Confidence Intervals for Common Mean of Lognormal Distributions

Narudee Smithpreecha(✉), Sa-Aat Niwitpong(✉), and Suparat Niwitpong(✉)

Department of Applied Statistics, Faculty of Applied Science,
King Mongkut's University of Technology North Bangkok, Bangkok 10800, Thailand
narudee.s@rmutp.ac.th, {sa-aat.n,suparat.n}@sci.kmutnb.ac.th

Abstract. This paper presents new confidence intervals for the common mean of lognormal distributions by transforming the lognormal data. Three approaches were based on generalized confidence intervals (GCI) and adjusted method of variance estimates recovery (adjusted MOVER). A Monte Carlo simulation was used to assess the coverage probability and average length. The simulation study found that the adjusted MOVER approach based on Angus's conservative method (AM2) is appropriate and had the smallest coverage error in all of the scenarios. The generalized confidence interval approach (GCI) had the second smallest coverage error and had the smallest average lengths among the three approaches when the coverage probabilities were close to nominal level 0.95. Real data examples illustrate this approach.

1 Introduction

One of the important right skewed distributions with a long tail is lognormal distribution. It is widely used in many fields, such as environmental study, survival analysis, biostatistics and other statistical fields. The lognormal distribution has closely resembled a normal distribution. Simple to implement and easy to understand, by taking the natural logarithm of a random variable, the random variable will have a normal distribution.

Interval estimation of lognormal means for one, two and several populations have received widespread attention in papers of science and statistical literature. Statistical methods for interval estimation involving common mean for several lognormal distributions have also appeared frequently in many journals, such as biometrical journal by Tian and Wu [1] defined the concept of generalized variable and the large sample criteria to provide approach for the confidence interval estimation and hypothesis testing of the common mean of several lognormal populations. Lin and Wang [2] in a journal of applied statistics focused on making inferences on several log-normal means based on the modification of the quadratic method. There are also many other journals such as journal of statistical research by Ahmed et al. [3], journal of probability and statistical science by Baklizi and Ebrahem [4] and measurement science review by Cimermanová [5] interested in construction of the confidence intervals for common mean of several lognormal distributions.

L. H. Anh et al. (eds.), *Econometrics for Financial Applications*, Studies in Computational Intelligence 760, https://doi.org/10.1007/978-3-319-73150-6_22

In this paper, the interest is to construct confidence intervals for common mean of lognormal distributions. A simple approach to construct the confidence intervals for lognormal mean by transform the lognormal data would be to log-transfer data prior to analyzing statistical. There have been several researchers used the log-transformed data to construct confidence intervals for mean of lognormal distribution. For example, the paper by Krishnamoorthy and Mathew [6] and Olsson [7]. It seems evident that the results for interval estimation of the common mean of several lognormal populations by taking the natural logarithm have not been studied. Furthermore, there has not been much discussion on methods for interval estimation of common lognormal means by taking the natural logarithm of data. Therefore, researchers have proposed new simple approaches to construct the confidence intervals for the common lognormal mean. The first approach is generalized confidence intervals (GCI) which is based on the concepts of generalized confidence interval and was introduced by Weerahandi [8]. The second and the third approaches are adjusted method of variance estimates recovery approach (adjusted MOVER) based on cox's method (AM1) and Angus's conservative method (AM2) which are based on the concepts of the method of variance of estimates recovery (MOVER) introduced by Zou and Donner [9]. The GCI approach, the MOVER approach and the adjusted MOVER approach have been successfully used to construct the confidence interval for many common parameters. As reviewed in Tian [10], Tian and Wu [1], Krishnamoorthy and Lu [11], Ye et al. [12], Donner and Zou [13], Suwan and Niwitpong [14], Li et al. [15], Wongkhao [16] and Thangjai and Niwitpong [17]. Therefore, the focus is to develop interval estimation procedures with three approaches for the common mean of lognormal distributions and then compare them to each of the situation in terms of coverage probability and average length.

This paper is organized as follows. The properties of lognormal distribution and the parameter of interest will be briefly introduced in Sect. 2. The three approaches developed and descriptions of computational procedures are presented in Sect. 3. Section 4 presents simulation results to evaluate performances of the three approaches on coverage probabilities and average lengths. Section 5 illustrates the proposed approaches with real examples. Finally, conclusions are given in Sect. 6.

2 Lognormal Distribution and the Parameter of Interest

Let $Y_1, Y_2, ..., Y_n$ be a random variable having lognormal distribution with two parameters. This means that the log-transformed variables $X_1 = \log Y_1, X_2 = \log Y_2, ..., X_n = \log Y_n$ are normally distributed, that has mean value $E(X) = \mu$ and variance $var(X) = \sigma^2$. The mean of Y is $E(Y) = \exp\left(\mu + \frac{\sigma^2}{2}\right)$, by taking the natural logarithm of a random variable we get $\log(E(Y)) = \mu + \frac{\sigma^2}{2}$.

According to Olsson [7], An estimator of $\log(E(Y))$ can be calculated from sample data as $\log\left(E(\widehat{Y})\right) = \overline{X} + \frac{S^2}{2}$ and an estimator of the variance of $\log\left(E(\widehat{Y})\right)$ is given by $var\left[\log\left(E(\widehat{Y})\right)\right] = \frac{S^2}{2} + \frac{S^4}{2(n-1)}$.

Consider k independent lognormal populations with a common mean α. Let $Y_{i1}, Y_{i2}, ..., Y_{in_i}$ be a random sample from the i-th lognormal population as follows:

$$X_{ij} = \log Y_{ij} \sim \left(\mu_i, \sigma_i^2\right), \text{ for } i = 1, 2, ..., k, j = 1, 2, ..., n_i.$$

Thus, the common mean is $\alpha = \exp\left(\mu_i + \frac{\sigma_i^2}{2}\right)$, $\log \alpha = \left(\mu_i + \frac{\sigma_i^2}{2}\right)$.

3 The Approaches of Confidence Interval Estimation

3.1 The Generalized Confidence Interval Approach

Weerahandi [8] introduced the concept of generalized confidence intervals (GCI) which is based on the generalized pivotal quantity (GPQ) for a parameter of interest θ and ν is a vector of nuisance parameters. A generalized pivot $R(X, x, \theta, \nu)$ for interval estimation, where x is an observed value of X, as a random variable having the following two properties:

1. $R(X, x, \theta, \nu)$ has a distribution free of the vector of nuisance parameters ν.
2. The value of $R(X, x, \theta, \nu)$ is θ.

Let R_α be the 100α-th percentile of R. Then R_α becomes the $100(1-\alpha)\%$ lower bound for θ and $\left(R_{\alpha/2}, R_{1-\alpha/2}\right)$ becomes a $100(1-\alpha)\%$ two-side generalized confidence interval for θ.

Consider k independent lognormal populations with a common mean α.

Thus, we have $\alpha = \exp\left(\mu_i + \frac{\sigma_i^2}{2}\right)$.

The common log-mean, $\theta = \log \alpha = \left(\mu_i + \frac{\sigma_i^2}{2}\right)$.

Let $\overline{X}_i$ and S_i^2 denote the sample mean and variance for data X_{ij} for the i-th sample and let $\overline{x}_i$ and s_i^2 denote the observed sample mean and variance respectively.

Thus $\sigma_i^2 = \frac{(n_i-1)S_i^2}{V_i}$ *where* $V_i \sim \chi^2_{n_i-1}$.
where V_i is χ^2 variates with degrees of freedom and $n_i - 1$, we have the generalized pivot

$$R_{\sigma_i^2} = \frac{(n_i - 1)\, s_i^2}{V_i} \sim \frac{(n_i - 1)\, s_i^2}{\chi^2_{n_i-1}}. \tag{1}$$

The generalized pivotal quantity to estimate μ_i based on the i-th sample can be defined as

$$R_{\mu_i} = \overline{x_i} - \frac{Z_i}{\sqrt{U_i}} \sqrt{\frac{(n_i - 1)s_i^2}{n_i}}, \tag{2}$$

where Z_i and U_i denote standard normal variate and χ^2 variate with degree of freedom $n_i - 1$ respectively.

The generalized pivotal quantity for estimating θ based on the i-th sample is

$$R_\theta^{(i)} = R_{\mu_i} + \frac{R_{\sigma_i^2}}{2}. \tag{3}$$

From the i-th sample, the maximum likelihood estimator of θ is

$$\widehat{\theta}^{(i)} = \widehat{\mu}_i + \frac{\widehat{\sigma}_i^2}{2}, \qquad where\ \widehat{\mu}_i = \overline{X}_i, \widehat{\sigma}_i^2 = S_i^2. \tag{4}$$

The variance for $\widehat{\theta}^{(i)}$ is

$$var\left(\widehat{\theta}^{(i)}\right) = \frac{\sigma_i^2}{n_i} + \frac{\sigma_i^4}{2(n_i-1)}, \qquad see\ Olsson\ [7]. \tag{5}$$

According to Ye et al. [12], the generalized pivotal quantity proposed for the common log-mean $\theta = \log\alpha$ is a weighted average of the generalized pivot $R_\theta^{(i)}$ based on k individual samples defined as

$$R_\theta = \frac{\sum_{i=1}^{k} R_w R_\theta^{(i)}}{\sum_{i=1}^{k} R_{w_i}}, \tag{6}$$

where

$$R_{w_i} = \frac{1}{R_{var\left(\widehat{\theta}^{(i)}\right)}}, \tag{7}$$

$$R_{var\left(\hat{\theta}^{(i)}\right)} = \frac{R_{\sigma_i^2}}{n_i} + \frac{R_{\sigma_i^4}}{2(n_i-1)}. \tag{8}$$

That is, $R_{var\left(\widehat{\theta}^{(i)}\right)}$ is $var\left(\widehat{\theta}^{(i)}\right)$ with σ_i^2 replaced by $R_{\sigma_i^2}$.

$\left[L_{Gci}, U_{Gci}\right] = \left(R_{\alpha/2}, R_{1-\alpha/2}\right)$ is the $100\,(1-\alpha)\%$ two-side generalized confidence interval of the common log-mean $\theta = \log\alpha$.

$\left[\exp\left(L_{Gci}\right), \exp\left(U_{Gci}\right)\right] = \left(\exp\left(R_{\alpha/2}\right), \exp\left(R_{1-\alpha/2}\right)\right)$ is the $100\,(1-\alpha)\%$ two-side generalized confidence interval of the common mean θ.

Computing algorithms

For a given data set X_{ij} for $i = 1, 2, ..., k, j = 1, 2, ..., n_i$, the generalized confidence intervals for θ can be computed by the following steps.

1. Compute $\overline{x}_i$ and s_i^2 for $i = 1, 2, ..., k$.
2. Generate $V_i \sim \chi^2_{n_i-1}$ and then calculate $R_{\sigma_i^2}$ from (1) for $i = 1, 2, ..., k$.
4. Generate $Z_i \sim N(0,1)$ and $U_i \sim \chi^2_{n_i-1}$ then calculate R_{μ_i} from (2) for $i = 1, 2, ..., k$.
4. Calculate $R_\theta^{(i)}$ from (3) for $i = 1, 2, ..., k$.
5. Repeat steps 2–3, calculate R_{w_i} from (7) and (8) for $i = 1, 2, ..., k$.
6. Compute R_θ following (6).
7. Repeat step 2–6 a total m times and obtain an array of R_θ's.
8. Rank this array of R_θ's from small to large. The 100α-th percentile of R_θ's, $R_\theta(\alpha)$, is an estimate of the lower bound of the one-sided $100\,(1-\alpha)\%$ confidence interval and $\left(R_\theta(\alpha/2), R_\theta(1-\alpha/2)\right)$ is a two-sided $100\,(1-\alpha)\%$ confidence interval.
9. Calculate the interval length.
10. Count the number of successes in 5,000 independent generated datasets.
11. Calculate coverage probability and average length.

3.2 The Adjusted Method of Variance Estimates Recovery Approach

The concepts of the method of variance estimates recovery (the Mover approach) and the large sample method are used to create the adjusted method of variance estimates recovery (The adjusted MOVER approach).

The Mover approach was introduced by Zou and Donner [9] which considers two parameters $\theta_1 + \theta_2$ which have $100\,(1-\alpha)\%$ confidence limits (l_1, u_1) and (l_2, u_2), respectively. Under the assumption of the point estimates $\widehat{\theta_1}$ and $\widehat{\theta_2}$ are independence, the lower limit L and the upper limit U are given by

$$\left[L_1, U_1\right] = \left(\widehat{\theta_1} + \widehat{\theta_2}\right) \pm z_{\alpha/2}\sqrt{\widehat{var}\left(\widehat{\theta_1}\right) + \widehat{var}\left(\widehat{\theta_2}\right)}, \tag{9}$$

where $\widehat{var}\left(\widehat{\theta_i}\right) = \frac{\left(\widehat{\theta_i}-l_i\right)^2}{z^2_{\alpha/2}}, \quad \widehat{var}\left(\widehat{\theta_i}\right) = \frac{\left(u_i-\widehat{\theta_i}\right)^2}{z^2_{\alpha/2}}.$

For $i = 1, 2$. Using these estimates with from (9), two-side $100\,(1-\alpha)\%$ confidence limits for $\theta_1 + \theta_2$ given as

$$L = \left(\widehat{\theta_1} + \widehat{\theta_2}\right) - \sqrt{\left(\widehat{\theta_1} - l_1\right)^2 + \left(\widehat{\theta_2} - l_2\right)^2}$$

$$U = \left(\widehat{\theta_1} + \widehat{\theta_2}\right) + \sqrt{\left(U_1 + \widehat{\theta_1}\right)^2 + \left(U_2 + \widehat{\theta_2}\right)^2}.$$

Let $\theta^{(1)}, \theta^{(2)}, ..., \theta^{(k)}$ be k parameters of interest, where the estimates $\widehat{\theta}^{(1)}, \widehat{\theta}^{(2)}, ..., \widehat{\theta}^{(k)}$ are independent. Use concept of Donner and Zou [13] to construct of a $100\,(1-\alpha)\%$ two-sided confidence interval (L, U) for $\theta^{(1)}, \theta^{(2)}, ..., \theta^{(k)}$.

Thus $\left[L, U\right] = \left(\widehat{\theta}^{(1)} + \widehat{\theta}^{(2)} + ... + \widehat{\theta}^{(k)}\right) \pm z_{\alpha/2}\sqrt{\widehat{var}\left(\widehat{\theta}^{(1)}\right) + ... + \widehat{var}\left(\widehat{\theta}^{(k)}\right)}.$

Where the variance estimate for $\widehat{\theta}^{(i)}$ at $\theta^{(i)} = l_i$ is $\widehat{var}\left(\widehat{\theta}^{(i)}\right) = \frac{\left(\widehat{\theta}^{(i)}-l_i\right)^2}{z^2_{\alpha/2}}.$

And the variance estimate at $\theta^{(i)} = u_i$ is $\widehat{var}\left(\widehat{\theta}^{(i)}\right) = \frac{\left(u_i-\widehat{\theta}^{(i)}\right)^2}{z^2_{\alpha/2}}.$

Therefore, the lower limit L and upper limit U for $\theta^{(1)}, \theta^{(2)}, ..., \theta^{(k)}$ is given by

$$L = \left(\widehat{\theta}^{(1)} + \widehat{\theta}^{(2)} + ... + \widehat{\theta}^{(k)}\right) - \sqrt{\left(\widehat{\theta}^{(1)} - l_1\right)^2 + ... + \left(\widehat{\theta}^{(k)} - l_k\right)^2},$$

$$U = \left(\widehat{\theta}^{(1)} + \widehat{\theta}^{(2)} + ... + \widehat{\theta}^{(k)}\right) + \sqrt{\left(u_1 - \widehat{\theta}^{(1)}\right)^2 + ... + \left(u_k - \widehat{\theta}^{(k)}\right)^2}.$$

The next step uses concept of the method of the large sample by Tain and Wu [1] for parameter with pooled estimate of the common parameter from k populations. It is defined as

$$\widehat{\theta} = \frac{\sum_{i=1}^{k} \frac{\widehat{\theta}^{(i)}}{var\left(\widehat{\theta}^{(i)}\right)}}{\sum_{i=1}^{k} \frac{1}{var\left(\widehat{\theta}^{(i)}\right)}}, \tag{10}$$

which gives a variance estimate for $\widehat{\theta}^{(i)}$ at $\theta^{(i)} = l_i$ and $\theta^{(i)} = u_i$ of

$$var\left(\widehat{\theta}^{(i)}\right) = \frac{1}{2}\left(\frac{\left(\widehat{\theta}^{(i)} - l_i\right)^2}{z^2_{\alpha/2}} + \frac{\left(u_i - \widehat{\theta}^{(i)}\right)^2}{z^2_{\alpha/2}}\right). \tag{11}$$

Therefore, the lower limit L for the common parameter θ is given by

$$L = \widehat{\theta} - z_{1-\alpha/2}\sqrt{\frac{1}{\sum_{i=1}^{k}\frac{1}{\frac{\left(\widehat{\theta}^{(i)}-l_i\right)^2}{z^2_{\alpha/2}}}}}. \tag{12}$$

Similarly, the upper limit U for the common parameter θ is given by

$$U = \widehat{\theta} + z_{1-\alpha/2}\sqrt{\frac{1}{\sum_{i=1}^{k}\frac{1}{\frac{\left(u_i-\widehat{\theta}^{(i)}\right)^2}{z^2_{\alpha/2}}}}}. \tag{13}$$

Hence, the adjusted MOVER solution for confidence interval estimation is

$$\left(\widehat{\theta} - z_{1-\alpha/2}\sqrt{1/\sum_{i=1}^{k}\frac{z^2_{\alpha/2}}{\left(\widehat{\theta}^{(i)} - l_i\right)^2}},\quad \widehat{\theta} + z_{1-\alpha/2}\sqrt{1/\sum_{i=1}^{k}\frac{z^2_{\alpha/2}}{\left(u_i - \widehat{\theta}^{(i)}\right)^2}}\right).$$

From the i-th sample, where $i = 1, 2, ..., k$. The common log-mean θ is

$$\theta = \log\alpha = \left(\mu_i + \frac{\sigma_i^2}{2}\right).$$

The maximum likelihood estimator of common log-mean θ is

$$\widehat{\theta}^{(i)} = \widehat{\mu_i} + \frac{\widehat{\sigma}_i^2}{2}, \qquad where\ \widehat{\mu_i} = \overline{X}_i, \widehat{\sigma}_i^2 = S_i^2.$$

Chami et al. [18] have presented Cox's method and Angus's conservative method for constructing confidence interval for log-mean of lognormal.

According to Cox's method, the confidence interval (CI_C) for $\widehat{\theta}^{(i)}$ is

$$l_{i1} = \overline{X}_i + \frac{S_i^2}{2} - Z_{1-\alpha/2}\sqrt{\frac{S_i^2}{n_i} + \frac{S_i^4}{2(n_i - 1)}}, \tag{14}$$

$$u_{i1} = \overline{X}_i + \frac{S_i^2}{2} + Z_{1-\alpha/2}\sqrt{\frac{S_i^2}{n_i} + \frac{S_i^4}{2(n_i - 1)}}, \tag{15}$$

According to Angus's conservative method, the confidence interval (CI_A) for $\widehat{\theta}^{(i)}$ is

$$l_{i2} = \overline{X}_i + \frac{S_i^2}{2} - \frac{t_{1-\alpha/2}}{\sqrt{n_i}} \sqrt{S_i^2 \left(1 + \frac{S_i^2}{2}\right)}, \tag{16}$$

$$u_{i2} = \overline{X}_i + \frac{S_i^2}{2} + \frac{q_{\alpha/2}}{\sqrt{n_i}} \sqrt{S_i^2 \left(1 + \frac{S_i^2}{2}\right)}, \tag{17}$$

Which $t_{1-\alpha/2}$ be $1-\alpha$ percentile of a t-distribution with n_i-1 degrees of freedom, and let $q_{\alpha/2} = \sqrt{\frac{n}{2}\left(\frac{n-1}{\chi_\alpha^2} - 1\right)}$ where χ_α^2 is the α- percentile of the chi-square distributions with $n_i - 1$ degrees of freedom.

Researchers use the method of the large sample for log-mean with pooled estimate. It is defined in Eq. (10) and variance estimate for $\widehat{\theta}^{(i)}$ in Eq. (11). Consequently, L and U are defined in Eqs. (12) and (13). One gains these two groups of confidence intervals (CI_C) for $\widehat{\theta}^{(i)}$ in Eqs. (14), (15) and (CI_A) for $\widehat{\theta}^{(i)}$ in Eqs. (16) and (17) defines AM1 and AM2.

Hence, the adjusted MOVER approach based on cox's method (AM1) for confidence interval estimation of common log-mean $\theta = \log \alpha$ is $\left[L_{AM1}, U_{AM1}\right]$

$$= \left(\widehat{\theta} - z_{1-\alpha/2}\sqrt{1/\sum_{i=1}^{k}\frac{z_{\alpha/2}^2}{\left(\widehat{\theta}^{(i)} - l_{i1}\right)^2}}, \widehat{\theta} + z_{1-\alpha/2}\sqrt{1/\sum_{i=1}^{k}\frac{z_{\alpha/2}^2}{\left(u_{i1} - \widehat{\theta}^{(i)}\right)^2}}\right). \tag{18}$$

or the confidence interval for common mean α is $\left[\exp\left(L_{AM1}\right), \exp\left(U_{AM1}\right)\right]$.

The adjusted MOVER approach based on Angus's conservative method (AM2) for confidence interval estimation of common log-mean $\theta = \log \alpha$ is $\left[L_{AM2}, U_{AM2}\right]$

$$= \left(\widehat{\theta} - z_{1-\alpha/2}\sqrt{1/\sum_{i=1}^{k}\frac{z_{\alpha/2}^2}{\left(\widehat{\theta}^{(i)} - l_{i1}\right)^2}}, \widehat{\theta} + z_{1-\alpha/2}\sqrt{1/\sum_{i=1}^{k}\frac{z_{\alpha/2}^2}{\left(u_{i1} - \widehat{\theta}^{(i)}\right)^2}}\right). \tag{19}$$

or the confidence interval for common mean α is $\left[\exp\left(L_{AM2}\right), \exp\left(U_{AM2}\right)\right]$.

Computing algorithms

For a given data set X_{ij} for $i = 1, 2, ..., k, j = 1, 2, ..., n_i$, the adjusted method of variance estimates recovery for θ can be computed by the following steps.

1. Compute $\overline{x}_i$ and s_i^2 for $i = 1, 2, ..., k$.
2. Compute $l_{i1}, u_{i1}, l_{i2}, u_{i2}$ from (14), (15), (16), (17) for $i = 1, 2, ..., k$.
3. Calculate $var((\widehat{\theta}^{(i)})$ from (11) for $i = 1, 2, ..., k$.
4. Compute $\widehat{\theta}$ following (10).
5. Calculate confidence interval estimation from (18), (19) for $i = 1, 2, ..., k$.
6. Calculate the interval length.
7. Count the number of successes in 5,000 independent generated datasets.
8. Calculate coverage probability and average length.

4 Simulation Studies

A simulation study was performed with the coverage probabilities and average lengths of the common mean of the lognormal distributions for various combinations of the number of samples $k = 2$, 3 and 10, the sample sizes $n = (n_1, , ..., n_k)$ and the population variance $\sigma^2 = (\sigma_1^2, , ..., \sigma_k^2)$, the values were different and the common $\theta = \log\alpha$ take 0 and 10. In this simulation study, 95% confidence intervals from three approaches were compared, comprising of the proposed procedure generalized confidence interval approach (GCI), the adjusted MOVER approach based on cox's method (AM1) and based on Angus's conservative method (AM2). For each parameter setting, 5000 random samples were generated, 2500 R_θ's were obtained for each of the random samples.

In Tables 1, 2 and 3, the following notation applies $n = (n_1, , ..., n_k)$ and $\sigma^2 = (\sigma_1^2, , ..., \sigma_k^2)$. For $k = 2$, we have $n_1^{(2)} = (15, 15), n_2^{(2)} = (10, 20), n_3^{(2)} = (20, 10)$, $n_4^{(2)} = (50, 50), n_5^{(2)} = (50, 100)$ and $\sigma_1^{2(2)} = (1, 1), \sigma_2^{2(2)} = (1, 9), \sigma_3^{2(2)} = (1, 25)$, $\sigma_4^{2(2)} = (1, 100)$. For $k = 3$, we have $n_1^{(3)} = (15, 15, 15), n_2^{(3)} = (10, 15, 20), n_3^{(3)} = (20, 15, 10), n_4^{(3)} = (30, 50, 100)$ and $\sigma_1^{2(3)} = (0.02, 0.2, 2), \sigma_2^{2(3)} = (1, 1, 1), \sigma_3^{2(3)} = (1, 4, 9), \sigma_4^{2(3)} = (1, 9, 100)$. For $k = 10$, we have $n_1^{(10)} = (15, 15, 15, 15, 15, 15, 15, 15, 15, 15), n_2^{(10)} = (15, 15, 1515, 15, 10, 10, 10, 10, 10)$, $n_3^{(10)} = (30, 30, 30, 30, 30, 30, 30, 30, 30, 30), n_4^{(10)} = (50, 50, 80, 100, 100, 50, 50, 80, 100, 100)$ and $\sigma_1^{2(10)} = (1, 1, 1, 1, 1, 1, 1, 1, 1, 1)$, $\sigma_2^{2(10)} = (1, 1, 9, 9, 25, 25, 49, 49, 81, 81), \sigma_3^{2(10)} = (0.01, 0.01, 0.1, 0.1, 1, 1, 10, 10, 100, 100)$.

Tables 1, 2, and 3 presents the coverage probabilities and average lengths for 2, 3 and 10 sample cases respectively. In sample case 2, the GCI approach overestimated the coverage probabilities for all of the scenarios. In other sample cases, the GCI approach tended to overestimate the coverage probabilities for all of the scenarios but this depends on n_i and σ_i^2, especially when the n_i is small and σ_i^2 are similar values. It is also shown that the GCI approach tends to be drop from the nominal level 0.95. In all sample cases, The AM1 approach provides the underestimate coverage probabilities and has the coverage probabilities close to the nominal level 0.95 when the number of samples went up. The AM2 approach performs very well for all of the scenarios, although it tends to produce a wider average lengths more than the GCI approach, but the average length is slightly higher than the GCI approach.

In this paper, the average lengths of all intervals are considered since the approaches provide the coverage probability above the nominal level of 0.95 for all cases. Finally, it was discovered that the AM2 approach provided much better results over the another approaches in terms of average lengths for all cases. However, the average lengths of the GCI approach were the shortest when the coverage probabilities were close to the nominal level of 0.95.

5 An Empirical Application

In this section, two real data examples are exhibited to illustrate the generalized confidence interval approach (GCI), the adjusted MOVER approach based on

Table 1. Coverage probabilities (CP) and average length (AL) of approximate 95% two – side confidence bounds for common mean α of lognormal distributions, $\mu_i = \log\alpha - \left(\sigma_i^2/2\right)$ (based on 5000 simulations, m of gci = 2500): 2 sample cases.

n	σ^2	θ	GCI		AM1		AM2	
			CP	AL	CP	AL	CP	AL
$n_1^{(2)}$	$\sigma_1^{2(2)}$	0	0.9578	1.0027	0.9076	0.8422	0.9804	1.1652
		10	0.9546	1.0033	0.8968	0.8423	0.9796	1.1654
	$\sigma_2^{2(2)}$	0	0.9558	1.4735	0.9144	1.2177	0.9812	1.6822
		10	0.9514	1.4784	0.9166	1.2204	0.9820	1.6859
	$\sigma_3^{2(2)}$	0	0.9586	1.4822	0.9224	1.2424	0.9856	1.7172
		10	0.9586	1.4795	0.9286	1.2381	0.9834	1.7113
	$\sigma_4^{2(2)}$	0	0.9600	1.4808	0.9260	1.2457	0.9866	1.7219
		10	0.9540	1.4831	0.9236	1.2468	0.9814	1.7233
$n_2^{(2)}$	$\sigma_1^{2(2)}$	0	0.9502	1.0042	0.8968	0.8361	0.9790	1.1672
		10	0.9532	1.0042	0.8990	0.8356	0.9752	1.1664
	$\sigma_2^{2(2)}$	0	0.9540	1.9166	0.8954	1.4605	0.9744	2.0687
		10	0.9550	1.9141	0.8956	1.4592	0.9770	2.0669
	$\sigma_3^{2(2)}$	0	0.9574	1.9796	0.9064	1.5040	0.9770	2.1362
		10	0.9584	1.9909	0.9144	1.5111	0.9802	2.1461
	$\sigma_4^{2(2)}$	0	0.9616	2.0357	0.9172	1.5351	0.9806	2.1811
		10	0.9642	2.0363	0.9142	1.5376	0.9830	2.1848
$n_3^{(2)}$	$\sigma_1^{2(2)}$	0	0.9532	1.0051	0.8976	0.8368	0.9770	1.1683
		10	0.9514	1.0022	0.8942	0.8342	0.9748	1.1647
	$\sigma_2^{2(2)}$	0	0.9502	1.2441	0.9112	1.0588	0.9828	1.4620
		10	0.9552	1.2474	0.9142	1.0603	0.9814	1.4641
	$\sigma_3^{2(2)}$	0	0.9524	1.2442	0.9236	1.0769	0.9838	1.4866
		10	0.9510	1.2422	0.9240	1.0777	0.9828	1.4877
	$\sigma_4^{2(2)}$	0	0.9524	1.2339	0.9346	1.0842	0.9852	1.4965
		10	0.9538	1.2276	0.9308	1.0800	0.9860	1.4908
$n_4^{(2)}$	$\sigma_1^{2(2)}$	0	0.9524	0.5026	0.9388	0.4757	0.9854	0.6928
		10	0.9504	0.5014	0.9316	0.4747	0.9866	0.6913
	$\sigma_2^{2(2)}$	0	0.9524	0.6985	0.9428	0.6678	0.9854	0.9722
		10	0.9528	0.6993	0.9380	0.6691	0.9874	0.9741
	$\sigma_3^{2(2)}$	0	0.9534	0.7089	0.9420	0.6791	0.9864	0.9889
		10	0.9538	0.7086	0.9416	0.6786	0.9892	0.9881
	$\sigma_4^{2(2)}$	0	0.9540	0.7119	0.9450	0.6798	0.9894	0.9898
		10	0.9522	0.7122	0.9418	0.6804	0.9904	0.9907
$n_5^{(2)}$	$\sigma_1^{2(2)}$	0	0.9514	0.4037	0.9408	0.3887	0.9860	0.5935
		10	0.9544	0.4038	0.9438	0.3888	0.9862	0.5934
	$\sigma_2^{2(2)}$	0	0.9558	0.6884	0.9464	0.6597	0.9866	0.9643
		10	0.9512	0.6857	0.9386	0.6577	0.9886	0.9614
	$\sigma_3^{2(2)}$	0	0.9540	0.7080	0.9452	0.6782	0.9876	0.9882
		10	0.9482	0.7058	0.9424	0.6768	0.9868	0.9861
	$\sigma_4^{2(2)}$	0	0.9508	0.7123	0.9458	0.6801	0.9880	0.9903
		10	0.9528	0.7113	0.9454	0.6791	0.9862	0.9889

Table 2. Coverage probabilities (CP) and average length (AL) of approximate 95% two – side confidence bounds for common mean α of lognormal distributions, $\mu_i = \log\alpha - \left(\sigma_i^2/2\right)$ (based on 5000 simulations, m of gci = 2500): 3 sample cases.

n_i	σ_i^2	θ	GCI		AM1		AM2	
			CP	AL	CP	AL	CP	AL
$n_1^{(3)}$	$\sigma_1^{2(3)}$	0	0.9542	0.1518	0.9248	0.1335	0.9786	0.1867
		10	0.9516	0.1511	0.9256	0.1332	0.9762	0.1862
	$\sigma_2^{2(3)}$	0	0.9512	0.8110	0.8876	0.6755	0.9774	0.9350
		10	0.9486	0.8146	0.8856	0.6777	0.9746	0.9381
	$\sigma_3^{2(3)}$	0	0.9380	1.3733	0.8818	1.1257	0.9736	1.5537
		10	0.9380	1.3662	0.8802	1.1181	0.9710	1.5432
	$\sigma_4^{2(3)}$	0	0.9562	1.4534	0.9136	1.2115	0.9814	1.6738
		10	0.9528	1.4615	0.9134	1.2125	0.9800	1.6751
$n_2^{(3)}$	$\sigma_1^{2(3)}$	0	0.9540	0.1909	0.9158	0.1572	0.9776	0.2262
		10	0.9514	0.1917	0.9100	0.1582	0.9728	0.2276
	$\sigma_2^{2(3)}$	0	0.9468	0.8142	0.8774	0.6725	0.9756	0.9364
		10	0.9440	0.8107	0.8806	0.6706	0.9746	0.9337
	$\sigma_3^{2(3)}$	0	0.9382	1.6750	0.8744	1.3048	0.9656	1.8374
		10	0.9352	1.6612	0.8628	1.2942	0.9674	1.8230
	$\sigma_4^{2(3)}$	0	0.9524	1.9390	0.8898	1.4702	0.9738	2.0834
		10	0.9564	1.9388	0.8986	1.4699	0.9758	2.0829
$n_3^{(3)}$	$\sigma_1^{2(3)}$	0	0.9544	0.1301	0.9384	0.1181	0.9804	0.1645
		10	0.9552	0.1300	0.9362	0.1181	0.9798	0.1645
	$\sigma_2^{2(3)}$	0	0.9460	0.8152	0.8796	0.6739	0.9724	0.9384
		10	0.9448	0.8141	0.8844	0.6739	0.9730	0.9383
	$\sigma_3^{2(3)}$	0	0.9346	1.1852	0.8876	0.9956	0.9750	1.3737
		10	0.9422	1.1955	0.8902	1.0011	0.9768	1.3812
	$\sigma_4^{2(3)}$	0	0.9536	1.2248	0.9220	1.0622	0.9848	1.4657
		10	0.9510	1.2147	0.9212	1.0584	0.9842	1.4604
$n_4^{(3)}$	$\sigma_1^{2(3)}$	0	0.9514	0.0981	0.9398	0.0929	0.9800	0.1316
		10	0.9534	0.0982	0.9404	0.0929	0.9812	0.1317
	$\sigma_2^{2(3)}$	0	0.9516	0.3716	0.9372	0.3528	0.9844	0.5294
		10	0.9518	0.3716	0.9376	0.3530	0.9886	0.5297
	$\sigma_3^{2(3)}$	0	0.9508	0.8136	0.9274	0.7564	0.9918	1.0740
		10	0.9516	0.8139	0.9280	0.7565	0.9902	1.0741
	$\sigma_4^{2(3)}$	0	0.9504	0.9190	0.9308	0.8545	0.9878	1.1996
		10	0.9542	0.9169	0.9342	0.8523	0.9892	1.1966

Table 3. Coverage probabilities (CP) and average length (AL) of approximate 95% two – side confidence bounds for common mean α of lognormal distributions, $\mu_i = \log\alpha - \left(\sigma_i^2/2\right)$ (based on 5000 simulations, m of gci = 2500): 10 sample cases.

n_i	σ_i^2	θ	GCI		AM1		AM2	
			CP	AL	CP	AL	CP	AL
$n_1^{(10)}$	$\sigma_1^{2(10)}$	0	0.8806	0.4456	0.8020	0.3598	0.9586	0.4983
		10	0.8724	0.4450	0.7896	0.3593	0.9566	0.4976
	$\sigma_2^{2(10)}$	0	0.9248	0.9981	0.8730	0.8216	0.9694	1.1360
		10	0.9308	0.9933	0.8704	0.8180	0.9722	1.1311
	$\sigma_3^{2(10)}$	0	0.9600	0.0763	0.9226	0.0650	0.9734	0.0909
		10	0.9548	0.0761	0.9178	0.0648	0.9738	0.0907
$n_2^{(10)}$	$\sigma_1^{2(10)}$	0	0.8648	0.5010	0.7488	0.3852	0.9412	0.5403
		10	0.8674	0.5018	0.7624	0.3857	0.9444	0.5410
	$\sigma_2^{2(10)}$	0	0.9182	1.0171	0.8598	0.8205	0.9660	1.1346
		10	0.9274	1.0147	0.8730	0.8192	0.9706	1.1329
	$\sigma_3^{2(10)}$	0	0.9566	0.0763	0.9204	0.0649	0.9690	0.0908
		10	0.9576	0.0764	0.9178	0.0650	0.9692	0.0909
$n_3^{(10)}$	$\sigma_1^{2(10)}$	0	0.9052	0.2965	0.8744	0.2660	0.9874	0.3729
		10	0.9086	0.2967	0.8736	0.2661	0.9886	0.3730
	$\sigma_2^{2(10)}$	0	0.9482	0.6563	0.9202	0.5953	0.9848	0.8342
		10	0.9444	0.6564	0.9182	0.5947	0.9830	0.8333
	$\sigma_3^{2(10)}$	0	0.9532	0.0509	0.9318	0.0471	0.9702	0.0664
		10	0.9522	0.0509	0.9318	0.0471	0.9712	0.0663
$n_4^{(10)}$	$\sigma_1^{2(10)}$	0	0.9202	0.1791	0.9118	0.1713	0.9896	0.2615
		10	0.9304	0.1793	0.9216	0.1714	0.9890	0.2618
	$\sigma_2^{2(10)}$	0	0.9510	0.4992	0.9312	0.4582	0.9714	0.6694
		10	0.9526	0.5011	0.9316	0.4599	0.9680	0.6718
	$\sigma_3^{2(10)}$	0	0.9524	0.0372	0.9424	0.0355	0.9658	0.0524
		10	0.9502	0.0372	0.9390	0.0355	0.9618	0.0524

cox's method (AM1) and Angus's conservative method (AM2). All examples have been studied and reported by Lin and Wang [2]. The first example (A) was the medical charge data divided into two groups, 119 of them were an American group and 106 of them were the white group. The second example (B) was the pharmacokinetics data equally divided into three groups, 22 of them were group 1, group 2 and group 3. The data sets are presented in Table 4 and the results of confidence interval for two data sets are presented in Table 5. It can be seen that the interval of the adjusted MOVER approach based on Angus's conservative method (AM2) has the confidence interval close to the sample mean and wider lengths more than other approaches.

Table 4. The summary statistics of the log-transformed data sets.

Data set	n_i	$\overline{X}_i$	s_i^2	$\widehat{\theta}_i$
(A) The medical charge data				
American group	119	9.067	1.825	9.979
White group	106	8.693	2.693	10.039
(B) The pharmacokinetics data				
Group 1	22	2.601	0.24	2.721
Group 2	22	2.596	0.20	2.696
Group 3	22	2.599	0.17	2.684

Table 5. The results of confidence interval for four data sets.

The approaches	Confidence interval	Length
Data set (A) the medical charge data		
The generalized confidence interval approach (GCI)	(9.724, 10.288)	0.564
The adjusted MOVER approach		
Based on cox's method (AM1)	(9.723, 10.274)	0.551
Based on Angus's conservative method (AM2)	(9.722, 10.601)	0.879
Data set (B) the pharmacokinetics data		
The generalized confidence interval approach (GCI)	(2.582, 2.825)	0.243
The adjusted MOVER approach		
Based on cox's method (AM1)	(2.585, 2.811)	0.226
Based on Angus's conservative method (AM2)	(2.578, 2.893)	0.315

6 Discussion and Conclusions

This paper has presented three simple approaches to construct confidence intervals for common mean of lognormal distributions. The proposed confidence intervals were constructed by the generalized confidence interval approach (GCI), the adjusted MOVER approach based on cox's method (AM1) and based on Angus's conservative method (AM2). By the simulation studies, coverage probabilities form the generalized confidence interval approach (GCI) was always close to the nominal confidence level at 0.95. But there were a few cases, it seems that the generalized confidence interval approach (GCI) performed well only when σ_i^2 were more various. The adjusted MOVER approach based on cox's method (AM1) provided underestimated coverage probabilities for all cases. The adjusted MOVER approach based on Angus's conservative method (AM2) yielded coverage probabilities which tended to be high compared with the nominal level of 0.95 for all almost cases and the average lengths were wide to a little bit as compared with the GCI approach. Overall, the generalized confidence interval approach

(GCI) provided coverage probabilities close to nominal level 0.95 and average lengths is shorter than other approaches for $k = 2$. The adjusted MOVER approach based on Angus's conservative method (AM2) provides stable coverage probabilities for all k. In conclusion, the adjusted MOVER approach based on Angus's conservative method (AM2) can be successfully used to estimate the common mean of lognormal distributions.

The results of the generalized confidence interval approach (GCI) for common mean ($k \geq 2$) are similar to the simulation of the generalized confidence interval approach (GCI) for single mean of lognormal distribution which is studied by Olsson [7]. However, the coverage probabilities for $k \geq 2$ is decrease when k increased. In addition, this paper is constructing the confidence intervals for common mean of lognormal distributions by transform the lognormal data prior to use the generalized confidence interval approach (GCI). It is simple to construct the confidence intervals for the common mean. The results show that, it provided coverage probabilities close to nominal level 0.95. However, the results are not good equal to results of Lin and Wang [2], but average lengths are shorter.

Acknowledgments. The first author gratefully acknowledges the financial support from Rajamangala University of Technology Phra Nakhon of Thailand.

References

1. Tian, L., Wu, J.: Inferences on the common mean of several log-normal populations: the generalized variable approach. Biometrical J. **49**, 944–951 (2007)
2. Lin, S.H., Wang, R.S.: Modified method on the means for several log-normal distributions. J. Appl. Stat. **40**(1), 194–208 (2013)
3. Ahmed, S.E., Tomkins, R.J., Volodin, A.I.: Test of homogeneity of parallel samples from lognormal populations with unequal variances. J. Stat. Res. **35**(2), 25–33 (2001)
4. Baklizi, A., Ebrahem, M.: Interval estimation of common lognormal mean of several populations. J. Probab. Stat. Sci. **3**(1), 1–16 (2005)
5. Cimermanová, K.: Estimation of confidence intervals for the log-normal means and for the ratio and the difference of log-normal means with application of breath analysis. Measur. Sci. Rev. **7**(4), 31–36 (2007)
6. Krishnamoorthy, K., Mathew, T.: Inferences on the means of lognormal distributions using generalized p-values and generalized confidence intervals. J. Stat. Plann. Infer. **115**, 103–121 (2003)
7. Olsson, U.: Confidence intervals for the mean of a log-normal distribution. J. Stat. Educ. **13**(1) (2005). Available via DIALOG. www.amstat.org/puplications/jse/v13n1olsson.html. Accessed 11 Aug 2016
8. Weerahandi, S.: Generalized confidence intervals. J. Am. Stat. Assoc. **88**, 899–905 (1993)
9. Zou, G.Y., Donner, A.: Construction of confidence limits about effect measures: a general approach. Stat. Med. **27**, 1693–1702 (2008)
10. Tian, L.: Inferences on the common coefficient of variation. Stat. Med. **24**, 2213–2220 (2005)

11. Krishnamoorthy, K., Lu, Y.: Inference on the common means of several normal populations based on the generalized variable method. Biometrics **59**, 237–247 (2003)
12. Ye, R.D., et al.: Inferences on the common mean of several inverse Gaussian populations. Comput. Stat. Data Anal. **54**, 906–915 (2010)
13. Donner, A., Zou, G.Y.: Closed-form confidence intervals for function of the normal mean and standard deviation. Stat. Meth. Med. Res. **21**(4), 347–359 (2012)
14. Suwan, S., Niwitpong, S.: Estimated variance ratio confidence interval of nonnormal distributions. Far. East J. Math. Sci. **4**, 339–350 (2013)
15. Li, H.Q., et al.: Confidence intervals for ratio of two poisson rates using the method of variance estimates recovery. Comput. Stat. **29**, 869–889 (2014)
16. Wongkhao, A.: Confidence intervals for parameters of normal distribution. Ph.D. thesis: King Mongkut's University of Technology North Bangkok (2014)
17. Thangjai, W., Niwitpong, S.: Confidence intervals for the weighted coefficients of variation of two-parameter exponential distributions. Cogent Math. **4**, 1–16 (2017)
18. Chami, P., Antoine, R., Sahai, A.: On efficient confidence intervals for the lognormal mean. J. Appl. Sci. **7**(13), 1790–1794 (2007)