Thailand Statistician
April 2024; 22(2): 328-347
http://statassoc.or.th
Contributed paper

Confidence Intervals for Common Mean of Delta-Lognormal Distributions Based on Left-Censored Data with Application to Rainfall Data in Thailand

Warisa Thangjai [a] and Sa-Aat Niwitpong*[b]
[a] Department of Statistics, Faculty of Science, Ramkhamhaeng University, Bangkok, Thailand
[b] Department of Applied Statistics, Faculty of Applied Science,
King Mongkut's University of Technology North Bangkok, Bangkok, Thailand
*Corresponding author; e-mail: sa-aat.n@sci.kmutnb.ac.th

Received: 11 August 2023
Revised: 16 October 2023
Accepted: 30 October 2023

Abstract

In environmental data analysis, it is common to encounter left-censored data, such as rainfall and particulate matter data, which follow the delta-lognormal distribution. This paper focuses on estimating confidence intervals for the common mean of delta-lognormal distributions based on left-censored data. The confidence intervals are constructed using four approaches: the generalized confidence interval approach, the Bayesian approach, the parametric bootstrap approach, and the adjusted method of variance estimates recovery approach. The performance of these approaches is evaluated through Monte Carlo simulations using RStudio programming. The results reveal that for the number of sample cases $k = 3$, the generalized confidence interval approach and the adjusted method of variance estimates recovery approach performed very well when the sample sizes were small, whereas the Bayesian approach performed exceptionally well for moderate and large sample sizes. For the number of sample cases $k = 6$, the generalized confidence interval approach and the adjusted method of variance estimates recovery approach performed very well for small and moderate sample sizes, while the Bayesian approach excelled for large sample sizes. The results are illustrated with rainfall data from three regions of Thailand.

Keywords: Common mean, confidence interval, delta-lognormal distribution, left-censored data, rainfall data.

1. Introduction

Thailand is geographically divided into six regions, each with uneven precipitation patterns. Some regions receive minimal rainfall while others experience abundant rainfall. Consequently, droughts and floods frequently occur, posing natural disasters that require management rather than eradication. The common mean is used to describe the average rainfall in different areas. Daily rainfall data is typically modeled using the delta-lognormal distribution because it includes both zero and positive values. Various researchers, such as Maneerat et al. (2020), Maneerat et al. (2021), Thangjai et al. (2022), Yosboonruang et al. (2022), Thangjai and Niwitpong (2023a), Thangjai and Niwitpong (2023b), and Thangjai et al. (2023), have estimated confidence intervals for the parameters of this distribution, including mean functions, the common coefficient of variation, and percentile functions.

Censored data sets contain observations within a limited value range, with some data reported as zero due to being below the measurement threshold. Researchers, like Owen and DeRouen (1980), Glass and Gray (2001), Krishnamoorthy et al. (2011), Thangjai and Niwitpong (2023a), and Thangjai and Niwitpong (2023b), have efficiently estimated quantiles and other descriptive statistics of the underlying continuous distribution using censored data.

In cases where we have k sample cases and $k \geqslant 2$ independent samples, the confidence interval for the common mean derived from several independent delta-lognormal samples based on left-censored data becomes crucial. Several researchers, such as Krishnamoorthy and Lu (2003), Lin and Lee (2005), Tian (2005), Tian and Wu (2007), Ye et al. (2010), Ng (2014), Thangjai and Niwitpong (2017), Thangjai et al. (2017a), Thangjai et al. (2017b), Thangjai and Niwitpong (2018), Thangjai and Niwitpong (2020), Thangjai et al. (2020a), Thangjai et al. (2020b), and Thangjai et al. (2022), have studied this problem.

Interval estimators are preferred over point estimators because they provide a range of values likely to contain the unknown parameter of interest. However, there has been little statistical research on interval estimation for the common mean of delta-lognormal distributions based on left-censored data. This paper aims to address this gap by proposing confidence intervals using various approaches: the generalized confidence interval (GCI) approach, Bayesian approach, parametric bootstrap approach, and the adjusted method of variance estimates recovery (adjusted MOVER) approach. These methods have been compared and evaluated in previous studies. The GCI approach utilizes the generalized pivotal quantity (GPQ) for constructing the confidence interval. Several researchers have compared the GCI approach with other approaches for constructing the confidence interval, as documented in studies by Tian (2005), Chen and Zhou (2006), Tian and Wu (2007), Ye et al. (2010), and Thangjai et al. (2018). The Bayesian approach involves the utilization of the posterior probability and enables the comparison with other approaches for constructing confidence intervals. This comparison is supported by studies such as Rao and DCunha (2016), Ma and Chen (2018), and Thangjai et al. (2021). The parametric bootstrap approach uses the sampling distribution and has been utilized by researchers such as Padgett and Tomlinson (2003), Zhang (2015), and Altunkaynak and Gamgam (2019) for constructing confidence intervals. The adjusted MOVER approach uses the exact formula to construct the confidence interval and has been studied by Thangjai and Niwitpong (2017), Thangjai et al. (2017a), and Thangjai et al. (2017b).

The paper is organized as follows. Section 2 presents the methodologies for constructing confidence intervals for the common mean of delta-lognormal distributions based on left-censored data. Section 3 assesses the performance of these approaches through Monte Carlo simulations. Section 4 provides an example using rainfall data. Section 5 includes a discussion, and finally, Section 6 presents the conclusion.

2. Methods

For one population, let $n = n_0 + n_1$ be the sample size, where n_0 is the number of zero values and n_1 is the number of positive values. Moreover, the number of zero values n_0 has a binomial distribution, and the number of positive values n_1 has a log-normal distribution. Let $Z = (Z_1, Z_2, ..., Z_n)$ be the random variable from a delta log-normal distribution with parameters mean μ, variance σ^2, and the probability of obtaining a zero observation δ. The distribution of Z is given by

$$G(z_j;\mu,\sigma^2,\delta) = \begin{cases} \delta; & z_j = 0 \\ \delta + (1-\delta)F(z_j;\mu,\sigma^2); & z_j > 0 \end{cases}, \tag{1}$$

where $F(z_j;\mu,\sigma^2)$ is the log-normal distribution function and $j = 1, 2, ..., n$.

The population mean of Z is defined by

$$\upsilon = (1-\delta)\exp\left(\mu + \frac{1}{2}\sigma^2\right). \tag{2}$$

For estimating the mean of log-normal distribution based on left-censored data, let $X = (X_1, X_2, ..., X_n)$ be the random variable from the log-normal distribution with parameters μ and σ^2. Let ξ be some censoring point and let $n = n_1 + n_2$ observations be a sample size, where n_1 observations less than or equal to $\log(\xi)$ are not known but they are assumed to be nonzero and $n_2 = n - n_1$ observations greater than some censoring point $\log(\xi)$. Suppose that $Y_j = \log(X_j)$ is the observations above $\log(\xi)$ and has the normal distribution, where $j = 1, 2, ..., n_2$. Let $h = \frac{n_1}{n}$ be the fraction of observations in the sample that is below $\log(\xi)$. The mean of Y is given by

$$\bar{Y} = \frac{1}{n_2}\sum_{j=1}^{n_2} Y_j.$$

The variance of Y is given by

$$S^2 = \frac{1}{n_2}\sum_{j=1}^{n_2}(Y_j - \bar{Y})^2. \tag{3}$$

Suppose that ϕ and Φ are the density function and the distributions function of the standard normal distribution, respectively. According to Krishnamoorthy et al. (2011), the maximum likelihood estimators of μ and σ^2 are given by

$$\hat{\mu} = \bar{Y} - \psi(h, a)\left(\bar{Y} - \log(\xi)\right)$$

$$\text{and} \qquad \hat{\sigma}^2 = S^2 + \psi(h, a)\left(\bar{Y} - \log(\xi)\right)^2,$$

where

$$a = \frac{\log(\xi) - \mu}{\sigma}, \quad W(a) = \frac{\phi(a)}{1 - \Phi(a)}, \quad V(h, a) = \frac{hW(-a)}{1 - h}, \quad \psi(h, a) = \frac{V(h, a)}{V(h, a) - a}.$$

The mean of log-normal distribution based on left-censored data is

$$\theta = \exp\left(\mu + \frac{1}{2}\sigma^2\right).$$

Therefore, the estimator of the mean of log-normal distribution based on left-censored data is

$$\hat{\theta} = \exp\left(\hat{\mu} + \frac{1}{2}\hat{\sigma}^2\right).$$

For k populations, for $i = 1, 2, ..., k$, let $X_i = (X_{i1}, X_{i2}, ..., X_{in_i})$ be the random variable from the log-normal distribution with parameters μ_i and σ_i^2. Let ξ_i be some censoring points and let $n_i = n_{i(1)} + n_{i(2)}$ observations be a sample sizes, where $n_{i(1)}$ observations less than or equal to $\log(\xi_i)$ and $n_{i(2)} = n_i - n_{i(1)}$ observations greater than some censoring point $\log(\xi_i)$. Suppose that $Y_{ij} = \log(X_{ij})$ is the observations above $\log(\xi_i)$, where $i = 1, 2, ..., k$ and $j = 1, 2, ..., n_{i(2)}$. Let $h_i = \frac{n_{i(1)}}{n_i}$ be the fraction of observations in the sample that is below $\log(\xi_i)$. Let $\bar{Y}_i$ and S_i^2 be the mean and variance of Y_i, respectively. The maximum likelihood estimators of μ_i and σ_i^2 are given by

$$\hat{\mu}_i = \bar{Y}_i - \psi(h_i, a_i)\left(\bar{Y}_i - \log(\xi_i)\right) \tag{4}$$

$$\text{and} \qquad \hat{\sigma}_i^2 = S_i^2 + \psi(h_i, a_i)\left(\bar{Y}_i - \log(\xi_i)\right)^2, \tag{5}$$

where

$$a_i = \frac{\log(\xi_i) - \mu_i}{\sigma_i}, W(a_i) = \frac{\phi(a_i)}{1 - \Phi(a_i)}, V(h_i, a_i) = \frac{h_i W(-a_i)}{1 - h_i}, \psi(h_i, a_i) = \frac{V(h_i, a_i)}{V(h_i, a_i) - a_i}. \tag{6}$$

The mean of log-normal distribution based on left-censored data is given by

$$\theta_i = \exp\left(\mu_i + \frac{1}{2}\sigma_i^2\right).$$

Therefore, the estimator of the mean of log-normal distribution based on left-censored data is given by

$$\hat{\theta}_i = \exp\left(\hat{\mu}_i + \frac{1}{2}\hat{\sigma}_i^2\right).$$

The variance of $\hat{\theta}_i$ is given by

$$Var(\hat{\theta}_i) = \left(\frac{2\hat{\sigma}_i^2 + \hat{\sigma}_i^4}{2n_i}\right)\exp\left(2\hat{\mu}_i + (\hat{\sigma})_i^2\right). \tag{7}$$

The common mean of log-normal distributions based on left-censored data is the weighted average of $\hat{\theta}_i$ based on k individual sample which is defined by

$$\hat{\theta} = \sum_{i=1}^{k} \frac{\hat{\theta}_i}{Var(\hat{\theta}_i)} \Bigg/ \sum_{i=1}^{k} \frac{1}{Var(\hat{\theta}_i)},$$

where $Var(\hat{\theta}_i)$ is defined in Eqn. (7).

2.1. Generalized confidence interval approach

A generalized pivotal quantity (GPQ) is used to construct the generalized confidence interval.
Definition: Let $R(X; x, \mu, \sigma^2, \delta)$ be a function of X, z, μ, σ^2, and δ. Then $R(X; x, \mu, \sigma^2, \delta)$ is said to be the GPQ if it satisfies the following conditions:
1. $R(X; x, \mu, \sigma^2, \delta)$ has a probability distribution free of unknown parameters.
2. The observed pivotal of $R(X; x, \mu, \sigma^2, \delta)$ does not depend on the nuisance parameter.

The $100(1-\alpha)\%$ two-sided confidence intervals for the common mean of delta-lognormal distributions based on left-censored data can be constructed using $[R(\alpha/2), R(1-\alpha/2)]$, where $R(\alpha/2)$ denote the $(\alpha/2)$-th quantile of $R(X; x, \mu, \sigma^2, \delta)$ and $R(1-\alpha/2)$ denote the $(1-\alpha/2)$-th quantile of $R(X; x, \mu, \sigma^2, \delta)$.

The GPQ for μ_i is given by

$$R_{\mu_i} = \hat{\mu}_i - \frac{\hat{\mu}_i^*}{\hat{\sigma}_i^*}\hat{\sigma}_i, \tag{8}$$

where $\hat{\mu}_i^*$ and $\hat{\sigma}_i^*$ are the maximum likelihood estimators based on a censored sample from standard normal distribution and $\hat{\mu}_i$ and $\hat{\sigma}_i$ are defined in Eqn. (4) and Eqn. (5).

The GPQ for σ_i is given by

$$R_{\sigma_i} = \frac{\hat{\sigma}_i}{\hat{\sigma}_i^*}, \tag{9}$$

where $\hat{\sigma}_i^*$ is the maximum likelihood estimator based on a censored sample from standard normal distribution.

The GPQ for θ_i is given by

$$R_{\theta_i} = \exp\left(R_{\mu_i} + \frac{1}{2}\left(R_{\sigma_i}\right)^2\right), \tag{10}$$

where R_{μ_i} and R_{σ_i} are defined in Eqn. (23) and Eqn. (24), respectively. The GPQ for $Var\left(\hat{\theta}_i\right)$ is given by

$$R_{Var(\hat{\theta}_i)} = \left(\frac{2\left(R_{\sigma_i}\right)^2 + \left(R_{\sigma_i}\right)^4}{2n_i}\right)\exp\left(2R_{\mu_i} + \left(R_{\sigma_i}\right)^2\right), \tag{11}$$

where R_{μ_i} and R_{σ_i} are defined in Eqns. (8) and (9), respectively.

The GPQ for the common mean of delta-lognormal distributions based on left-censored data is a weighted average of the GPQ of θ_i based on k individual sample. The GPQ for the common mean of delta-lognormal distributions based on left-censored data is given by

$$R_\theta = \sum_{i=1}^{k} \frac{R_{\theta_i}}{R_{Var(\hat{\theta}_i)}} \Bigg/ \sum_{i=1}^{k} \frac{1}{R_{Var(\hat{\theta}_i)}}, \tag{12}$$

where R_{θ_i} is defined in Eqn. (10) and $R_{Var(\hat{\theta}_i)}$ is defined in Eqn. (11).

Therefore, the $100(1-\alpha)\%$ two-sided confidence interval for the common mean of delta-lognormal distributions based on left-censored data using the GCI approach is given by

$$CI_{GCI} = [L_{GCI}, U_{GCI}] = [R_\theta(\alpha/2), R_\theta(1-\alpha/2)],$$

where $R_\theta(\alpha/2)$ and $R_\theta(1-\alpha/2)$ denote the $100(\alpha/2)$-th and $100(1-\alpha/2)$-th percentiles of R_θ, respectively.

Algorithm 1 is used to construct the GCI for the common mean of delta-lognormal distributions based on left-censored data.

Algorithm 1.

Step 1: Generate sample from the standard normal distribution and compute $\hat{\mu}_i^*$ and $\hat{\sigma}_i^*$

Step 2: Compute R_{μ_i} from Eqn. (8)

Step 3: Compute R_{σ_i} from Eqn. (9)

Step 4: Compute R_{θ_i} from Eqn. (10)

Step 5: Compute $R_{Var(\hat{\theta}_i})$ from Eqn. (11)

Step 6: Compute R_θ from Eqn. (12)

Step 7: Repeat step 1 - step 6, a total m times and obtain an array of R_θs

Step 8: Compute L_{GCI} and U_{GCI}

2.2. Bayesian approach

The Bayesian approach is a method for updating probabilities using the Bayes' theorem. In this paper, the Jeffreys Independence prior is updated using Bayes' rule. The prior is defined as $p(\mu_i, \sigma_i^2) = p(\mu_i)p(\sigma_i^2)$. The posterior distribution of σ_i^2 is the inverse gamma distribution. The posterior distribution of σ_i^2 is defined by

$$\sigma_i^2 | y_i \sim IG\left(\frac{n_{i(2)}-1}{2}, \frac{(n_{i(2)}-1)\hat{\sigma}_i^2}{2}\right), \tag{13}$$

where $\hat{\sigma}_i^2$ is defined in Eqn. (5).

The posterior distribution of μ_i given σ_i^2 is the normal distribution. The posterior distribution of μ_i given σ_i^2 is defined by

$$\mu_i | \sigma_i^2, y_i \sim N\left(\hat{\mu}_i, \frac{\sigma_i^2}{n_{i(2)}}\right), \tag{14}$$

where $\hat{\mu}_i$ is defined in Eqn. (4) and σ_i^2 is defined in Eqn. (13).

The posterior distribution of θ_i is defined by

$$\theta_i = \exp\left(\mu_i + \frac{1}{2}\sigma_i^2\right), \tag{15}$$

where σ_i^2 and μ_i are defined in Eqn. (13) and Eqn. (14), respectively.

The variance of θ_i is given by

$$Var(\theta_i) = \left(\frac{2\sigma_i^2 + \sigma_i^4}{2n_i}\right) \exp\left(2\mu_i + \sigma_i^2\right), \tag{16}$$

where σ_i^2 and μ_i are defined in Eqn. (13) and Eqn. (14), respectively.

The common mean of delta-lognormal distributions based on left-censored data is defined by

$$\theta_{BS} = \sum_{i=1}^{k} \frac{\theta_i}{Var(\theta_i)} \Bigg/ \sum_{i=1}^{k} \frac{1}{Var(\theta_i)}, \tag{17}$$

where θ_i is defined in Eqn. (15) and $Var(\theta_i)$ is defined in Eqn. (16).

Therefore, the $100(1-\alpha)\%$ two-sided credible interval for the common mean of delta-lognormal distributions based on left-censored data using the Bayesian approach is given by

$$CI_{BS} = [L_{BS}, U_{BS}],$$

where L_{BS} and U_{BS} denote the lower and upper limits of the shortest $100(1-\alpha)\%$ highest posterior density interval of θ_{BS}, respectively.

Algorithm 2 is used to construct the Bayesian credible interval for the common mean of delta-lognormal distributions based on left-censored data.

Algorithm 2.
Step 1: Compute $\sigma_i^2|y_i$ from Eqn. (13)
Step 2: Compute $\mu_i|\sigma_i^2, y_i$ from Eqn. (14)
Step 3: Compute θ_i from Eqn. (15)
Step 4: Compute $Var(\theta_i)$ from Eqn. (16)
Step 5: Compute θ_{BS} from Eqn. (17)
Step 6: Repeat step 1 - step 5, a total m times and obtain an array of θ_{BS}s
Step 7: Compute L_{BS} and U_{BS}

2.3. Parametric bootstrap approach

Let $Y_{i1}^*, Y_{i2}^*, ..., Y_{in_{i(2)}}^*$ be the sample with replacement from $Y_{i1}, Y_{i2}, ..., Y_{in_{i(2)}}$. Moreover, let $\bar{Y}_i^*$ be the estimator of the population mean. It is given by

$$\bar{Y}_i^* = \frac{1}{n_{i(2)}} \sum_{j=1}^{n_{i(2)}} Y_{ij}^*. \tag{18}$$

Suppose S_i^{2*} is the estimator of the population variance. It is given by

$$S_i^{2*} = \frac{1}{n_{i(2)}} \sum_{j=1}^{n_{i(2)}} \left(Y_{ij}^* - \bar{Y}_i^*\right)^2. \tag{19}$$

The maximum likelihood estimators of μ_i^* and σ_i^{2*} are given by

$$\hat{\mu}_i^* = \bar{Y}_i^* - \psi(h_i, a_i)\left(\bar{Y}_i^* - \log(\xi_i)\right) \tag{20}$$

and

$$\hat{\sigma}_i^{2*} = S_i^{2*} + \psi(h_i, a_i)\left(\bar{Y}_i^* - \log(\xi_i)\right)^2, \tag{21}$$

where $\psi(h_i, a_i)$ is defined by Eqn. (6).

The estimator of the mean of log-normal distribution based on left-censored data is given by

$$\hat{\theta}_i^* = \exp\left(\hat{\mu}_i^* + \frac{1}{2}\hat{\sigma}_i^{2*}\right). \tag{22}$$

The variance of $\hat{\theta}_i^*$ is given by

$$Var(\hat{\theta}_i^*) = \left(\frac{2\hat{\sigma}_i^{2*} + \hat{\sigma}_i^{4*}}{2n_i}\right)\exp\left(2\hat{\mu}_i^* + \hat{\sigma}_i^{2*}\right). \tag{23}$$

The common mean of log-normal distributions based on left-censored data is the weighted average of $\hat{\theta}_i^*$ based on k individual sample which is defined by

$$\hat{\theta}^* = \sum_{i=1}^{k} \frac{\hat{\theta}_i^*}{Var(\hat{\theta}_i^*)} \Big/ \sum_{i=1}^{k} \frac{1}{Var(\hat{\theta}_i^*)}, \tag{24}$$

where $Var(\hat{\theta}_i^*)$ is defined in Eqn. (23).

The lower and upper limits of the confidence interval for the common mean of delta-lognormal distributions based on left-censored data using the parametric bootstrap approach are given by

$$L_{PB} = \overline{\hat{\theta}^*} - z_{1-\alpha/2} sd(\hat{\theta}^*) \tag{25}$$

$$\text{and} \qquad U_{PB} = \overline{\hat{\theta}^*} + z_{1-\alpha/2} sd(\hat{\theta}^*), \tag{26}$$

where $\overline{\hat{\theta}^*}$ is the mean of $\hat{\theta}^*$, $sd(\hat{\theta}^*)$ is the standard deviation of $\hat{\theta}^*$, and $z_{1-\alpha/2}$ is the $100(1-\alpha/2)$-th percentile of the standard normal distribution.

Therefore, the $100(1-\alpha)\%$ two-sided confidence interval for the common mean of delta-lognormal distributions based on left-censored data using the parametric bootstrap approach is given by

$$CI_{PB} = [L_{PB}, U_{PB}],$$

where L_{PB} and U_{PB} are defined in Eqn. (25) and Eqn. (26), respectively.

Algorithm 3 is used to construct the parametric bootstrap confidence interval for the common mean of delta-lognormal distributions based on left-censored data.

Algorithm 3.

Step 1: In parametric bootstrapping, the underlying assumption is that the data originates from a recognized distribution with unspecified parameters. To address this, we calculate these parameters based on the data and subsequently employ the estimated distributions to generate simulated samples. Generate $y_{i1}^*, y_{i2}^*, ..., y_{in_{i(2)}}^*$ from normal distributions with $\hat{\mu}_i$ and $\hat{\sigma}_i^2$

Step 2: Compute $\bar{y}_i^*$ from Eqn. (18)

Step 3: Compute s_i^{2*} from Eqn. (19)

Step 4: Compute $\hat{\mu}_i^*$ from Eqn. (20)

Step 5: Compute $\hat{\sigma}_i^{2*}$ from Eqn. (21)

Step 6: Compute $\hat{\theta}_i^*$ from Eqn. (22)

Step 7: Compute $Var(\hat{\theta}_i^*)$ from Eqn. (23)

Step 8: Compute $\hat{\theta}^*$ from Eqn. (24)

Step 9: Repeat step 1 - step 8, a total m times and obtain an array of $\hat{\theta}^*$s

Step 10: Compute L_{PB} and U_{PB}

2.4. Adjusted method of variance estimates recovery approach

Thangjai and Niwitpong (2023a) proposed the confidence interval for the mean of delta-lognormal distribution based on left-censored data. The lower and upper limits for the mean of delta-lognormal distribution based on left-censored data are defined by

$$l_{\theta_i} = \exp\left[\left(\hat{\mu}_i + \frac{1}{2}\hat{\sigma}_i^2\right) - \sqrt{(\hat{\mu}_i - l_{\mu_i})^2 + \left(\frac{1}{2}\hat{\sigma}_i^2 - \frac{1}{2}l_{\sigma_i^2}\right)^2}\right]$$

$$\text{and} \qquad u_{\theta_i} = \exp\left[\left(\hat{\mu}_i + \frac{1}{2}\hat{\sigma}_i^2\right) + \sqrt{(u_{\mu_i} - \hat{\mu}_i)^2 + \left(\frac{1}{2}u_{\sigma_i^2} - \frac{1}{2}\hat{\sigma}_i^2\right)^2}\right],$$

where $$l_{\mu_i} = \hat{\mu}_i - z_{1-\alpha/2}\sqrt{\frac{(n_{i(2)}-1)\hat{\sigma}_i^2}{n_{i(2)}\chi^2_{n_{i(2)}-1}}}, \quad u_{\mu_i} = \hat{\mu}_i + z_{1-\alpha/2}\sqrt{\frac{(n_{i(2)}-1)\hat{\sigma}_i^2}{n_{i(2)}\chi^2_{n_{i(2)}-1}}}$$
$$l_{\sigma_i^2} = \frac{(n_{i(2)}-1)\hat{\sigma}_i^2}{\chi^2_{1-\alpha/2,n_{i(2)}-1}}, \quad u_{\sigma_i^2} = \frac{(n_{i(2)}-1)\hat{\sigma}_i^2}{\chi^2_{\alpha/2,n_{i(2)}-1}},$$

and $\hat{\mu}_i$ is defined in Eqn. (4), $\hat{\sigma}_i$ is defined in Eqn. (5), $z_{1-\alpha/2}$ is the $100(1-\alpha/2)$-th percentile of the standard normal distribution, $\chi^2_{n_{i(2)}-1}$ is the chi-squared distribution with $n_{i(2)} - 1$ degrees of freedom, and $\chi^2_{1-\alpha/2,n_{i(2)}-1}$ and $\chi^2_{\alpha/2,n_{i(2)}-1}$ are the $100(1-\alpha/2)$-th and $100(\alpha/2)$-th percentiles of the chi-squared distribution with $n_{i(2)} - 1$ degrees of freedom, respectively.

The common mean of log-normal distributions based on left-censored data is the weighted average of $\hat{\theta}_i$ based on k individual sample which is defined by

$$\hat{\theta} = \sum_{i=1}^{k} \frac{\hat{\theta}_i}{Var(\hat{\theta}_i)} \bigg/ \sum_{i=1}^{k} \frac{1}{Var(\hat{\theta}_i)},$$

where $$\hat{\theta}_i = \exp\left(\hat{\mu}_i + \frac{1}{2}\hat{\sigma}_i^2\right) \text{ and } \widehat{Var}(\hat{\theta}_i) = \frac{1}{2}\left(\frac{\left(\hat{\theta}_i - l_{\theta_i}\right)^2}{z^2_{\alpha/2}} + \frac{\left(u_{\theta_i} - \hat{\theta}_i\right)^2}{z^2_{\alpha/2}}\right).$$

According to Thangjai and Niwitpong (2020), the lower and upper limits of the confidence interval for the weighted average of θ_i based on adjusted MOVER approach are defined by

$$L_{AM} = \hat{\theta} - \sqrt{\sum_{i=1}^{k} \frac{\left(\hat{\theta}_i - l_{\theta_i}\right)^2}{\left(\widehat{Var}\left(\hat{\theta}_{l_i}\right)\right)^2} \bigg/ \sum_{i=1}^{k} \frac{1}{\left(\widehat{Var}\left(\hat{\theta}_{l_i}\right)\right)^2}} \tag{27}$$

and $$U_{AM} = \hat{\theta} + \sqrt{\sum_{i=1}^{k} \frac{\left(u_{\theta_i} - \hat{\theta}_i\right)^2}{\left(\widehat{Var}\left(\hat{\theta}_{u_i}\right)\right)^2} \bigg/ \sum_{i=1}^{k} \frac{1}{\left(\widehat{Var}\left(\hat{\theta}_{u_i}\right)\right)^2}}, \tag{28}$$

where $$\widehat{Var}\left(\hat{\theta}_{l_i}\right) = \frac{\left(\hat{\theta}_i - l_{\theta_i}\right)^2}{z^2_{\alpha/2}}, \quad \widehat{Var}\left(\hat{\theta}_{u_i}\right) = \frac{\left(u_{\theta_i} - \hat{\theta}_i\right)^2}{z^2_{\alpha/2}}.$$

Therefore, the $100(1-\alpha)\%$ two-sided confidence interval for the common mean of delta-lognormal distributions based on left-censored data using the adjusted MOVER approach is given by

$$CI_{AM} = [L_{AM}, U_{AM}],$$

where L_{AM} and U_{AM} are defined in Eqns. (27) and (28), respectively.

3. Results

In this section, we present the numerical results obtained from Monte Carlo simulation studies, which are used to compare the four approaches proposed in Section 2. The coverage probability and average length of the confidence intervals were calculated using the RStudio programming. The performance of the confidence intervals is assessed using the coverage probability and the average length. The best-performing confidence interval is identified as having a coverage probability greater than or equal to 0.95 and the shortest average length. The RStudio programming language is used to construct the confidence intervals based on the GCI, Bayesian, parametric bootstrap, and adjusted MOVER approaches. Each approach is executed with $M =$ 5,000 runs, and the GCI, Bayesian, and parametric bootstrap approaches use $m =$ 2,500 runs.

In this study, five configuration factors were considered to evaluate the performance of the confidence intervals: the number of groups, sample sizes, population means, population standard deviations, probabilities of obtaining zero observation, and censoring points. Algorithm 4 is used to compute the coverage probabilities and average lengths of the confidence intervals for the common mean of delta-lognormal distributions based on left-censored data.

Algorithm 4.

Step 1: Generate z_i from delta-lognormal distributions with parameters μ_i, σ_i, and δ_i and set x_i from log-normal distributions with parameters μ_i and σ_i, where $i = 1, 2, ..., k$
Step 2: Compute $y_i = \log(x_i)$ and select $y_i > \log(\xi_i)$
Step 3: Compute $n_{i(1)}$, $n_{i(2)}$, $\hat{\mu}_i$, $\hat{\sigma}_i$, $\hat{\theta}_i$, and $\hat{\theta}$
Step 4: Construct the confidence intervals CI_{GCI}, CI_{BS}, CI_{PB}, and CI_{AM}
Step 5: If $L \leqslant \theta \leqslant U$, set $p = 1$; else set $p = 0$
Step 6: Compute $U - L$
Step 7: Repeat step 1 - step 6, a total M times
Step 8: Compute mean of p defined by the coverage probability
Step 9: Compute mean of $U - L$ defined by the average length

For the number of sample cases $k = 3$, the population means, population standard deviations, probabilities of obtaining zero observation, and censoring points are presented in Table 1. The sample sizes were set as (n_1, n_2, n_3) = (20,20,20), (30,30,30), (20,20,30), (50,50,50), (30,30,50), (100,100,100), and (50,50,100). From Table 2, the results indicate that the coverage probabilities of the adjusted MOVER approach were greater than the nominal confidence level of 0.95 for all cases. The GCI and adjusted MOVER approaches were greater than the nominal confidence level of 0.95 for small sample sizes ($n_i \leqslant$ 20). Moreover, the average lengths of the GCI and adjusted MOVER approaches were shorter than the average length of the adjusted MOVER approach when the sample sizes were small. Therefore, the GCI and adjusted MOVER approaches are recommended for constructing the confidence interval when the sample sizes are small. Additionally, the Bayesian approach performed satisfactorily in terms of coverage probability and average length for moderate and large sample sizes.

For the number of sample cases $k = 6$, the population means, population standard deviations, probabilities of obtaining zero observation, and censoring points are presented in Table 3. The sample sizes were set as $(n_1, n_2, n_3, n_4, n_5, n_6)$ = (20,20,20,20,20,20), (30,30,30,30,30,30), (20,20,20,30,30,30), (50,50,50,50,50,50), (30,30,30,50,50,50), (100,100,100,100,100,100), and (50,50,50,100,100,100). From Table 4, the results show that the GCI and the adjusted MOVER approach performed satisfactorily in terms of coverage probability and average length for small and moderate sample sizes. Additionally, the Bayesian approach provided the best credible interval for large sample sizes.

Table 1 Values selected for the population means, population standard deviations, probabilities of obtaining zero observation, and censoring points: 3 sample cases

Run number	(μ_1, μ_2, μ_3)	$(\sigma_1, \sigma_2, \sigma_3)$	$(\delta_1, \delta_2, \delta_3)$	(ξ_1, ξ_2, ξ_3)
1	(0.00,0.00,0.00)	(1.00,1.00,1.00)	(0.10,0.10,0.10)	(0.10,0.10,0.10)
2	(0.00,0.00,0.00)	(1.00,1.00,1.00)	(0.10,0.10,0.10)	(0.10,0.10,0.25)
3	(0.00,0.00,0.00)	(1.00,1.00,1.00)	(0.10,0.10,0.25)	(0.10,0.10,0.10)
4	(0.00,0.00,0.00)	(1.00,1.00,1.00)	(0.10,0.10,0.25)	(0.10,0.10,0.25)
5	(0.00,0.00,0.00)	(1.00,1.00,2.00)	(0.10,0.10,0.10)	(0.10,0.10,0.10)
6	(0.00,0.00,0.00)	(1.00,1.00,2.00)	(0.10,0.10,0.10)	(0.10,0.10,0.25)
7	(0.00,0.00,0.00)	(1.00,1.00,2.00)	(0.10,0.10,0.25)	(0.10,0.10,0.10)
8	(0.00,0.00,0.00)	(1.00,1.00,2.00)	(0.10,0.10,0.25)	(0.10,0.10,0.25)

Table 2 The coverage probabilities (CPs) and average lengths (ALs) of 95% two-sided confidence intervals for the common mean of delta-lognormal distributions based on left-censored data: 3 sample cases

(n_1, n_2, n_3)	Run number	CP (AL)			
		CI_{GCI}	CI_{BS}	CI_{PB}	CI_{AM}
(20,20,20)	1	0.9468	0.9098	0.7590	0.9810
		(1.5410)	(1.4516)	(1.1274)	(2.9523)
	2	0.9342	0.8856	0.7162	0.9712
		(1.5063)	(1.3700)	(1.0803)	(2.7297)
	3	0.9582	0.9266	0.7896	0.9814
		(1.7308)	(1.6458)	(1.2143)	(3.4897)
	4	0.9304	0.8838	0.6962	0.9802
		(1.6113)	(1.4907)	(1.1173)	(3.1804)
	5	0.9686	0.9472	0.8292	0.9778
		(2.1064)	(1.9209)	(1.4099)	(3.6141)
	6	0.9714	0.9518	0.8282	0.9786
		(2.1248)	(1.9369)	(1.4119)	(3.6956)
	7	0.9704	0.9482	0.8316	0.9800
		(2.1200)	(1.9408)	(1.4225)	(3.7170)
	8	0.9652	0.9408	0.8142	0.9802
		(2.0744)	(1.8877)	(1.3833)	(3.5957)
(30,30,30)	1	0.9616	0.9356	0.8210	0.9856
		(1.3008)	(1.2409)	(0.9765)	(2.2711)
	2	0.9448	0.9108	0.7742	0.9810
		(1.2636)	(1.1580)	(0.9283)	(2.0616)
	3	0.9682	0.9454	0.8338	0.9878
		(1.4369)	(1.3877)	(1.0447)	(2.5271)
	4	0.9538	0.9108	0.7530	0.9870
		(1.3612)	(1.2738)	(0.9748)	(2.3667)
	5	0.9818	0.9652	0.8752	0.9868
		(1.7310)	(1.6157)	(1.2156)	(2.5664)
	6	0.9792	0.9658	0.8814	0.9878
		(1.7123)	(1.5999)	(1.2085)	(2.5550)
	7	0.9804	0.9660	0.8720	0.9896
		(1.7132)	(1.5985)	(1.2070)	(2.5699)
	8	0.9784	0.9626	0.8746	0.9918
		(1.7063)	(1.5891)	(1.2019)	(2.5657)
(20,20,30)	1	0.9530	0.9184	0.7868	0.9854
		(1.4522)	(1.3778)	(1.0808)	(2.6118)
	2	0.9344	0.8926	0.7390	0.9744
		(1.4082)	(1.2769)	(1.0206)	(2.2924)
	3	0.9626	0.9316	0.7936	0.9870
		(1.6605)	(1.5794)	(1.1763)	(3.3816)
	4	0.9286	0.8776	0.6890	0.9758
		(1.5022)	(1.3830)	(1.0539)	(2.7844)
	5	0.9722	0.9502	0.8402	0.9780
		(2.1906)	(1.9909)	(1.4414)	(3.7320)
	6	0.9718	0.9512	0.8368	0.9796
		(2.1506)	(1.9597)	(1.4205)	(3.6412)
	7	0.9758	0.9528	0.8294	0.9842
		(2.1461)	(1.9474)	(1.4112)	(3.6694)
	8	0.9690	0.9460	0.8246	0.9780
		(2.1171)	(1.9180)	(1.3968)	(3.6484)

Table 2 Continued

(n_1, n_2, n_3)	Run number	CP (AL)			
		CI_{GCI}	CI_{BS}	CI_{PB}	CI_{AM}
(50,50,50)	1	0.9766 (1.0328)	0.9598 (0.9960)	0.8850 (0.7902)	0.9962 (1.7003)
	2	0.9638 (1.0087)	0.9328 (0.9283)	0.8252 (0.7552)	0.9900 (1.5560)
	3	0.9850 (1.1473)	0.9706 (1.1198)	0.9004 (0.8537)	0.9948 (1.8696)
	4	0.9672 (1.0736)	0.9294 (1.0165)	0.8044 (0.7905)	0.9944 (1.7279)
	5	0.9836 (1.3461)	0.9740 (1.2795)	0.9146 (0.9843)	0.9920 (1.8455)
	6	0.9858 (1.3262)	0.9784 (1.2691)	0.9214 (0.9814)	0.9944 (1.8245)
	7	0.9860 (1.3278)	0.9732 (1.2606)	0.9068 (0.9707)	0.9946 (1.8330)
	8	0.9866 (1.3300)	0.9750 (1.2669)	0.9182 (0.9775)	0.9942 (1.8431)
(30,30,50)	1	0.9728 (1.1995)	0.9506 (1.1506)	0.8542 (0.9139)	0.9910 (1.9614)
	2	0.9464 (1.1454)	0.9068 (1.0412)	0.7928 (0.8488)	0.9794 (1.6842)
	3	0.9790 (1.3698)	0.9562 (1.3222)	0.8566 (1.0044)	0.9902 (2.4498)
	4	0.9472 (1.2442)	0.9082 (1.1543)	0.7422 (0.8889)	0.9862 (2.0535)
	5	0.9806 (1.7838)	0.9670 (1.6589)	0.8824 (1.2404)	0.9876 (2.5643)
	6	0.9806 (1.7642)	0.9682 (1.6570)	0.8886 (1.2446)	0.9850 (2.5725)
	7	0.9790 (1.7408)	0.9632 (1.6180)	0.8706 (1.2111)	0.9866 (2.5510)
	8	0.9780 (1.7431)	0.9628 (1.6246)	0.8744 (1.2178)	0.9874 (2.5826)
(100,100,100)	1	0.9900 (0.7350)	0.9836 (0.7128)	0.9434 (0.5692)	0.9992 (1.1939)
	2	0.9810 (0.7205)	0.9582 (0.6633)	0.8956 (0.5465)	0.9964 (1.0901)
	3	0.9896 (0.8166)	0.9858 (0.8018)	0.9416 (0.6177)	0.9990 (1.2826)
	4	0.9774 (0.7679)	0.9540 (0.7333)	0.8738 (0.5731)	0.9976 (1.2068)
	5	0.9850 (0.9429)	0.9872 (0.9095)	0.9526 (0.7123)	0.9952 (1.2424)
	6	0.9846 (0.9337)	0.9856 (0.9107)	0.9484 (0.7155)	0.9982 (1.2477)
	7	0.9860 (0.9300)	0.9862 (0.8976)	0.9462 (0.7016)	0.9980 (1.2468)
	8	0.9868 (0.9256)	0.9856 (0.8983)	0.9422 (0.7045)	0.9970 (1.2433)

Table 2 Continued

(n_1, n_2, n_3)	Run number	CP (AL)			
		CI_{GCI}	CI_{BS}	CI_{PB}	CI_{AM}
(50,50,100)	1	0.9830	0.9732	0.9092	0.9974
		(0.9027)	(0.8737)	(0.6980)	(1.3941)
	2	0.9620	0.9294	0.8372	0.9888
		(0.8606)	(0.7785)	(0.6426)	(1.1753)
	3	0.9884	0.9798	0.9204	0.9970
		(1.0686)	(1.0410)	(0.7955)	(1.7992)
	4	0.9598	0.9166	0.7912	0.9902
		(0.9485)	(0.8821)	(0.6854)	(1.4459)
	5	0.9854	0.9800	0.9238	0.9938
		(1.3924)	(1.3241)	(1.0162)	(1.8392)
	6	0.9856	0.9776	0.9234	0.9918
		(1.3597)	(1.3186)	(1.0167)	(1.8355)
	7	0.9854	0.9754	0.9142	0.9924
		(1.3525)	(1.2847)	(0.9859)	(1.8308)
	8	0.9844	0.9760	0.9112	0.9942
		(1.3432)	(1.2909)	(0.9937)	(1.8327)

Table 3 Values selected for the population means, population standard deviations, probabilities of obtaining zero observation, and censoring points: 6 sample cases

Run number	$(\mu_1, \mu_2, \mu_3,$ $\mu_4, \mu_5, \mu_6)$	$(\sigma_1, \sigma_2, \sigma_3,$ $\sigma_4, \sigma_5, \sigma_6)$	$(\delta_1, \delta_2, \delta_3,$ $\delta_4, \delta_5, \delta_6)$	$(\xi_1, \xi_2, \xi_3,$ (ξ_4, ξ_5, ξ_6)
1	(0.00,0.00,0.00, 0.00,0.00,0.00)	(1.00,1.00,1.00, 1.00,1.00,1.00)	(0.10,0.10,0.10, 0.10,0.10,0.10)	(0.10,0.10,0.10, 0.10,0.10,0.10)
2	(0.00,0.00,0.00, 0.00,0.00,0.00)	(1.00,1.00,1.00, 1.00,1.00,1.00)	(0.10,0.10,0.10, 0.10,0.10,0.10)	(0.10,0.10,0.10, 0.25,0.25,0.25)
3	(0.00,0.00,0.00, 0.00,0.00,0.00)	(1.00,1.00,1.00, 1.00,1.00,1.00)	(0.10,0.10,0.10, 0.25,0.25,0.25)	(0.10,0.10,0.10, 0.10,0.10,0.10)
4	(0.00,0.00,0.00, 0.00,0.00,0.00)	(1.00,1.00,1.00, 1.00,1.00,1.00)	(0.10,0.10,0.10, 0.25,0.25,0.25)	(0.10,0.10,0.10, 0.25,0.25,0.25)
5	(0.00,0.00,0.00, 0.00,0.00,0.00)	(1.00,1.00,1.00, 2.00,2.00,2.00)	(0.10,0.10,0.10, 0.10,0.10,0.10)	(0.10,0.10,0.10, 0.10,0.10,0.10)
6	(0.00,0.00,0.00, 0.00,0.00,0.00)	(1.00,1.00,1.00, 2.00,2.00,2.00)	(0.10,0.10,0.10, 0.10,0.10,0.10)	(0.10,0.10,0.10, 0.25,0.25,0.25)
7	(0.00,0.00,0.00, 0.00,0.00,0.00)	(1.00,1.00,1.00, 2.00,2.00,2.00)	(0.10,0.10,0.10, 0.25,0.25,0.25)	(0.10,0.10,0.10, 0.10,0.10,0.10)
8	(0.00,0.00,0.00, 0.00,0.00,0.00)	(1.00,1.00,1.00, 2.00,2.00,2.00)	(0.10,0.10,0.10, 0.25,0.25,0.25)	(0.10,0.10,0.10, 0.25,0.25,0.25)

Table 4 The coverage probabilities (CPs) and average lengths (ALs) of 95% two-sided confidence intervals for the common mean of delta-lognormal distributions based on left-censored data: 6 sample cases

$(n_1, n_2, n_3, n_4, n_5, n_6)$	Run number	CP (AL)			
		CI_{GCI}	CI_{BS}	CI_{PB}	CI_{AM}
(20,20,20,20,20,20)	1	0.7976 (0.9965)	0.7246 (0.9693)	0.5510 (0.8297)	0.9720 (2.3012)
	2	0.7110 (0.9545)	0.6372 (0.8986)	0.4544 (0.7815)	0.9622 (2.0889)
	3	0.8430 (1.1522)	0.7600 (1.1457)	0.5692 (0.9152)	0.9796 (2.8571)
	4	0.6972 (1.0340)	0.5906 (0.9912)	0.3716 (0.8139)	0.9634 (2.4771)
	5	0.9340 (1.5244)	0.8920 (1.4398)	0.7460 (1.1665)	0.9746 (3.0086)
	6	0.9402 (1.4911)	0.9002 (1.4114)	0.7354 (1.1407)	0.9816 (2.9246)
	7	0.9282 (1.4923)	0.8802 (1.4203)	0.7232 (1.1642)	0.9738 (2.9674)
	8	0.9272 (1.4825)	0.8778 (1.4040)	0.7192 (1.1465)	0.9736 (2.9844)
(30,30,30,30,30,30)	1	0.8766 (0.8755)	0.8216 (0.8525)	0.6528 (0.7203)	0.9894 (1.9309)
	2	0.7744 (0.8390)	0.7042 (0.7798)	0.5282 (0.6749)	0.9792 (1.7279)
	3	0.9128 (1.0194)	0.8548 (1.0147)	0.7064 (0.8038)	0.9912 (2.2930)
	4	0.7568 (0.9129)	0.6548 (0.8722)	0.4448 (0.7129)	0.9828 (2.0178)
	5	0.9632 (1.3326)	0.9344 (1.2680)	0.8272 (1.0205)	0.9858 (2.2862)
	6	0.9644 (1.3065)	0.9400 (1.2524)	0.8286 (1.0074)	0.9872 (2.2772)
	7	0.9536 (1.2954)	0.9254 (1.2342)	0.7984 (0.9962)	0.9864 (2.2565)
	8	0.9492 (1.2871)	0.9192 (1.2263)	0.8016 (0.9949)	0.9832 (2.2553)
(20,20,20,30,30,30)	1	0.8584 (0.9399)	0.7912 (0.9162)	0.6256 (0.7803)	0.9838 (2.0672)
	2	0.7228 (0.8839)	0.6672 (0.8237)	0.4818 (0.7146)	0.9712 (1.7766)
	3	0.8788 (1.1062)	0.8016 (1.0902)	0.6388 (0.8814)	0.9850 (2.7627)
	4	0.7036 (0.9589)	0.5848 (0.9103)	0.3816 (0.7473)	0.9708 (2.1986)
	5	0.9460 (1.5886)	0.9062 (1.4922)	0.7650 (1.1810)	0.9698 (3.0023)
	6	0.9388 (1.5578)	0.9084 (1.4755)	0.7664 (1.1723)	0.9750 (3.0085)
	7	0.9378 (1.5465)	0.8902 (1.4493)	0.7274 (1.1545)	0.9762 (3.0043)
	8	0.9348 (1.5261)	0.8942 (1.4345)	0.7366 (1.1457)	0.9778 (3.0017)

Table 4 Continued

$(n_1, n_2, n_3, n_4, n_5, n_6)$	Run number	CP (AL)			
		CI_{GCI}	CI_{BS}	CI_{PB}	CI_{AM}
(50,50,50,50,50,50)	1	0.9488	0.9140	0.8018	0.9982
		(0.7171)	(0.6983)	(0.5784)	(1.5655)
	2	0.8602	0.7940	0.6362	0.9946
		(0.6899)	(0.6323)	(0.5442)	(1.3864)
	3	0.9628	0.9256	0.8360	0.9978
		(0.8382)	(0.8339)	(0.6526)	(1.7927)
	4	0.8450	0.7394	0.5420	0.9924
		(0.7560)	(0.7194)	(0.5800)	(1.5954)
	5	0.9832	0.9680	0.9002	0.9950
		(1.0853)	(1.0435)	(0.8339)	(1.7153)
	6	0.9772	0.9642	0.8982	0.9932
		(1.0623)	(1.0352)	(0.8283)	(1.7150)
	7	0.9750	0.9572	0.8742	0.9918
		(1.0500)	(1.0088)	(0.8067)	(1.7021)
	8	0.9714	0.9554	0.8786	0.9926
		(1.0404)	(1.0044)	(0.8075)	(1.7037)
(30,30,30,50,50,50)	1	0.9188	0.8732	0.7414	0.9962
		(0.7919)	(0.7713)	(0.6506)	(1.6687)
	2	0.7906	0.7288	0.5644	0.9890
		(0.7435)	(0.6812)	(0.5906)	(1.4154)
	3	0.9380	0.8912	0.7648	0.9930
		(0.9520)	(0.9388)	(0.7505)	(2.2184)
	4	0.7710	0.6474	0.4376	0.9832
		(0.8190)	(0.7691)	(0.6242)	(1.7447)
	5	0.9664	0.9408	0.8410	0.9850
		(1.3847)	(1.3154)	(1.0410)	(2.2483)
	6	0.9576	0.9390	0.8328	0.9860
		(1.3478)	(1.3062)	(1.0398)	(2.2754)
	7	0.9566	0.9226	0.8024	0.9844
		(1.3399)	(1.2702)	(1.0094)	(2.2578)
	8	0.9582	0.9334	0.8224	0.9882
		(1.3332)	(1.2744)	(1.0210)	(2.2864)
(100,100,100,100,100,100)	1	0.9858	0.9654	0.9142	1.0000
		(0.5190)	(0.5052)	(0.4113)	(1.1510)
	2	0.9178	0.8556	0.7340	0.9988
		(0.5000)	(0.4535)	(0.3873)	(1.0087)
	3	0.9928	0.9828	0.9386	0.9998
		(0.6120)	(0.6067)	(0.4697)	(1.2829)
	4	0.9202	0.8032	0.6282	0.9994
		(0.5508)	(0.5230)	(0.4150)	(1.1626)
	5	0.9908	0.9862	0.9490	0.9994
		(0.7771)	(0.7546)	(0.6026)	(1.1972)
	6	0.9870	0.9860	0.9472	0.9994
		(0.7637)	(0.7568)	(0.6048)	(1.2025)
	7	0.9846	0.9776	0.9352	0.9988
		(0.7596)	(0.7373)	(0.5888)	(1.1988)
	8	0.9876	0.9808	0.9386	0.9984
		(0.7553)	(0.7417)	(0.5932)	(1.2041)

Table 4 Continued

$(n_1, n_2, n_3, n_4, n_5, n_6)$	Run number	CP (AL)			
		CI_{GCI}	CI_{BS}	CI_{PB}	CI_{AM}
(50,50,50,100,100,100)	1	0.9736 (0.6029)	0.9520 (0.5869)	0.8774 (0.4851)	0.9994 (1.2516)
	2	0.8572 (0.5609)	0.7976 (0.5043)	0.6466 (0.4319)	0.9962 (1.0304)
	3	0.9828 (0.7579)	0.9614 (0.7470)	0.8996 (0.5857)	0.9990 (1.7060)
	4	0.8370 (0.6324)	0.6924 (0.5864)	0.4816 (0.4676)	0.9978 (1.2876)
	5	0.9784 (1.1461)	0.9638 (1.1044)	0.9024 (0.8783)	0.9936 (1.7129)
	6	0.9732 (1.1104)	0.9642 (1.1063)	0.9058 (0.8814)	0.9934 (1.7260)
	7	0.9726 (1.0874)	0.9536 (1.0463)	0.8672 (0.8334)	0.9926 (1.6972)
	8	0.9736 (1.0847)	0.9646 (1.0663)	0.8902 (0.8508)	0.9942 (1.7230)

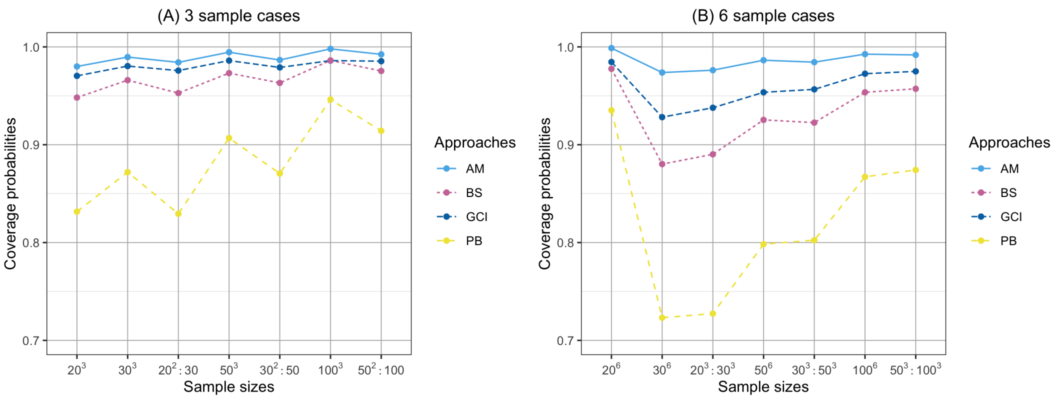

Figure 1 Comparison of the coverage probabilities of proposed approaches according to sample sizes

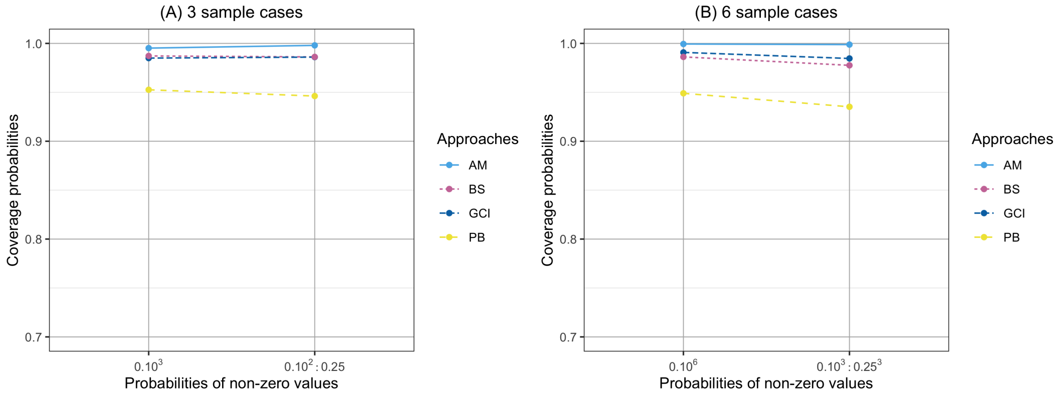

Figure 2 Comparison of the coverage probabilities of proposed approaches according to probabilities of non-zero values

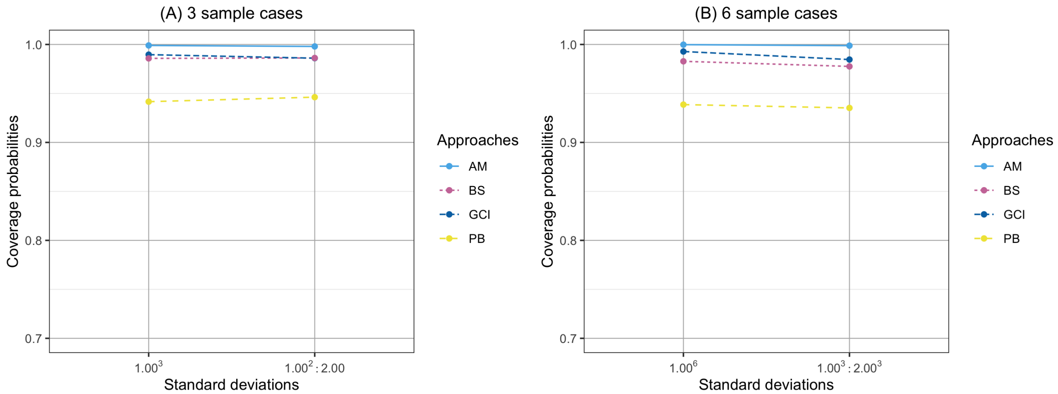

Figure 3 Comparison of the coverage probabilities of proposed approaches according to standard deviations

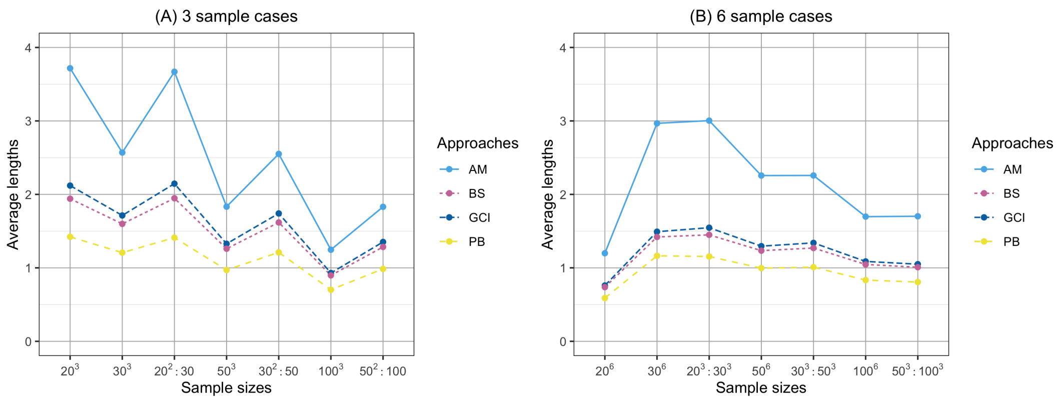

Figure 4 Comparison of the average lengths of proposed approaches according to sample sizes

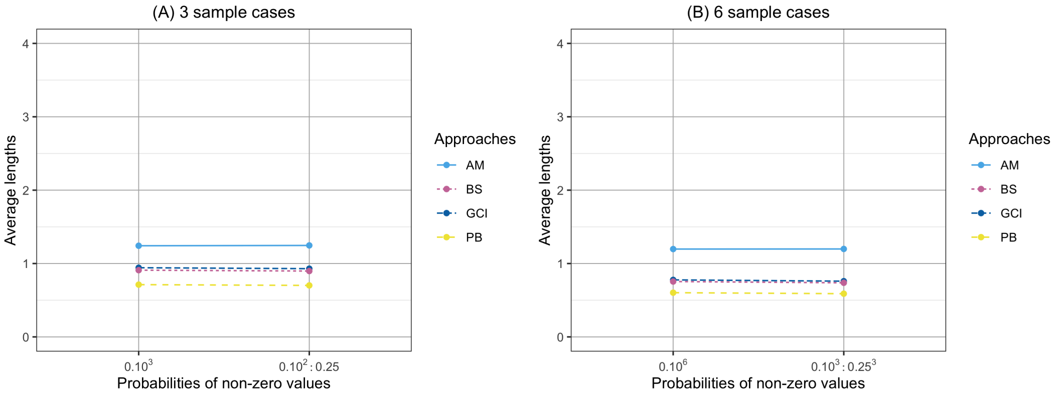

Figure 5 Comparison of the average lengths of proposed approaches according to probabilities of non-zero values

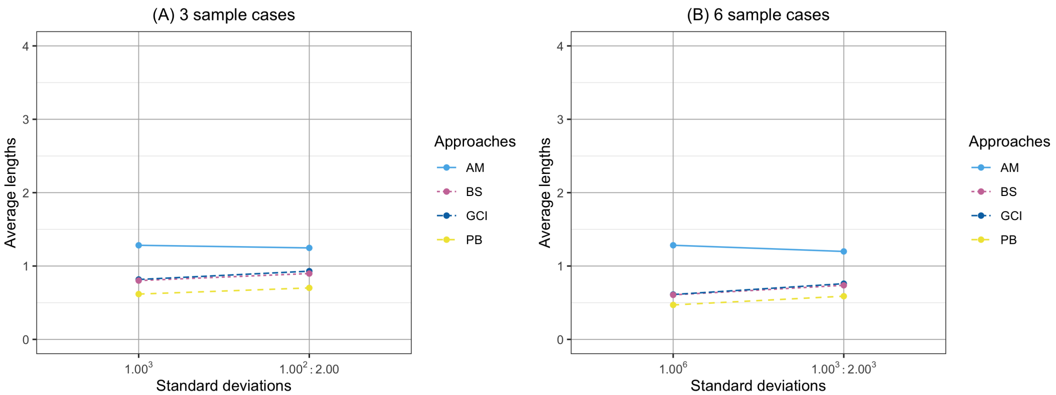

Figure 6 Comparison of the average lengths of proposed approaches according to standard deviations

4. Empirical Application

Real datasets are used to illustrate the efficacy of the proposed confidence intervals. A Monte Carlo simulation consisting of 2,500 repetitions was run to construct the GCI, Bayesian, and parametric bootstrap confidence intervals.

On September 1, 2021, the Thai Meteorological Department reported rainfall data for five regions in Thailand. Thangjai et al. (2022) found that the rainfall data for the northern, northeastern, and eastern regions of Thailand follow delta-lognormal distributions. The statistics of the rainfall data for the northern, northeastern, and eastern regions are presented in Table 5. The common mean of the rainfall data for the northern, northeastern, and eastern regions is $\hat{\theta} = 8.18$ mm.

The 95% two-sided confidence intervals for the common mean of delta-lognormal distributions based on left-censored data, using the rainfall data for the northern, northeastern, and eastern regions, are shown in Table 6. From the 95% confidence interval estimates, the length of the Bayesian credible interval is the shortest, but it does not contain the true common mean. Moreover, the length of the parametric bootstrap confidence interval is shorter than the lengths of the GCI and adjusted MOVER confidence intervals. However, the coverage probabilities of the parametric bootstrap confidence interval are less than the nominal confidence level of 0.95 in the simulation proposed in Section 3. Additionally, the length of the GCI is shorter than the length of the adjusted MOVER confidence interval. Therefore, for the number of sample cases $k = 3$, the GCI approach is recommended to construct the common mean of delta-lognormal distributions based on left-censored data when the sample sizes are small.

Furthermore, it is important to mention that confidence intervals for the common mean of delta-lognormal distributions, obtained from data with left-censoring, can be applied in the realm of environmental, meteorological, and climatological data. These datasets often contain positive values and exhibit right-skewed distributions, as observed in variables like particulate matter 2.5 (PM2.5) and particulate matter 10 (PM10).

Table 5 Sample statistics of the rainfall data for the northern, northeastern, and eastern regions

Statistics	Northern	Northeastern	Eastern
n_i	29	28	15
$n_{i(1)}$	6	10	1
$n_{i(2)}$	23	18	14
$\hat{\mu}_i$	-0.23	-0.66	1.05
$\hat{\sigma}_i^2$	4.44	9.09	4.88
$\hat{\theta}_i$	7.34	48.74	32.67

Table 6 The 95% two-sided confidence intervals for the common mean of delta-lognormal distributions based on left-censored data using the rainfall data for three regions

Approaches	Confidence intervals		
	Lower	Upper	Length
GCI	2.7476	39.8908	37.1432
Bayesian	0.5644	0.6773	0.1129
Parametric bootstrap	0.3057	17.0749	16.7692
Adjusted MOVER	2.0329	80.0942	78.0613

5. Discussion

Thangjai and Niwitpong (2023a) estimated confidence intervals for the mean of a delta-lognormal distribution based on left-censored data using the GCI, Bayesian, and parametric bootstrap approaches. They also constructed confidence intervals for the difference between means of delta-lognormal distributions based on left-censored data using the GCI, Bayesian, parametric bootstrap, and MOVER approaches. In a separate study, Thangjai and Niwitpong (2023b) proposed confidence intervals for the ratio of means of delta-lognormal distributions based on left-censored data using the GCI, Bayesian, parametric bootstrap, and adjusted MOVER approaches. In this paper, we utilized the concepts introduced by Thangjai and Niwitpong (2023a) and Thangjai and Niwitpong (2023b) to construct confidence intervals for the common mean of delta-lognormal distributions based on left-censored data using the GCI, Bayesian, parametric bootstrap, and adjusted MOVER approaches.

The GCI, Bayesian, and adjusted MOVER approaches are recommended for constructing confidence intervals for the common mean of delta-lognormal distributions based on left-censored data. The results are consistent with the findings from previous studies conducted by Thangjai and Niwitpong (2017), Thangjai et al. (2017c), Thangjai and Niwitpong (2018), Thangjai et al. (2021), Thangjai and Niwitpong (2022), and Thangjai et al. (2022).

6. Conclusion

The confidence intervals for the common mean of delta-lognormal distributions based on left-censored data were constructed using the GCI, Bayesian, parametric bootstrap, and adjusted MOVER approaches. For the number of sample cases $k = 3$, the GCI and adjusted MOVER approaches were recommended for constructing the confidence intervals when the sample sizes were small, while the Bayesian approach was suggested for constructing credible interval when the sample sizes were moderate and large. For the number of sample cases $k = 6$, the GCI and adjusted MOVER approaches were recommended for constructing the confidence interval when the sample sizes were small and moderate, while the Bayesian approach were recommended for constructing the confidence interval when the sample sizes were large.

Acknowledgements

This work (Grant No. RGNS 65-178) was supported by Office of the Permanent Secretary, Ministry of Higher Education, Science, Research and Innovation (OPS MHESI), Thailand Science Research and Innovation (TSRI) and Ramkhamhaeng University.

Funding

This work (Grant No. RGNS 65-178) was financially supported by Office of the Permanent Secretary, Ministry of Higher Education, Science, Research and Innovation.

References

Altunkaynak B, Gamgam H. Bootstrap confidence intervals for the coefficient of quartile variation. Commun Stat - Simul C. 2019; 48: 2138-2146.

Chen YH, Zhou XH. Generalized confidence intervals for the ratio or difference of two means for lognormal populations with zeros. UW Biostatistics Working Paper Series. 2006: 1-16.

Glass DC, Gray CN. Estimating mean exposures from censored data: exposure to benzene in the Australian petroleum industry. Ann Occup Hyg. 2001: 45; 275-282.

Krishnamoorthy K, Lu Y. Inference on the common means of several normal populations based on the generalized variable method. Biometrics. 2003: 59; 237-247.

Krishnamoorthy K, Mallick A, Mathew T. Inference for the lognormal mean and quantiles based on samples with left and right Type I censoring. Technometrics. 2011: 53; 72-83.

Lin SH, Lee JC. Generalized inferences on the common mean of several normal populations. J Stat Plan Infer. 2005: 134; 568-582.

Ma Z, Chen G. Bayesian methods for dealing with missing data problems. J Korean Stat Soc. 2018: 47; 297-313.

Maneerat P, Niwitpong S-A, Niwitpong S. Statistical estimation of mean of delta-lognormal distribution. Thail Stat. 2020: 18; 439-456.

Maneerat P, Niwitpong S-A, Niwitpong S. Bayesian confidence intervals for a single mean and the difference between two means of delta-lognormal distributions. Commun Stat - Simul C. 2021; 50: 2906-2934.

Ng CK. Inference on the common coefficient of variation when populations are lognormal: A simulation-based approach. J Stat Adv Theory Appl. 2014; 11: 117-134.

Owen WJ, DeRouen TA. Estimation of the mean for lognormal data containing zeroes and left-censored values, with applications to the measurement of worker exposure to air contaminants. Biometrics. 1980; 36: 707-719.

Padgett WJ, Tomlinson, M.A. Lower confidence bounds for percentiles of Weibull and Birnbaum-Saunders distributions. J Stat Comput Simul. 2003; 73: 429-443.

Rao KA, DCunha JG. Bayesian inference for median of the lognormal distribution. J Mod App Stat Meth. 2016; 15: 526-535.

Tian L. Inferences on the common coefficient of variation. Stat Med. 2005; 24: 2213-2220.

Tian L, Wu J. Inferences on the common mean of several log-normal populations: the generalized variable approach. Biometrical J. 2007; 49: 944 - 951.

Thangjai W, Niwitpong, S-A. Confidence intervals for the weighted coefficients of variation of two-parameter exponential distributions. Cogent Math. 2017; 4: 1-16.

Thangjai W, Niwitpong S. Confidence intervals for common variance of several one-parameter exponential populations. Adv Appl Stat. 2018, 53, 285-311.

Thangjai W, Niwitpong S-A. Confidence intervals for common signal-to-noise ratio of several lognormal distributions. Iran J Sci Technol Trans A: Sci. 2020; 44: 99-107.

Thangjai W, Niwitpong S-A. The relative potency of two drugs using the confidence interval for ratio of means of two normal populations with unknown coefficients of variation. J Stat Appl Probab. 2022; 11: 1-14.

Thangjai W, Niwitpong S-A. Confidence intervals for mean and difference between means of delta-lognormal distributions based on left-censored data. Symmetry. 2023a; 15(1216): 1-24.

Thangjai W, Niwitpong S-A. Confidence intervals for ratio of means of delta-lognormal distributions based on left-censored data with application to rainfall data in Thailand. PeerJ. 2023b; in press.

Thangjai W, Niwitpong S-A., Niwitpong S. Confidence intervals for the common mean of several normal populations. Robustness in Econometrics. 2017a; 692: 321-331.

Thangjai W, Niwitpong S-A., Niwitpong S. Confidence intervals for the common mean of several one-parameter exponential populations. Adv Appl Stat. 2017b; 51: 245-260.

Thangjai W, Niwitpong S., Niwitpong, S-A. On large sample confidence intervals for the common inverse mean of several normal populations. Adv Appl Stat. 2017c; 51: 59-84.

Thangjai W, Niwitpong S-A, Niwitpong S. Confidence intervals for variance and difference between variances of one-parameter exponential distributions. Adv Appl Stat. 2018, 53, 259-283.

Thangjai W, Niwitpong S-A, Niwitpong S. Adjusted generalized confidence intervals for the common coefficient of variation of several normal populations. Commun Stat - Simul C. 2020a; 49: 194-206.

Thangjai W, Niwitpong S-A, Niwitpong S. Confidence intervals for the common coefficient of variation of rainfall in Thailand. PeerJ. 2020b; 8: e10004.

Thangjai W, Niwitpong S-A, Niwitpong S. A Bayesian approach for estimation of coefficients of variation of normal distributions. Sains Malaysiana. 2021; 50: 261-278.

Thangjai W, Niwitpong S-A, Niwitpong S. Estimation of common percentile of rainfall datasets in Thailand using delta-lognormal distributions. PeerJ. 2022; 10: 1-39.

Thangjai W, Niwitpong S-A, Niwitpong S, Smithpreecha N. Confidence interval estimation for the ratio of the percentiles of two delta-lognormal distributions with application to rainfall data. Symmetry. 2023; 15(794): 1-13.

Ye RD, Ma TF, Wang SG. Inferences on the common mean of several inverse Gaussian populations. Comput Stat Data Anal. 2010; 54: 906-915.

Yosboonruang N, Niwitpong S-A, Niwitpong S. Bayesian computation for the common coefficient of variation of delta-lognormal distributions with application to common rainfall dispersion in Thailand. PeerJ. 2022; 10: 1-24.

Zhang G. A parametric bootstrap approach for one-way ANOVA under unequal variances with unbalanced data. Commun Stat - Simul C. 2015; 44: 827-832.