GU J Sci 30(3): 209-222 (2017)

Gazi University

Journal of Science

http://dergipark.gov.tr/gujs

Comparison of Tests for the Equality of Several Log-Normal Means

Esra GÖKPINAR[1,*]

[1] *Gazi University, Department of Statistics, Ankara, Turkey*

Article Info

Received: 19/04/2017
Accepted: 12/06/2017

Keywords

Log-normal means
Likelihood-based methods
Generalized pivotal-based methods
Bootstrap-based methods

Abstract

In this paper, we focused on testing for the equality of several log-normal means since the log-normal distribution is one of the most common distributions for analyzing positive and right-skewed data. Recently, many researchers have proposed a lot of methods based on likelihood-based methods, generalized pivotal-based methods, and bootstrap-based methods for this case. Apparently, since there is not an exact result regarding which test is better than the others in which cases, our goal shed light on this important issue. For this reason, we investigate these methods and compare them with each other by a simulation study.

1. INTRODUCTION

The log-normal distribution is one of the most common distributions that is used to model positive and right-skewed data. It has many practical applications such as economic data, data on response of biological material to stimuli, and certain types of life data [3]. A very common problem in applied statistics is to compare the means of several log-normal distributions. By using log-transformation version of data, the classical ANOVA F-test can be used for this problem. However, the classical ANOVA F-test has high type I error rate when the variances of groups are not equal. In the literature, there are many tests dealing with this issue [23, 24, 11, 21, 10].

Traditional tests, including the classical likelihood ratio test (LRT), are asymptotic in nature, and hence do not perform well for small sample sizes. Therefore, in recent years, methods called as likelihood-based methods, generalized pivotal (GV)-based methods, bootstrap-based methods have been proposed by researchers for testing the equality of means of several log-normal distributions. Gill [4] proposed a method introduced by Skovgaard [20] , which is a correction to the LRT for small-sample situations. This method is referred to as the modified likelihood ratio test (MLRT). A limited simulation study was carried out for comparing this test with F-ratio test, only in cases where k=5 and n=5. Besides, Krishnamoorthy and Oral [14] proposed a method to improve the LRT, which is called the standardized likelihood ratio test (SLRT), for testing the equality of means of several log-normal distributions. They compared it with the MLRT and the generalized variable test (GVT). Simulation studies showed that the SLRT appears to be the best in terms of type I error rates and powers. They also noted that the MLRT is not appropriate for applications as it is not defined for some samples.

Furthermore, there are alternative modifications based on likelihood ratio method in the literature. For example, Wu et al. [22] proposed two methods based on likelihood approach, which are called the signed likelihood ratio and modified signed likelihood ratio methods, to construct a confidence interval for the mean of a log normal distribution. Lin [16] proposed higher order likelihood method known as the modified signed likelihood ratio method to construct a confidence interval for the common mean of several log normal distributions. However, these modified versions are very difficult to extend to compare two or more log-normal means.

*Corresponding author, e-mail: eyigit@gazi.edu.tr

In the literature, methods based on GV are suggested for testing the equality of means of several log-normal distributions. Krishnamoorthy and Mathew [13] suggested the GV approach for testing the equality of two log-normal means. Li [17] generalized the GV approach of Krishnamoorthy and Mathew [13] by adopting a quadratic procedure for testing the equality of several log-normal means. Li's simulation study showed that when comparing this test with Welch test, it is better than the other. However, Lin and Wang [15] noted that the quadratic procedure cannot be suitable for asymmetric distributions, thus they proposed a modification of the quadratic method for testing the equality of means of several log-normal distributions. They also showed that Li's [17] results are acceptable only under some combinations of the parameters of log-normal distribution.

In the literature, another proposed approach is bootstrap approach (PB) for testing the equality of means of several log-normal distributions. The PB is a type of Monte Carlo method applied on observed data. A computational approach test (CAT), which is a type of parametric bootstrap method, was firstly proposed by Pal et al. [19]. The CAT method based on simulation and numerical computations uses the maximum likelihood estimates (MLEs), and does not require the knowledge of any sampling distribution. Some papers related to the CAT are given as, Chang et al. [1, 2], Gökpınar and Gökpınar [6,8,9], Gökpınar et al. [7], Mutlu et al. [18], etc.. Gökpınar and Gökpınar [5] proposed a test based on CAT for testing the equality of several log-normal means. They compared it with Welch test and GV approach by Krishnamoorthy and Methew [13] when the number of groups was 2 (k=2). Their numerical results show that the proposed test is better than the others in most of the considered cases. Jafari and Abdollahnezhad [12] proposed three tests based on CAT for testing the equality of several log-normal means. Two of these tests are CAT modifications of the LRT and MLRT. They compared them with the other tests-the LRT, the MLRT, GV approach by Li [17] and GV approach by Krishnamoorthy and Methew [13] when k=2- only in terms of the type I error rates. Their simulation study showed that the type I error rates of CAT modifications of the LRT and MLRT are closer to the nominal level than the others.

As mentioned above, in general, for testing the equality of means of several log-normal distributions, classified methods in the literature are three folds: likelihood-based methods, generalized pivotal-based methods and bootstrap-based methods. For this reason, in this study we investigated these methods and compared them with each other by a comprehensive simulation study.

The rest of this study was organized as follows. In Section 2, we defined the null and alternative hypotheses of interest, and presented the likelihood-based methods-LRT, MLRT, SLRT-, generalized pivotal-based methods-GV approach by Krishnamoorthy and Mathew [13] when k=2, GV approach by Li [17]-,bootstrap-based methods-CAT approach by Gökpınar and Gökpınar [5], CAT approaches by Jafari and Abdollahnezhad [12]-. In Section 3, we conducted simulation study to assess the type I error rates and powers of these tests for different cases. Concluding remarks were summarized in Section 4.

2. TEST STATISTICS

Let $Y_{ij},\ i=1,\cdots,k,\quad j=1,\cdots,n_i$, be a random sample from population with log-normal$\left(\mu_i,\sigma_i^2\right)$ distribution. The mean of the *i*th population is given as $M_i=\exp(\theta_i)$, where $\theta_i=\mu_i+\sigma_i^2/2$. It is well known that $X_{ij}=\ln(Y_{ij})$ is distributed as normal distribution with mean μ_i and variance σ_i^2. The unbiased estimators of μ_i and σ_i^2 are defined as

$$\bar{X}_i=\sum\nolimits_{j=1}^{n_i}X_{ij}\Big/n_i\,,\qquad S_i^2=\sum\nolimits_{j=1}^{n_i}\left(X_{ij}-\bar{X}_i\right)^2\Big/\left(n_i-1\right). \tag{1}$$

We want to test H_0 against H_1 given in Eq. (2):

$$H_0: M_1=M_2=\cdots=M_k,\qquad H_1: M_i\neq M_{i'},\quad \exists i\neq i'\left(i,i'=1,\cdots,k\right). \tag{2}$$

It is clear that testing H_0 in Eq. (2) is equivalent to testing $H_0^{(1)}$ in Eq. (3) given as

$$H_0^{(1)}:\theta_1=\theta_2=\cdots=\theta_k=\theta,\qquad H_1^{(1)}:\theta_i\neq\theta_{i'},\quad \exists i\neq i'\left(i,i'=1,\cdots,k\right). \tag{3}$$

In the following sections, for testing H_0 in Eq. (2) (or $H_0^{(1)}$ in Eq. (3)), likelihood-based methods, generalized pivotal-based methods and bootstrap-based methods were presented.

2.1. Likelihood-Based Tests

2.1.1. The likelihood ratio test

The log-likelihood function is

$$l=-0.5\sum_{i=1}^{k}n_i\log\left(\sigma_i^2\right)-0.5\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{\left(x_{ij}-\mu_i\right)^2}{\sigma_i^2}.$$

The maximum likelihood (ML) estimates of μ_i and σ_i^2 are denoted as

$$\hat{\mu}_i=\bar{X}_i=\sum\nolimits_{j=1}^{n_i}X_{ij}\Big/n_i\qquad\qquad \hat{\sigma}_i^2=\sum\nolimits_{j=1}^{n_i}\left(X_{ij}-\bar{X}_i\right)^2\Big/n_i\,. \tag{4}$$

Under the null hypothesis the log-likelihood function, that is, the restricted log-likelihood function is

$$l_0=-0.5\sum_{i=1}^{k}n_i\log\left(\sigma_i^2\right)-0.5\sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{\left(x_{ij}-\theta+0.5\sigma_i^2\right)^2}{\sigma_i^2}.$$

The restricted maximum likelihood (RML) estimates of μ_i and σ_i^2 are given as

$$\tilde{\theta}=\frac{\left(0.5+\sum_{i=1}^{k}\frac{n_i\bar{x}_i}{\tilde{\sigma}_i^2}\right)}{\sum_{i=1}^{k}\frac{n_i}{\tilde{\sigma}_i^2}}\ ,\qquad \tilde{\sigma}_i^2=-2+2\left(1+\frac{1}{n_i}\sum_{i=1}^{k}\left(x_{ij}-\tilde{\theta}\right)^2\right)^{1/2} \tag{5}$$

As seen from equations, the RMLEs of the θ and σ_i^2 have no closed forms. The RMLEs of θ and σ_i^2 (i=1,2,...,k) can be solved iteratively. Jafari and Abdollahnezhad [12] used an initial value

$\theta^{(0)} = \sum_{i}^{k} n_i \bar{x}_i \Big/ \sum_{i}^{k} n_i$, Gökpınar and Gökpınar [5] used an initial value $\theta^{(0)} = \dfrac{\sum_{i=1}^{k} n_i \left(\bar{X}_i + S_i^2/2\right)\big/ S_i^2}{\sum_{i=1}^{k} n_i / S_i^2}$.

The likelihood ratio test (LRT) is

$$\Lambda = 2(l - l_0) = \sum_{i}^{k} n_i \log\left(\frac{n_i \tilde{\sigma}_i^2}{(n_i - 1) s_i^2}\right) + \sum_{i}^{k} \sum_{j}^{n_i} \frac{\left(x_{ij} - \tilde{\theta} + 0.5\tilde{\sigma}_i^2\right)}{\tilde{\sigma}_i^2} - n \tag{6}$$

where $n = \sum_{i}^{k} n_i$. Λ has an approximate chi-square distribution with k-1 degrees of freedom under the null hypothesis.

2.1.2. The Modified Likelihood Ratio Test

Gill [4] proposed a correction given by Skovgaard [20] to the likelihood ratio test statistic which leads to more accurate inference in small-sample situations. The modified likelihood ratio test (MLRT) is given by

$$\Lambda^* = \Lambda\left(1 - \Lambda^{-1} \log \gamma\right)^2. \tag{7}$$

Here γ is defined as

$$\gamma = \frac{\left\{\left(t - \tilde{\tau}\right)' \tilde{\Sigma}^{-1} \left(t - \tilde{\tau}\right)\right\}^{k/2} \left|\tilde{\Sigma}\right|^{1/2}}{\Lambda^{k/2-1} \left(\hat{\beta} - \tilde{\beta}\right)' \left(t - \tilde{\tau}\right)' \left|\hat{\Sigma}\right|^{1/2}}. \tag{8}$$

where $\beta = \left(\frac{\mu_1}{\sigma_1^2}, ..., \frac{\mu_k}{\sigma_k^2}, \frac{1}{\sigma_1^2}, ..., \frac{1}{\sigma_k^2}\right)'$ is the vector of canonical parameters and $t = \left(\bar{x}_1, ..., \bar{x}_k, -0.5\sum_{j=1}^{n_1} x_{1j}^2, ..., -0.5\sum_{j=1}^{n_k} x_{kj}^2\right)'$ is the vector of canonical sufficient statistics. Further, $\tau = E(t) = \left[\mu_1, ..., \mu_k, -0.5n_1\left(\mu_1^2 + \sigma_1^2\right), ..., -0.5n_k\left(\mu_k^2 + \sigma_k^2\right)\right]'$ and

$$\Sigma = V(t) = \begin{bmatrix} diag\left(\frac{\sigma_i^2}{n_i}\right)_{i=1,..k} & diag\left(-\frac{\mu_i \sigma_i^2}{n_i}\right)_{i=1,..k} \\ diag\left(-\frac{\mu_i \sigma_i^2}{n_i}\right)_{i=1,..k} & diag\left(n_i \sigma_i^2 \left(\mu_i^2 + 0.5\sigma_i^2\right)\right)_{i=1,..k} \end{bmatrix}$$

Let $\tilde{\tau} = \left[\left(\tilde{\theta} - 0.5\tilde{\sigma}_1^2\right), ..., \left(\tilde{\theta} - 0.5\tilde{\sigma}_k^2\right), -0.5n_1\left(\tilde{\mu}_1^2 + \tilde{\sigma}_1^2\right), ..., -0.5n_k\left(\tilde{\mu}_k^2 + \tilde{\sigma}_k^2\right)\right]'$ be an estimate of τ under the null hypothesis. Similarly, $\tilde{\beta}$ and $\tilde{\Sigma}$ are estimates of β and Σ under the null hypothesis. $\hat{\beta}$ and $\hat{\Sigma}$ are ML estimates of β and Σ. Λ^* has an approximate chi-square distribution with *k*-1 degrees of freedom under the null hypothesis.

Remark 1. Krishnamoorthy and Oral [14] noted that, as the quantity γ in Eq. (8) could be negative for some samples, the MLRT is not appropriate for these cases. We also observed this case in the simulation study, then we ignored this test in the simulation study.

2.1.3. The standardized likelihood ratio test

Krishnamoorthy and Oral [14] proposed the standardize likelihood ratio test (SLRT) for testing the equality of means of several log-normal distributions. The SLRT is defined as

$$\Lambda_S = \sqrt{2(\mathrm{k}-1)}\left(\frac{\Lambda - m(\Lambda)}{SD(\Lambda)}\right) + (k-1), \tag{9}$$

where $m(\Lambda)$ and $SD(\Lambda)$ are the mean and standard deviation of Λ, respectively. Λ_S has an approximate chi-square distribution with *k*-1 degrees of freedom under the null hypothesis. Krishnamoorthy and Oral [14] estimated the expressions of $m(\Lambda)$ ve $SD(\Lambda)$ through simulation because they are difficult to obtain. The SLRT can be computed through the following steps.

For a given data $(x_{i1}, x_{i2}, \cdots, x_{in_i})$, calculate $\bar{x}_i$ and s_i^2.

1. Calculate the LRT statistic in Eq. (6)
2. Generate an artificial sample $\mathrm{X_{ij}} \sim \mathrm{N}\left(\tilde{\theta} - 0.5\tilde{\sigma}_i^2, \tilde{\sigma}_i^2\right)$, $\mathrm{j} = 1,\dots,\mathrm{n_i}$, $\mathrm{i} = 1,\dots,\mathrm{k}$ where $\tilde{\theta}$ and $\tilde{\sigma}_i^2$ in Eq. (5).
3. Calculate the LRT statistic in Eq. (6) for this replicated sample.
4. Repeat the steps 2 and 3 for a large number of times.
5. Calculate the mean and standard deviation of these simulated LRT, and find the SLRT statistic Λ_S in Eq. (9).
6. If $\Lambda_S > \chi^2_{k-1;1-\alpha}$, then H_0 is rejected.

2.2. The Generalized Variable-based Tests

2.2.1. Krishnamoorthy and Mathew's Test

Krishnamoorthy and Mathew [13] suggested the GV approach for testing the equality of two log-normal means. The GV is defined as:

$$\mathrm{T} = \mathrm{T}_{\theta_1} - \mathrm{T}_{\theta_2},$$

where

$$\mathrm{T}_{\theta_i} = \bar{x}_i - \frac{\mathrm{Z_i}}{\mathrm{U_i}}\sqrt{s_i^2(\mathrm{n_i}-1)/\mathrm{n_i}} + \frac{s_i^2(\mathrm{n_i}-1)}{2\mathrm{U}_i^2}, \qquad \mathrm{i} = 1,2. \tag{10}$$

Here $Z_i = \sqrt{n_i}\left(\bar{X}_i - \mu_i\right)/\sigma_i \sim N(0,1)$ and $U_i^2 = (n_i - 1)S_i^2/\sigma_i^2 \sim \chi^2_{n_i-1}$, $i = 1,2$. The GV method can be computed through the following steps.

For a given data $\left(x_{i1}, x_{i2}, \cdots, x_{in_i}\right)$, calculate $\bar{x}_i$ and s_i^2.

1. Generate $Z_{\mathrm{i}} \sim N(0,1)$ and $U_{\mathrm{i}}^2 \sim \chi^2_{n_{\mathrm{i}}-1}$, $i=1,2$.

2. Compute $T_{\theta_i} = \bar{x}_i - Z_i s_i \sqrt{n_i - 1} \Big/ \left(U_i \sqrt{n_i}\right) + s_i^2 \left(n_i - 1\right) \big/ 2U_i^2$, $i=1,2$ in Eq.(10).

3. Compute $\mathrm{T} = \mathrm{T}_{\theta_1} - \mathrm{T}_{\theta_2}$.

4. Let A_l =1 if T≥0.

5. Repeat step 1-4 a total L times. Let $p^* = \sum_{l=1}^{L} A_l \big/ L$.

6. Calculate the p-value as p= 2×min(p*, 1- p*). In the case of $p<\alpha$, H_0 is rejected.

2.2.2. Li's Test

Li [17] proposed the GV approach for testing the equality of several log-normal means by using the GV suggested by Krishnamoorthy and Mathew [13]. The GV method can be computed through the following steps.

1. Generate $Z_{\mathrm{i}} \sim N(0,1)$ and $U_{\mathrm{i}}^2 \sim \chi^2_{n_{\mathrm{i}}-1}$, $i=1,\ldots,\mathrm{k}$.

2. Compute $D_\theta = \left(T_{\theta_1} - T_{\theta_k}, \ldots, T_{\theta_{k-1}} - T_{\theta_k}\right)$.

3. Repeat the steps 1 and 2 for a large number of times, say, M.

4. Calculate the mean and covariance of these values of D_θ, denote them by $\hat{\mu}_T$ and $\hat{\Sigma}_T$, respectively.

5. Calculate $\hat{\mu}_T' \hat{\Sigma}_T^{-1} \hat{\mu}_T$.

6. Calculate $Q^* = \left(D_\theta - \hat{\mu}_T\right)' \hat{\Sigma}_T^{-1} \left(D_\theta - \hat{\mu}_T\right)$, for each of M values of D_θ.

7. Repeat step 1-6 a large number of times, say, L. Calculate the estimate of the generalized p value, that is, $\hat{p} = \#\left(Q^* > \hat{\mu}_T' \hat{\Sigma}_T^{-1} \hat{\mu}_T\right) \Big/ L$.

We refer to this test as L-GV in the simulation study.

2.2.3. Lin and Wang's Test

Lin and Wang [15] proposed a modified method based on GV approach for testing the equality of several log-normal means. $\mathrm{H}_0^{(1)}$ in Eq. (3) can be expressed as follows:

$\mathrm{H}_0^{(2)} : \mathbf{C}_0^* \boldsymbol{\theta} = \boldsymbol{\theta}_0$,

where $\boldsymbol{\theta} = \left(\theta_1, \theta_2, \ldots, \theta_k\right)'$, $\boldsymbol{\theta}_0 = \mathbf{0}_{\mathrm{k}-1}$ and

$$\mathbf{C}_0^* = \begin{bmatrix} 1 & -1 & 0 & \ldots & 0 \\ 1 & 0 & -1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 & 0 & 0 & \ldots & -1 \end{bmatrix}$$

The GV given in Eq. (10) can be written as:

$$T_{\theta_i} = \bar{x}_i - \frac{s_i}{\sqrt{n_i}} t_i + \frac{s_i^2 (n_i - 1)}{2U_i^2}, \qquad i = 1, \ldots, k. \tag{11}$$

The proposed modified method based on GV approach can be computed through the following steps.

For a given data $\left(x_{i1}, x_{i2}, \cdots, x_{in_i}\right)$, calculate $\bar{x}_i$ and s_i^2.

1. Generate $t_i \sim t_{n_i - 1}$ and $U_i^2 \sim \chi^2_{n_i - 1}$, $i = 1, \ldots, k$.

2. Compute T_{θ_i}, i=1,2...,k.

3. Compute $\mathbf{R}$=$\mathbf{C}^*\mathbf{G}$, where $\mathbf{C}^* = \mathbf{C}_0^*$ and $\mathbf{G}$=$\left(T_{\theta_1}, T_{\theta_2}, \ldots, T_{\theta_k}\right)'$.

4. Repeat step 1-3 m times and obtain an array of $\mathbf{R}$s.

5. Calculate the mean and covariance matrix of $\mathbf{R}$, denote them by $\mathbf{m_R}$ and $\mathbf{S}_\mathbf{R}^*$, respectively. Obtain $\tilde{R}_q = \left(\mathbf{S}_\mathbf{R}^*\right)^{-1/2}\left(\mathbf{R}_q - \mathbf{m_R}\right)$ for q=1,2,…,m.

6. Compute $d_q = \left(\sum_{p=1}^{m} \left\| \tilde{\mathbf{R}}_p - \tilde{\mathbf{R}}_q \right\| \right) \Big/ m, \; q = 1, \ldots, m.$

7. Let $d_0 = \sum_{p=1}^{m} \left\| \tilde{\mathbf{R}}_p - \tilde{\mathbf{R}}_0 \right\|$ with $\tilde{\mathbf{R}}_0 = \left(\mathbf{S}_\mathbf{R}^*\right)^{-1/2}\left(\boldsymbol{\theta}_0 - \mathbf{m_R}\right)$, then calculate Monte Carlo estimate of the generalized p-value as $\hat{p} = \#\left(d_0 \le d_q\right) \big/ m$, for q=1,…m.

We refer to this test as LW-GV in the simulation study.

2.3. Parametric Bootstrap-based Tests

2.3.1. Gökpınar and Gökpınar's CAT Test

Gökpınar and Gökpınar [5] proposed a test based on CAT which is a type of parametric bootstrap method. $H_0^{(1)}$ in Eq. (3) is expressed in terms of suitable scalar η as follows:

$$H_0^{(3)} : \eta = \eta\left(\theta_1, \cdots, \theta_k\right) = \sum_{i=1}^{k} \frac{n_i \left(\theta_i - \bar{\theta}\right)^2}{\sigma_i^2} = 0, \tag{12}$$

where $\bar{\theta} = \sum_{i=1}^{k} n_i \theta_i \Big/ \sum_{i=1}^{k} n_i$. Using the MLEs of μ_i and σ_i^2 (i=1,2,...,k), we can obtain the estimator of η which can be used as a test statistic for testing $H_0^{(3)}$. The CAT method can be computed through the following steps.

1. Calculate the MLE of η, that is, $\hat{\eta} = \sum_{i=1}^{k} n_i \left(\hat{\theta}_i - \hat{\bar{\theta}}\right)^2 \Big/ \hat{\sigma}_i^2$, where $\hat{\theta}_i = \bar{X}_i + S_i^2 / 2$ and $\hat{\bar{\theta}} = \sum_{i=1}^{k} n_i \hat{\theta}_i \Big/ \sum_{i=1}^{k} n_i$,

2. Assume that $\mathrm{H}_0^{(1)}$ is true, find the restricted MLEs of $\left(\sigma^2{}_1,\cdots,\sigma^2{}_{1k},\theta\right)$ in Eq. (5), and denoted as $\left(\tilde{\sigma}_1^2,\cdots,\tilde{\sigma}_k^2,\tilde{\theta}\right)$.

3. Generate artificial sample $X_{ij} \sim \mathrm{N}\left(\tilde{\theta}\text{-}\tilde{\sigma}_i^2/2,\, \tilde{\sigma}_i^2\right),\quad j=1,...,\mathrm{n_i},\, i=1,...,\mathrm{k}$ a large of number of times (say, L times). For each of these replicated samples, recalculate the values of $\hat{\eta}^{(l)},\quad l=1,...,L.$

4. Calculate the p-value as $p=\#\left(\hat{\eta}^{(l)}>\hat{\eta}\right)/L$. In the case of p<α, H_0 is rejected.

We refer to this test as CAT1 in the simulation study.

2.3.2. Jafari and Abdollahnezhad's CAT Tests

Jafari and Abdollahnezhad [12] proposed some tests based on CAT for testing the equality of several log-normal means. $\mathrm{H}_0^{(1)}$ in Eq. (3) can be expressed as follows:

$\mathrm{H}_0^{(4)}: \mathrm{C}\boldsymbol{\theta}=\mathbf{0},$

where $\boldsymbol{\theta}=\left(\theta_1,\theta_2,...,\theta_k\right)'$ and $\mathrm{C}=\begin{bmatrix} 1 & -1 & 0 & \dots & 0 \\ 1 & 0 & -1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 1 & 0 & 0 & \dots & -1 \end{bmatrix}$

The test statistic is given as:

$$T^*=\left(C\hat{\boldsymbol{\theta}}\right)'\left[CVC'\right]^{-1}\left(C\hat{\boldsymbol{\theta}}\right)=\sum_{i=1}^{k} w_i\hat{\theta}_i^2-\frac{1}{\sum_{i=1}^{k} w_i}\left(\sum_{i=1}^{k} w_i\hat{\theta}_i\right)^2, \tag{13}$$

where $V=\left[diag\left(v_i\right)\right]$ and $w_i=\dfrac{1}{v_i}=\left(\dfrac{S_i^2}{n_i}+\dfrac{S_i^4}{2\left(n_i-1\right)}\right)^{-1}$. The proposed test based on CAT method can be computed through the following steps.

1. Calculate the value of T^* in Eq. (13), denote it by $\hat{T}^*$.

2. Find the restricted MLEs of $\left(\sigma^2{}_1,\cdots,\sigma^2{}_{1k},\theta\right)$ in Eq. (5), that is, $\left(\tilde{\sigma}_1^2,\cdots,\tilde{\sigma}_k^2,\tilde{\theta}\right)$.

3. Generate artificial sample $X_{ij} \sim \mathrm{N}\left(\tilde{\theta}\text{-}0.5\tilde{\sigma}_i^2,\, \tilde{\sigma}_i^2\right),\quad j=1,...,\mathrm{n_i},\, i=1,...,\mathrm{k}$ a large of number of times (say, L times). For each of these replicated samples, recalculate the values of T^* and denote it by $\hat{T}^{*(l)},\quad l=1,...,L.$

4. Calculate the p-value as $p=\#\left(\hat{T}^{*(l)}>\hat{T}^*\right)/L$. In the case of p<α, H_0 is rejected.

We refer to this test as CAT2 in the simulation study.

Jafari and Abdollahnezhad [12] also proposed two tests based on CAT. These approaches are very similar to the previous test, in the first approach, they used the likelihood ratio statistic in Eq. (6), instead of T statistic in Eq. (13). We refer to this test based on the likelihood ratio statistic as CAT3. Therefore, we do not mention how to calculate these tests' steps again. In the second approach, they also used the modified likelihood ratio statistic in Eq. (7), instead of T statistic in Eq. (13). As mentioned in the remark, because the MLRT is not appropriate for some samples, we ignored this test based on CAT in the simulation study.

3. SIMULATION STUDY

In this section, the SLRT, L-GV, LW-GV, CAT1, CAT2 and CAT3 were compared with respect to their estimated type I error rates and powers. We considered the cases of k=3, 5 with small and moderately large sample sizes.

To estimate the type I error rates and powers of all the tests under the specified nominal level of 0.05, we generated 2000 random numbers with sample size n_i (i=1,…,k) from the log-normal distribution. We also took L=m=2000 to estimate the p-values. The simulation study was carried out in the MATLAB. The estimated type I error rates and powers of all the tests were given in the tables.

Table 1. *The estimated type I error rates of the all tests for k=3*

σ	n	CAT	CAT2	CAT3	SLRT	L-GV	LW-GV
1,2,2	6,6,6	0.033	0.033	0.047	0.047	0.009	0.010
	10,10,10	0.042	0.047	0.062	0.063	0.018	0.020
	20,20,20	0.052	0.048	0.052	0.052	0.028	0.031
	30,30,30	0.054	0.047	0.048	0.051	0.029	0.035
	6,6,10	0.026	0.038	0.048	0.047	0.007	0.008
	6,10,20	0.041	0.036	0.046	0.045	0.012	0.014
	10,20,30	0.040	0.041	0.047	0.048	0.018	0.022
1,2,4	6,6,6	0.028	0.045	0.050	0.051	0.013	0.015
	10,10,10	0.044	0.042	0.052	0.050	0.018	0.020
	20,20,20	0.046	0.046	0.048	0.048	0.029	0.032
	30,30,30	0.044	0.043	0.044	0.042	0.034	0.034
	6,6,10	0.035	0.033	0.044	0.046	0.006	0.013
	6,10,20	0.028	0.036	0.056	0.057	0.010	0.012
	10,20,30	0.042	0.050	0.060	0.060	0.027	0.030
1,4,9	6,6,6	0.017	0.034	0.044	0.042	0.017	0.019
	10,10,10	0.037	0.051	0.054	0.054	0.035	0.038
	20,20,20	0.033	0.059	0.058	0.059	0.047	0.048
	30,30,30	0.051	0.053	0.054	0.052	0.045	0.044
	6,6,10	0.017	0.046	0.061	0.062	0.020	0.021
	6,10,20	0.012	0.047	0.048	0.048	0.014	0.018
	10,20,30	0.033	0.046	0.045	0.047	0.028	0.031

Table 2. *The estimated type I error rates of the all tests for k=5*

σ	n	CAT1	CAT2	CAT3	SLRT	L-GV	LW-GV
1,1,2,2,2	6,6,6,6,6	0.042	0.034	0.049	0.052	0.004	0.006
	10,10,10,10,10	0.039	0.044	0.057	0.058	0.012	0.016
	20,20,20,20,20	0.046	0.045	0.050	0.049	0.026	0.029
	30,30,30,30,30	0.045	0.052	0.050	0.052	0.035	0.041
	6,6,6,10,10	0.042	0.031	0.053	0.052	0.008	0.011
	6,6,10,10,20	0.044	0.036	0.060	0.060	0.011	0.014
	10,10,20,20,30	0.052	0.043	0.044	0.044	0.016	0.022
1,1,2,2,4	6,6,6,6,6	0.036	0.032	0.052	0.055	0.006	0.010
	10,10,10,10,10	0.043	0.043	0.049	0.051	0.017	0.021
	20,20,20,20,20	0.043	0.050	0.048	0.048	0.028	0.035
	30,30,30,30,30	0.040	0.052	0.048	0.049	0.034	0.038
	6,6,6,10,10	0.040	0.027	0.045	0.044	0.008	0.009
	6,6,10,10,20	0.036	0.031	0.049	0.048	0.007	0.007
	10,10,20,20,30	0.040	0.045	0.049	0.049	0.014	0.021
1,2,2,4,9	6,6,6,6,6	0.029	0.033	0.039	0.038	0.013	0.013
	10,10,10,10,10	0.035	0.045	0.050	0.049	0.022	0.022
	20,20,20,20,20	0.044	0.041	0.046	0.043	0.029	0.032
	30,30,30,30,30	0.041	0.045	0.045	0.043	0.032	0.033
	6,6,6,10,10	0.025	0.047	0.061	0.061	0.015	0.017
	6,6,10,10,20	0.036	0.030	0.050	0.051	0.010	0.012
	10,10,20,20,30	0.043	0.045	0.049	0.050	0.017	0.023

Table 3. *The estimated powers of the all tests for k=3 and* $\boldsymbol{\theta}$*=(0.5, 1, 1.5)*

σ	n	CAT1	CAT2	CAT3	SLRT	L-GV	LW-GV
1,2,2	6,6,6	0.085	0.060	0.128	0.132	0.009	0.011
	10,10,10	0.163	0.121	0.187	0.189	0.040	0.047
	20,20,20	0.358	0.335	0.385	0.383	0.249	0.266
	30,30,30	0.551	0.522	0.560	0.559	0.454	0.467
	6,6,10	0.131	0.079	0.146	0.150	0.028	0.034
	6,10,20	0.208	0.201	0.193	0.194	0.094	0.112
	10,20,30	0.325	0.353	0.305	0.305	0.219	0.234
1,2,4	6,6,6	0.072	0.020	0.070	0.070	0.005	0.007
	10,10,10	0.118	0.038	0.112	0.113	0.008	0.013
	20,20,20	0.249	0.097	0.235	0.238	0.076	0.089
	30,30,30	0.364	0.193	0.325	0.323	0.189	0.202
	6,6,10	0.113	0.025	0.090	0.092	0.006	0.005
	6,10,20	0.182	0.092	0.134	0.134	0.026	0.037
	10,20,30	0.268	0.176	0.197	0.195	0.092	0.101
1,4,9	6,6,6	0.044	0.024	0.047	0.047	0.008	0.011
	10,10,10	0.074	0.026	0.065	0.067	0.019	0.020
	20,20,20	0.115	0.011	0.092	0.089	0.011	0.011
	30,30,30	0.158	0.021	0.132	0.133	0.039	0.039
	6,6,10	0.057	0.022	0.062	0.061	0.010	0.009
	6,10,20	0.075	0.026	0.078	0.075	0.007	0.009
	10,20,30	0.142	0.027	0.094	0.097	0.019	0.021

Table 4. *The estimated powers of the all tests for k=3 and θ=(0.5, 1, 2)*

σ	n	CAT1	CAT2	CAT3	SLRT	L-GV	LW-GV
1,2,2	6,6,6	0.119	0.112	0.238	0.240	0.022	0.027
	10,10,10	0.306	0.314	0.417	0.417	0.109	0.135
	20,20,20	0.666	0.719	0.751	0.753	0.590	0.612
	30,30,30	0.875	0.913	0.922	0.921	0.865	0.871
	6,6,10	0.224	0.168	0.271	0.273	0.063	0.077
	6,10,20	0.370	0.376	0.380	0.380	0.224	0.237
	10,20,30	0.608	0.656	0.618	0.623	0.515	0.523
1,2,4	6,6,6	0.102	0.029	0.138	0.141	0.004	0.006
	10,10,10	0.206	0.067	0.220	0.217	0.011	0.020
	20,20,20	0.496	0.248	0.453	0.453	0.187	0.208
	30,30,30	0.650	0.456	0.626	0.634	0.418	0.441
	6,6,10	0.185	0.048	0.173	0.175	0.008	0.011
	6,10,20	0.334	0.195	0.268	0.269	0.071	0.084
	10,20,30	0.492	0.369	0.420	0.422	0.223	0.243
1,4,9	6,6,6	0.049	0.018	0.061	0.058	0.006	0.010
	10,10,10	0.108	0.018	0.086	0.088	0.013	0.013
	20,20,20	0.203	0.016	0.151	0.150	0.021	0.024
	30,30,30	0.282	0.034	0.232	0.233	0.058	0.072
	6,6,10	0.091	0.021	0.082	0.080	0.008	0.011
	6,10,20	0.155	0.025	0.111	0.112	0.006	0.007
	10,20,30	0.258	0.051	0.181	0.185	0.031	0.034

Table 5. *The estimated powers of the all tests for k=5 and θ=(0.5, 0.5, 1, 1, 1.5)*

σ	n	CAT1	CAT2	CAT3	SLRT	L-GV	LW-GV
1,1,2,2,2	6,6,6,6,6	0.078	0.036	0.124	0.129	0.002	0.004
	10,10,10,10,10	0.146	0.097	0.206	0.206	0.018	0.029
	20,20,20,20,20	0.377	0.280	0.430	0.427	0.170	0.209
	30,30,30,30,30	0.565	0.517	0.618	0.625	0.448	0.469
	6,6,6,10,10	0.124	0.066	0.161	0.161	0.011	0.016
	6,6,10,10,20	0.219	0.146	0.212	0.213	0.064	0.079
	10,10,20,20,30	0.357	0.339	0.365	0.364	0.183	0.210
1,1,2,2,4	6,6,6,6,6	0.069	0.023	0.103	0.103	0.006	0.008
	10,10,10,10,10	0.133	0.036	0.140	0.140	0.007	0.011
	20,20,20,20,20	0.277	0.109	0.245	0.240	0.071	0.087
	30,30,30,30,30	0.418	0.217	0.389	0.392	0.189	0.214
	6,6,6,10,10	0.126	0.026	0.122	0.121	0.006	0.008
	6,6,10,10,20	0.184	0.054	0.130	0.126	0.011	0.016
	10,10,20,20,30	0.303	0.169	0.226	0.226	0.087	0.102
1,2,2,4,9	6,6,6,6,6	0.042	0.024	0.049	0.050	0.005	0.007
	10,10,10,10,10	0.079	0.028	0.069	0.072	0.012	0.015
	20,20,20,20,20	0.124	0.022	0.091	0.092	0.017	0.019
	30,30,30,30,30	0.164	0.027	0.146	0.145	0.030	0.040
	6,6,6,10,10	0.068	0.023	0.077	0.074	0.005	0.006
	6,6,10,10,20	0.079	0.025	0.065	0.067	0.006	0.006
	10,10,20,20,30	0.158	0.045	0.096	0.094	0.012	0.021

Table 6. *The estimated powers of the all tests for k=5 and* $\boldsymbol{\theta}$*=(0.5, 1, 1.5, 1.5, 2)*

σ	n	CAT1	CAT2	CAT3	SLRT	L-GV	LW-GV
1,1,2,2,2	6,6,6,6,6	0.118	0.072	0.172	0.174	0.008	0.012
	10,10,10,10,10	0.286	0.237	0.343	0.348	0.077	0.100
	20,20,20,20,20	0.667	0.643	0.687	0.688	0.500	0.540
	30,30,30,30,30	0.875	0.873	0.887	0.889	0.813	0.828
	6,6,6,10,10	0.194	0.128	0.241	0.238	0.033	0.046
	6,6,10,10,20	0.320	0.269	0.319	0.318	0.117	0.145
	10,10,20,20,30	0.576	0.564	0.524	0.531	0.352	0.394
1,1,2,2,4	6,6,6,6,6	0.109	0.045	0.150	0.149	0.003	0.007
	10,10,10,10,10	0.235	0.103	0.223	0.222	0.034	0.049
	20,20,20,20,20	0.512	0.397	0.509	0.512	0.305	0.341
	30,30,30,30,30	0.754	0.668	0.733	0.732	0.612	0.638
	6,6,6,10,10	0.176	0.062	0.153	0.152	0.012	0.015
	6,6,10,10,20	0.294	0.137	0.197	0.199	0.033	0.042
	10,10,20,20,30	0.465	0.351	0.358	0.364	0.194	0.218
1,2,2,4,9	6,6,6,6,6	0.069	0.019	0.084	0.084	0.004	0.005
	10,10,10,10,10	0.126	0.016	0.123	0.123	0.006	0.009
	20,20,20,20,20	0.260	0.058	0.246	0.250	0.049	0.061
	30,30,30,30,30	0.425	0.174	0.411	0.408	0.192	0.219
	6,6,6,10,10	0.092	0.023	0.103	0.105	0.003	0.004
	6,6,10,10,20	0.154	0.022	0.125	0.128	0.004	0.004
	10,10,20,20,30	0.280	0.099	0.202	0.202	0.041	0.053

From the numerical results in Table 1 and Table 2, it appears that the estimated type I error rates of the CAT3 and SLRT are close to the nominal level for all of the considered cases. The estimated type I error rates of the CAT1 and CAT2 are close to the nominal level for large sample sizes while the estimated type I error rates of these tests are smaller than the nominal level for small sample sizes. The L-GV and LW-GV have the estimated type I error rates smaller than the nominal level even for large sample sizes.

We also evaluated the powers of these tests for some sample sizes and parameter configurations. From the numerical results in Table 3, it appears that the powers of the CAT3 and especially the SLRT are better than the others. As the sample sizes increase, the powers of the CAT1 are close to those of the CAT3 and SLRT. Furthermore, the CAT1 are affected positively from the increase of the variances of the groups. In these cases, even for small sample sizes, the CAT1 has higher power when compared to the others while the estimated type I error rate of the CAT1 are smaller than the nominal level according to the CAT3 and SLRT. The L-GV and LW-GV have far smaller powers when compared to the others, even the values of these powers are smaller than the nominal level. As the sample sizes increase, the powers of these tests increase, but these powers are smaller than those of the others.

As expected, the powers of all tests increase with the difference of between the values of θ. From the numerical results in Table 4, it appears that the powers of the CAT3 and SLRT are better than the others especially for small sample sizes. As the variances of the groups increase, especially for large sample sizes, the power of the CAT1 is higher than the others.

The patterns that we noticed in Table 3 and Table 4 continue to hold in Table 5 and Table 6 as well.

4. CONCLUSION

In this paper, we investigated some methods based on likelihood-based methods, generalized pivotal (GV)-based methods and bootstrap-based methods for testing the equality of several log-normal means. We compared them with each other in terms of type I error rate and power by using Monte Carlo simulation for different number of groups and sample sizes. Simulation results indicate that the powers of the CAT3 and SLRT are better than the others especially for small sample sizes when the difference among the variances of the groups is small. As the variances of the groups increase, especially for large sample sizes, the power of the CAT1 is higher than the others.

CONFLICT OF INTEREST

No conflict of interest was declared by the authors

REFERENCES

[1] Chang, C.H. and Pal, N., "A revisit to the Behren-Fisher Problem: comparison of five test methods", Communications in Statistics-Simulation and Computation, 37(6):1064-1085, (2008).

[2] Chang, C.H., Lin, J.J. and Pal, N., "Testing the equality of several gamma means: a parametric bootstrap method with applications", Computational Statistics, 26(1):55-76, (2011).

[3] Crow, E.L. and Shimizu, K., Lognormal distribution, Mercel Dekker, New York, (1998).

[4] Gill, P.S., "Small-sample inference for the comparison of means of log-normal distributions", Biometrics, 60(2):525–527, (2004).

[5] Gökpınar, F. and Gökpınar, E., "Testing the equality of several log-normal means based on a computational approach," Communications in Statistics-Simulation and Computation 46(3): 1998-2010, (2017).

[6] Gökpınar, E. and Gökpınar, F., "A test based on computational approach for equality of means under unequal variance assumption", Hacettepe Journal of Mathematics and Statistics, 41(4):605-613, (2012).

[7] Gökpınar, E., Polat, E., Gökpınar, F. and Gunay, S., "A new computational approach for testing equality of ınverse gaussian means under heterogenity", Hacettepe Journal of Mathematics and Statistics, 42(5):581-590, (2013).

[8] Gökpınar, F. and Gökpınar, E., "A Computational approach for testing of coefficients of variation in k normal population", Hacettepe Journal of Mathematics and Statistics**,** 44(5):1197-1213, (2015a).

[9] Gökpınar, E. and Gökpınar, F., "Testing equality of variances for several normal populations", Communications in Statistics-Simulation and Computation, DOI: 10.1080/03610918.2014.955110, (2015b).

[10] Gupta, R.C. and Li, X., "Statistical inference for the common mean of two log-normal distributions and some applications in reliability", Computational Statistics & Data Analysis, 50:3141-3164, (2006).

[11] Guo, J. and Luh, W., "Testing methods for the one-way fixed effects ANOVA models of log-normal samples", Journal of Applied Statistics, 27:731-738, (2000).

[12] Jafari, A.A. and Abdollahnezhad, K., "Testing the equality means of several log-normal distributions", Communications in Statistics-Simulation and Computation, 46(3): 2311-2320, (2017).

[13] Krishnamoorthy, K. and Mathew, T., "Inferences on the means of lognormal distributions using generalized p-values and generalized confidence intervals", Journal of Statistical Planning and Inference, 115:103-121, (2003).

[14] Krishnamoorthy, K., and Oral, E., "Standardized likelihood ratio test for comparing several log-normal means and confidence interval for the common mean," Statistical methods in medical research, DOI:10.1177/0962280215615160, (2015).

[15] Lin, S.H., and Wang, R.S., "Modified method on the means for several log-normal distributions," Journal of Applied Statistics, 40(1): 194-208, (2013).

[16] Lin, S.H., "The higher order likelihood method for the common mean of several log-normal distributions," Metrika, 76(3): 381-392, (2013).

[17] Li, X., “A generalized p-value approach for comparing the means of several log-normal populations”, Statistics and Probability Letters, 79:1404-1408, (2009).

[18] Mutlu, H.T., Gökpınar, F., Gökpınar, E. and Gül, H.H., "A New Computational Approach Test for One-Way ANOVA under Heteroscedasticity”, Communications in Statistics - Theory and Methods, DOI: 10.1080/03610926.2016.117708, (2016).

[19] Pal, N., Lim, W. K. and Ling, C.H., “A computational approach to statistical inferences”, Journal of Applied Probability & Statistics, 2:13-35, (2007).

[20] Skovgaard, I. M., “Likelihood asymptotics”, Scandinavian Journal of Statistics, 28:3-32, (2001).

[21] Wu, J., Jiang, G., Wong, A.C.M. and Sun, X., “Likelihood analysis for the ratio of means of two independent log-normal distributions”, Biometrics, 58:463–469, (2002).

[22] Wu, J., Wong, A.C.M. and Jiang, G., “Likelihood-based confidence intervals for a lognormal mean”, Stat. Med., 22:1849–1860, (2003).

[23] Zhou, X. H., Gao, S. and Hui, S. L., “Methods for comparing the means of two independent log-normal samples”, Biometrics, 53:1129-1135, (1997).

[24] Zhou, X.and Tu, W., “Comparison of several independent population means when their samples contain log-normal and possibly zero observations”, Biometrics, 55:645-651, (1999).