) comes from Jensen's inequality. We can show $E(\hat{\theta}_{\text{Shen}}) - \theta < 0$ in the same way. Moreover, since $\sigma^2(n-1)^2(4n/(n-1) + k\sigma^2)/2n(2n(n+1) + k\sigma^2(n-1)) \rightarrow 0$, we have $\text{Bias}(\hat{\theta}_k) \rightarrow 0$, as $n \rightarrow \infty$. Now, for $k \leq 3$ and given $S^2 = s^2$, we have $2(n+4)(n-1) + 3s^2 > 2n(n+1) + ks^2$, which results in $\Pr(\hat{\theta}_{\text{Shen}} < \hat{\theta}_k | S^2) = 1$. Hence, $E(\hat{\theta}_{\text{Shen}} - \theta) < E(\hat{\theta}_k - \theta) < 0$. \square

Appendix D: Proof of Proposition 3

Proof:

$$\begin{aligned} V(\hat{\theta}_k) &= E(\hat{\theta}_k^2) - E^2(\hat{\theta}_k) \\ &= E\left[\exp\left(\bar{X} + \frac{(n-1)S^2}{2n(n+1) + kS^2}\right)\right]^2 - E^2\left[\bar{X} + \exp\left(\frac{(n-1)S^2}{2n(n+1) + kS^2}\right)\right] \\ &= E[\exp(2\bar{X})]E\left[\exp\left(\frac{2(n-1)S^2}{2n(n+1) + kS^2}\right)\right] - E^2[\exp(\bar{X})]E^2\left[\exp\left(\frac{(n-1)S^2}{2n(n+1) + kS^2}\right)\right] \\ &= \exp\left(2\mu + \frac{2\sigma^2}{n}\right)Q_2(k) - \exp\left(2\mu + \frac{\sigma^2}{n}\right)Q_1^2(k) \\ &= \exp\left(2\mu + \frac{\sigma^2}{n}\right)\left[\exp\left(\frac{\sigma^2}{n}\right)Q_2(k) - Q_1^2(k)\right], \end{aligned}$$

where $Q_i(k) = E[\exp(i(n-1)S^2/(2n(n+1) + kS^2))]$, for $i = 1, 2$. Let $W \sim \chi_{(n-1)}^2$ and $g(w)$ be a probability density function of W . Since $S^2/\sigma^2 \sim \chi_{(n-1)}^2$,

$$Q_i(k) = E\left[\exp\left(\frac{i\sigma^2(n-1)W}{2n(n+1) + k\sigma^2 W}\right)\right] = \int_0^\infty \exp\left(\frac{i\sigma^2(n-1)w}{2n(n+1) + k\sigma^2 w}\right)g(w)dw.$$

Furthermore, since $\lim_{n \rightarrow \infty} Q_i(k) = \exp(i\sigma^2/2)$, by Lemma 1,

$$\lim_{n \rightarrow \infty} \left[\exp\left(\frac{\sigma^2}{n}\right)Q_2(k) - Q_1^2(k)\right] = e^0 e^{\sigma^2} - \left(e^{\frac{\sigma^2}{2}}\right)^2 = 0,$$

resulting in $V(\hat{\theta}_k) \rightarrow 0$ as $n \rightarrow \infty$. \square

Appendix E: Proof of Proposition 4

Proof: Let

$$\mathbf{U}_n = \begin{bmatrix} \bar{X} \\ S_{n-1}^2 \end{bmatrix}, \boldsymbol{\eta} = \begin{bmatrix} \mu \\ \sigma^2 \end{bmatrix}, \text{ and } \psi(\mathbf{w}) = \exp\left(a + \frac{(n-1)b}{2n(n+1)/(n-1) + kb}\right),$$

where $S_{n-1}^2 = S^2/(n-1)$, and $\mathbf{w} = [a, b]^T$. Note that,

$$\psi(\mathbf{U}_n) = \exp\left[\bar{X} + \frac{(n-1)S_{n-1}^2}{2n(n+1)/(n-1) + kS_{n-1}^2}\right] = \exp\left[\bar{X} + \frac{(n-1)S^2}{2n(n+1) + kS^2}\right] = \hat{\theta}_k,$$

and

$$\psi(\boldsymbol{\eta}) = \exp\left[\mu + \frac{(n-1)\sigma^2}{2n(n+1)/(n-1) + k\sigma^2}\right].$$

Since \bar{X} and S_{n-1}^2 are independent, $\text{Cov}(\mathbf{U}_n) = \begin{bmatrix} \sigma^2/n & 0 \\ 0 & 2\sigma^4/(n-1) \end{bmatrix}$ and the gradient of $\psi(\boldsymbol{\eta})$ is given by

$$\nabla\psi(\boldsymbol{\eta}) = \left[\psi(\boldsymbol{\eta}), \psi(\boldsymbol{\eta}) \cdot \frac{2n(n+1)(n-1)^2}{(2n(n+1) + k\sigma^2(n-1))^2} \right]^T.$$

Since $\bar{X} \xrightarrow{P} \mu$ and $S_{n-1}^2 \xrightarrow{P} \sigma^2$, the first order Taylor expansion of $\psi(\mathbf{U}_n)$ is

$$\psi(\mathbf{U}_n) = \psi(\boldsymbol{\eta}) + [\nabla\psi(\boldsymbol{\eta})]^T (\mathbf{U}_n - \boldsymbol{\eta}) + o_p(1),$$

and the variance of $\hat{\theta}_k$ can be approximated by,

$$\begin{aligned} V(\hat{\theta}_k) &= V[\psi(\mathbf{U}_n)] \approx [\nabla\psi(\boldsymbol{\eta})]^T \text{Cov}(\mathbf{U}_n) [\nabla\psi(\boldsymbol{\eta})] \\ &= \psi^2(\boldsymbol{\eta}) \left[\frac{\sigma^2}{n} + \frac{2\sigma^4}{n-1} \cdot \frac{4n^2(n+1)^2(n-1)^4}{(2n(n+1) + k\sigma^2(n-1))^4} \right]. \end{aligned}$$

Since $\psi(\boldsymbol{\eta}) \rightarrow \theta$, $\sigma^2/n \rightarrow 0$, $2\sigma^4/(n-1) \rightarrow 0$ and $4n^2(n+1)^2(n-1)^4/(2n(n+1) + k\sigma^2(n-1))^4 \rightarrow 1/4$, we have $\lim_{n \rightarrow \infty} V[\psi(\mathbf{U}_n)] = 0$. □

Appendix F: Lemma 2

Lemma 2. Let Y be a chi-squared random variable with a probability density function $g(y)$. Define $h(w)$, c_2 and c_1 as,

$$h(w) = \frac{\sigma^2(n-1)w}{2n(n+1)}, \quad c_2 = e^{\sigma^2(\frac{3}{n}-1)}, \quad c_1 = e^{\frac{\sigma^2}{2}(\frac{1}{n}-1)}.$$

Then, for $n \geq 3$,

$$\int_0^\infty \left(\frac{c_1}{c_2} - e^{h(w)} \right) g(w) dw < 0.$$

Proof:

$$\begin{aligned} \int_0^\infty \left(\frac{c_1}{c_2} - e^{h(w)} \right) g(w) dw &= \int_0^\infty \left(e^{\frac{n-3}{2n}\sigma^2} - e^{\frac{\sigma^2(n-1)w}{2n(n+1)}} \right) g(w) dw \\ &= e^{\frac{n-3}{2n}\sigma^2} - \int_0^\infty e^{\frac{\sigma^2(n-1)w}{2n(n+1)}} g(w) dw \\ &= \sum_{j=0}^\infty \frac{1}{j!} \left(\frac{n-3}{2n}\sigma^2 \right)^j - \int_0^\infty \sum_{j=0}^\infty \left(\frac{\sigma^2(n-1)w}{2n(n+1)} \right)^j g(w) dw \quad (\text{F.1}) \\ &= \sum_{j=0}^\infty \frac{1}{j!} \left(\frac{n-3}{2n}\sigma^2 \right)^j - \sum_{j=0}^\infty \frac{1}{j!} \left(\frac{\sigma^2(n-1)}{2n(n+1)} \right)^j E(W^j) < 0. \end{aligned}$$

The inequality of the last part of (F.1) can be demonstrated by mathematical induction. Let $A(j)$ and $B(j)$ be the j^{th} term of

$$\sum_{j=0}^{\infty} \frac{1}{j!} \left(\frac{(n-3)\sigma^2}{2n} \right)^j \quad \text{and} \quad \sum_{j=0}^{\infty} \frac{1}{j!} \left(\frac{(n-1)\sigma^2}{2n(n+1)} \right)^j E(W^j),$$

respectively. Note that, $E(W^j) = \prod_{r=1}^j (n-3+2r) = (n-1)(n+1)(n+3) \cdots (n-3+2j)$. This implies $E(W^{j+1}) = E(W^j)(n+2j-1)$. Clearly, $A(0) = B(0)$. When $j = 1$,

$$A(1) = \frac{(n-3)\sigma^2}{2n} < B(1) = \frac{(n-1)\sigma^2}{2n(n+1)} E(W) = \frac{(n-1)^2\sigma^2}{2n(n+1)} = \frac{\sigma^2}{2n} \left(n-3 + \frac{4}{n+1} \right).$$

For $j \geq 2$, suppose $A(j) < B(j)$. Then, we obtain the following result:

$$\frac{A(j+1)}{B(j+1)} = \frac{A(j)}{B(j)} \cdot \frac{(n-3)(n+1)}{(n-1)(n+2j-1)} < \frac{n+1}{n+2j-1} < 1.$$

Therefore, $A(0) = B(0)$ and $A(j) < B(j)$, for all $j \geq 1$, resulting in $\sum_{j=0}^{\infty} A(j) - \sum_{j=0}^{\infty} B(j) < 0$. □

Appendix G: Proof of Theorem 1

Proof: From the definition of S^2 in (2.2), $W = S^2/\sigma^2 \sim \chi_{(n-1)}^2$. Then, with $g(w)$, which is a p.d.f. of the chi-squared distribution with degrees of freedom $n-1$, $Q_1(k)$ in (3.3) given by the following,

$$Q_1(k) = E \left[\exp \left(\frac{(n-1)S^2}{2n(n+1) + kS^2} \right) \right] = \int_0^{\infty} \exp \left(\frac{\sigma^2(n-1)w}{2n(n+1) + k\sigma^2w} \right) g(w) dw = \int_0^{\infty} q_1(w, k) dw, \tag{G.1}$$

where

$$q_1(w, k) = \exp \left(\frac{\sigma^2(n-1)w}{2n(n+1) + k\sigma^2w} \right) g(w). \tag{G.2}$$

Note that $q_1(w, k)$ is a continuous function, and $g(w) = w^{(n-1)/2-1} e^{-w/2} / (\Gamma((n-1)/2) \cdot 2^{(n-1)/2}) \leq 1/2$ for $n \geq 3$. First, we demonstrate that $q_1(w, k)$ and its improper integral are bounded. By deleting $2n(n+1)$ from $q_1(w, k)$, we obtain the following,

$$|q_1(w, k)| \leq e^{\frac{n-1}{k}} g(w) \leq \frac{1}{2} e^{\frac{n-1}{k}} \quad \text{and} \quad \int_0^{\infty} |q_1(w, k)| dw \leq \int_0^{\infty} e^{\frac{n-1}{k}} g(w) dw = e^{\frac{n-1}{k}}.$$

Furthermore, $\partial q_1(w, k)/\partial k$, which is continuous, and its improper integral is bounded, because,

$$\left| \frac{\partial q_1(w, k)}{\partial k} \right| = \frac{\sigma^4(n-1)w^2}{(2n(n+1) + k\sigma^2w)^2} q_1(w, k) \leq \frac{n-1}{k^2} q_1(w, k) \leq \frac{n-1}{2k^2} e^{\frac{n-1}{k}},$$

and

$$\int_0^{\infty} \left| \frac{\partial q_1(w, k)}{\partial k} \right| dw \leq \int_0^{\infty} \frac{n-1}{k^2} q_1(w, k) dw \leq \frac{n-1}{k^2} e^{\frac{n-1}{k}}.$$

Similarly, we can apply the same argument to $q_2(w, k)$ to demonstrate that $q_2(w, k)$, $\partial q_2(w, k)/\partial k$, and their improper integrals are bounded. To simplify the expressions, we define

$$h \equiv h(w, k) = \frac{\sigma^2(n-1)w}{2n(n+1) + k\sigma^2 w}, \quad c_2 = e^{\sigma^2(\frac{2}{n}-1)}, \quad c_1 = e^{\frac{\sigma^2}{2}(\frac{1}{n}-1)}, \quad (\text{G.3})$$

with

$$h' = -\frac{h^2}{(n-1)}, \quad h'' = -\frac{2}{n-1}hh'. \quad (\text{G.4})$$

Now, we can obtain the first and second derivatives of $R(\hat{\theta}_k, \theta)$ by interchanging the integral and derivative because, for $i = 1, 2$, all $q_i(w, k)$, $\partial q_i(w, k)/\partial k$, and their improper integrals are bounded,

$$\begin{aligned} \frac{\partial R(\hat{\theta}_k, \theta)}{\partial k} &= \frac{\partial}{\partial k} \int_0^\infty c_2 e^{2h(w,k)} g(w) dw - \frac{\partial}{\partial k} \int_0^\infty 2c_1 e^{h(w,k)} g(w) dw \\ &= \int_0^\infty \frac{\partial}{\partial k} c_2 e^{2h(w,k)} g(w) dw - \int_0^\infty \frac{\partial}{\partial k} 2c_1 e^{h(w,k)} g(w) dw \\ &= \int_0^\infty 2c_2 h'(w, k) e^{2h(w,k)} g(w) dw - \int_0^\infty 2c_1 h'(w, k) e^{h(w,k)} g(w) dw \quad (\text{G.5}) \\ &= \int_0^\infty 2c_2 h'(w, k) e^{h(w,k)} \left(e^{h(w,k)} - \frac{c_1}{c_2} \right) g(w) dw \\ &= \int_0^\infty \frac{2c_2}{n-1} [h(w, k)]^2 e^{h(w,k)} \left(\frac{c_1}{c_2} - e^{h(w,k)} \right) g(w) dw. \end{aligned}$$

Because $2c_2[h(w, k)]^2 e^{h(w,k)}/(n-1) > 0$, and is bounded for $w > 0$, the sign of $\partial R(\hat{\theta}_k, \theta)/\partial k$ depends on the sign of $\int_0^\infty (c_1/c_2 - e^{h(w,k)}) g(w) dw$. We can show that $1 \leq c_1/c_2 = e^{\sigma^2(n-3)/2n} < e^{\sigma^2/2}$, and $\lim_{k \rightarrow \infty} e^{h(w,k)} = 1$, implying that $\lim_{k \rightarrow \infty} \partial R(\hat{\theta}_k, \theta)/\partial k > 0$. However, for $k = 0$, $n \geq 3$, by Lemma 2, we have the following,

$$\int_0^\infty \left(\frac{c_1}{c_2} - e^{h(w,0)} \right) g(w) dw < 0,$$

implying $\partial R(\hat{\theta}_k, \theta)/\partial k|_{k=0} < 0$. Furthermore, we define $r_1(k)$ and $r_2(k)$ as

$$r_1(k) = \int_0^\infty 2c_2 h'(w, k) e^{2h(w,k)} g(w) dw \quad \text{and} \quad r_2(k) = - \int_0^\infty 2c_1 h'(w, k) e^{h(w,k)} g(w) dw,$$

respectively. Then, the third line of (G.5) states $\partial R(\hat{\theta}_k, \theta)/\partial k = r_1(k) + r_2(k)$. Recall that $h''(w, k) > 0$ from (G.3) and (G.4) because $h > 0$. Now, we observe that $r_1(k)$ is an increasing function, and $r_2(k)$ is a decreasing function because

$$\frac{\partial}{\partial k} r_1(k) = \int_0^\infty 2(c_2 h''(w, k) + 2c_2 (h'(w, k))^2) e^{2h(w,k)} g(w) dw > 0$$

and

$$\frac{\partial}{\partial k} r_2(k) = - \int_0^\infty 2(c_1 h''(w, k) + c_1 (h'(w, k))^2) e^{h(w,k)} g(w) dw < 0.$$

This implies that a unique critical point $k = k^*$ exists,

$$\int_0^{\infty} c_2 h'(w, k^*) e^{2h(w, k^*)} g(w) dw = \int_0^{\infty} c_1 h'(w, k^*) e^{h(w, k^*)} g(w) dw.$$

Moreover, we have the following

$$\begin{aligned} \frac{\partial^2 R(\hat{\theta}_k, \theta)}{\partial k^2} &= \int_0^{\infty} 2c_2 \left(h''(w, k) + 2[h'(w, k)]^2 \right) e^{2h(w, k)} g(w) dw \\ &\quad - \int_0^{\infty} 2c_1 \left(h''(w, k) + [h'(w, k)]^2 \right) e^{h(w, k)} g(w) dw \\ &= \int_0^{\infty} 2c_2 \left(\frac{2}{n-1} h(w, k) h'(w, k) + 2[h'(w, k)]^2 \right) e^{2h(w, k)} g(w) dw \\ &\quad - \int_0^{\infty} 2c_1 \left(\frac{2}{n-1} h(w, k) h'(w, k) + [h'(w, k)]^2 \right) e^{h(w, k)} g(w) dw \\ &= \int_0^{\infty} 2c_2 h'(w, k) e^{2h(w, k)} \left(\frac{2}{n-1} h(w, k) + 2h'(w, k) \right) g(w) dw \\ &\quad - \int_0^{\infty} 2c_1 h'(w, k) e^{h(w, k)} \left(\frac{2}{n-1} h(w, k) + h'(w, k) \right) g(w) dw, \end{aligned}$$

and at $k = k^*$,

$$\begin{aligned} \left. \frac{\partial^2 R(\hat{\theta}_k, \theta)}{\partial k^2} \right|_{k=k^*} &= \int_0^{\infty} 2c_2 h'(w, k^*) e^{2h(w, k^*)} \left(\frac{2}{n-1} h(w, k^*) + 2h'(w, k^*) \right) g(w) dw - \\ &\quad - \int_0^{\infty} 2c_1 h'(w, k^*) e^{2h(w, k^*)} \left(\frac{2}{n-1} h(w, k^*) + h'(w, k^*) \right) g(w) dw \\ &= \int_0^{\infty} 2c_2 [h'(w, k^*)]^2 e^{2h(w, k^*)} g(w) dw > 0. \end{aligned}$$

Therefore, a point $k = k^*$ exists, at which $R(\hat{\theta}_k, \theta)$ has a unique minimum. \square

References

- Chelmos D, Penzias AS, Kaufman G, and Cetrulo C (1995). Costs of triplet pregnancy, *American Journal of Obstetrics & Gynecology*, **172**, 677–682.
- Evans IG and Shaban SA (1974). A note on estimation in lognormal models, *Journal of the American Statistical Association*, **69**, 779–781.
- Finney DJ (1941). On the distribution of a variate whose logarithm is normally distributed, *Supplement to the Journal of the Royal Statistical Society*, **7**, 144–161.
- Hull JC (2018). *Options, Futures, and Other Derivatives*, Pearson, New York.
- Lacey LF, Keene ON, Pritchard JF, and Bye A (1997). Common noncompartmental pharmacokinetic variables: are they normally or log-normally distributed?, *Journal of Biopharmaceutical Statistics*, **7**, 171–178.
- Land C (1972). An evaluation of appropriate confidence interval methods for lognormal means, *Technometrics*, **14**, 145–158.

- Longford NT (2009). Inference with the lognormal distribution. *Journal of Statistical Planning and Inference*, **139**, 2329–2340.
- Ohring M and Kasprzak L (1998). *Reliability and Failure of Electronic Materials and Devices*, Academic Press, San Diego.
- Rukhin AL (1986). Improved estimation in lognormal models. *Journal of the American Statistical Association*, **81**, 1046–1049.
- Shen H, Brown LD, and Zhi H (2006). Efficient estimation of log-normal means with application to pharmacokinetic data, *Statistics in Medicine*, **25**, 3023–3038.
- Zellner A (1971). Bayesian and non-Bayesian analysis of the log-normal distribution and log-normal regression, *Journal of the American Statistical Association*, **66**, 327–330.
- Zhou XH, Hui SL, and Gao S (1997). Methods for comparing the means of two independent log-normal samples, *Biometrics*, **53**, 1129–1135.
- Zhou XH (1998). Estimation of the log-normal mean, *Statistics in Medicine*, **17**, 2251–2264.

Received January 07, 2021; Revised April 09, 2021; Accepted June 16, 2021