

Research Article

Improved Small Sample Inference on the Ratio of Two Coefficients of Variation of Two Independent Lognormal Distributions

A. Wong ¹ and L. Jiang²

¹Department of Mathematics and Statistics, York University, 4700 Keele Street, Toronto, Ontario, Canada M3J 1P3

²Beijing Education Examinations Authority, Beijing, China

Correspondence should be addressed to A. Wong; august@yorku.ca

Received 5 July 2018; Revised 29 October 2018; Accepted 27 January 2019; Published 3 March 2019

Guest Editor: Xu Zhang

Copyright © 2019 A. Wong and L. Jiang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Without the ability to use research tools and procedures that yield consistent measurements, researchers would be unable to draw conclusions, formulate theories, or make claims about generalizability of their results. In statistics, the coefficient of variation is commonly used as the index of reliability of measurements. Thus, comparing coefficients of variation is of special interest. Moreover, the lognormal distribution has been frequently used for modeling data from many fields such as health and medical research. In this paper, we proposed a simulated Bartlett corrected likelihood ratio approach to obtain inference concerning the ratio of two coefficients of variation for lognormal distribution. Simulation studies show that the proposed method is extremely accurate even when the sample size is small.

1. Introduction

In health and medical research, it is common that the variable of interest, X , such as the survival time, takes only positive values and the underlying distribution of this variable is highly skewed to the right. In this case, the frequently assumed normal distribution for X is not suitable. A standard approach to first transform X such that the transformed variable $Y = g(X)$ is normally distributed. Then the existing statistical theories developed for the normal distribution can be applied. For $X > 0$ and the distribution of X is highly skewed to the right, the most common transformation is the logarithmic transformation. In other words, $Y = \log(X)$ is normally distributed. Hence, X is lognormally distributed. Detailed review of the theories of the lognormal distribution can be found in Aitchison and Brown [1], and Crow and Simizu [2]. In practice, Fears et al. [3] investigated the variability and reproducibility of hormone assays used by laboratories with the capability of performing large numbers of tests. They assumed the hormone samples used in laboratories are independent lognormally distributed.

In this case, it is of special interest to know if each sample yields consistent measurements.

The coefficient of variation (τ) is defined as the ratio of the standard deviation to the mean, where the mean is assumed to be non zero. It is an important index for assessment of the reliability of a measuring procedure. Hence, the problem considered in Fears et al. [3] can be viewed as testing if the coefficients of variation used in each laboratory are the same or not.

Mathematically, if a random variable X is distributed as lognormal(μ, σ), then $Y = \log(X)$ is distributed as normal with mean μ and variance σ^2 . It is well-known that

$$E(X) = \exp\left\{\mu + \frac{\sigma^2}{2}\right\}, \quad (1)$$

$$\text{and } \text{var}(X) = \exp\{2\mu + \sigma^2\} [\exp\{\sigma^2\} - 1].$$

Hence, the coefficient of variation, τ , is

$$\tau = \frac{\sqrt{\text{var}(X)}}{E(X)} = \sqrt{\exp\{\sigma^2\} - 1} > 0. \quad (2)$$

Nam and Kwon [4] compared various approximate interval estimations of the ratio of two coefficients of variation for independent lognormal distributions. And their simulation results showed that empirical coverage rates of these methods are satisfactorily close to the nominal coverage rate for medium sample size. The aim of this paper is to develop a more accurate method to obtain inference for the ratio of two coefficients of variation for independent lognormal distributions. Moreover, the proposed method can be generalized to test if the coefficients of variation from k independent lognormal distributions are heterogeneous.

The rest of the paper is organized as follows. Section 2 reviewed the existing methods for obtaining inference concerning the ratio of two coefficients of variation from independent lognormal distribution. The simulated Bartlett corrected likelihood method is proposed in Section 3. A real data example is presented in Section 4 to illustrate the application of the method discussed in this paper. Simulation studies are performed to compare the accuracy of the methods discussed in this paper in Section 5. Extension to testing for homogeneity of coefficients of variations from k independent lognormal distributions is discussed in Section 6. Some concluding remarks are recorded in Section 7.

2. Existing Methods for Inference on the Ratio of Two Coefficients of Variation of Two Independent Lognormal Distributions

Let $(x_{i1}, \dots, x_{in_i})$ be the i^{th} sample from the lognormal (μ_i, σ_i) distribution, where $i = 1, \dots, k$. Then $(y_{i1}, \dots, y_{in_i}) = (\log x_{i1}, \dots, \log x_{in_i})$ is the i^{th} sample from the normal distribution with mean μ_i and variance σ_i^2 . From (2), the i^{th} coefficient of variation is $\tau_i = \sqrt{\exp\{\sigma_i^2\} - 1}$. Nam and Kwon [4] compared four methods in obtaining confidence intervals for $\psi = \tau_1/\tau_2$. The following is the summary of the methods discussed in Nam and Kwon [4]:

(1) Wald type method

Let the observed test statistic be

$$z_W(\psi) = \frac{\hat{\psi} - \psi}{\sqrt{\widehat{\text{var}}(\hat{\psi})}} \quad (3)$$

where $\hat{\psi} = \hat{\tau}_1/\hat{\tau}_2$, $\hat{\tau}_i = \sqrt{\exp\{\hat{\sigma}_i^2\} - 1}$, $\hat{\sigma}_i^2 = (1/n_i) \sum_{j=1}^{n_i} (\log x_{ij} - \sum_{h=1}^{n_i} \log x_{ih}/n_i)^2$, and $\widehat{\text{var}}(\hat{\psi}) = (n_1 \psi^4 [(1 + \hat{\tau}_2^2) \hat{\sigma}_2^2]^2 + n_2 [(1 + \hat{\tau}_1^2) \hat{\sigma}_1^2]^2) / 2n_1 n_2 (\hat{\tau}_1 \hat{\tau}_2)^2$. Then $Z_W(\psi)$ is asymptotically distributed as standard normal distribution. The significance function of ψ is $p(\psi) = \Phi(z_W(\psi))$, where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution.

(2) Fieller type method

Let the observed test statistic be

$$z_F(\psi) = \frac{\hat{\tau}_1 - \psi \hat{\tau}_2}{\sqrt{\widehat{\text{var}}(\hat{\tau}_1) + \psi^2 \widehat{\text{var}}(\hat{\tau}_2)}} \quad (4)$$

where

$$\widehat{\text{var}}(\hat{\tau}_i) = \frac{\hat{\sigma}_i^2 (1 + \hat{\tau}_i)^2}{2n_i \hat{\tau}_i^2} \quad i = 1, 2. \quad (5)$$

Then $Z_F(\psi)$ is also asymptotically distributed as standard normal distribution. The significance function of ψ is $p(\psi) = \Phi(z_F(\psi))$.

(3) Log method

Let the observed test statistic be

$$z_L(\psi) = \frac{\log \hat{\tau}_1 - \log \hat{\tau}_2 - \log \psi}{\sqrt{\widehat{\text{var}}(\log \hat{\tau}_1) + \widehat{\text{var}}(\log \hat{\tau}_2)}} \quad (6)$$

where

$$\widehat{\text{var}}(\log \hat{\tau}_i) = \frac{\widehat{\text{var}}(\hat{\tau}_i)}{\hat{\tau}_i^2} \quad i = 1, 2. \quad (7)$$

Then $Z_L(\psi)$ is also asymptotically distributed as standard normal distribution. The significance function of ψ is $p(\psi) = \Phi(z_L(\psi))$.

(4) Method of variance estimates recovery (MOVER)

This is a method that will directly obtain an approximate $(1 - \alpha)100\%$ confidence interval for ψ only. Let

$$l_i = \log \hat{\tau}_i - z_{\alpha/2} \sqrt{\widehat{\text{var}}(\log \hat{\tau}_i)}, \quad (8)$$

$$\text{and } u_i = \log \hat{\tau}_i + z_{\alpha/2} \sqrt{\widehat{\text{var}}(\log \hat{\tau}_i)}.$$

Then an approximate $(1 - \alpha)100\%$ confidence interval for $\log \psi$ is (L, U) where

$$\begin{aligned} L &= (\log \hat{\tau}_1 - \log \hat{\tau}_2) \\ &\quad - \sqrt{(\log \hat{\tau}_1 - l_1)^2 + (u_2 - \log \hat{\tau}_2)^2}, \\ U &= (\log \hat{\tau}_1 - \log \hat{\tau}_2) \\ &\quad + \sqrt{(\log \hat{\tau}_1 - l_1)^2 + (u_2 - \log \hat{\tau}_2)^2}. \end{aligned} \quad (9)$$

Thus, an approximate $(1 - \alpha)100\%$ confidence interval for ψ is $(\exp\{L\}, \exp\{U\})$. If $\widehat{\text{var}}(\log \hat{\tau}_i)$, for $i = 1, 2$, to be the same as that obtained in the Log method, the MOVER method is identical to the Log method. Note that Hasan and Krishamoorthy [5] proposed an improved version of the MOVER method.

3. Proposed Method

In this section, we will first review the likelihood based methods and the Bartlett corrected likelihood ratio method. Since the required Bartlett adjustment for the Bartlett corrected likelihood ratio method is very difficult to obtain, a numerical algorithm is proposed to approximate the Bartlett adjustment. Then the methods are applied to obtain inference for the ratio of two coefficients of variation of two independent lognormal distribution.

3.1. Likelihood Based Methods and Bartlett Corrected Likelihood Ratio Method. Let (x_1, \dots, x_n) be a sample from a known distribution with probability density function $f(\cdot, \theta)$, where θ is a p -dimensional vector of parameters. Let $\psi = \psi(\theta)$, which has dimension $d < p$ be the parameter of interest. The log-likelihood function is

$$\ell(\theta) = \ell(\theta; x_1, \dots, x_n) = \sum_{i=1}^n \log f(x_i; \theta). \quad (10)$$

Under the regularity conditions stated in Barndorff-Nielsen and Cox [6], we have the standardized maximum likelihood estimate (MLE) statistic $(\hat{\theta} - \theta)' [\text{var}(\hat{\theta})]^{-1} (\hat{\theta} - \theta)$ and the likelihood ratio statistic $2[\ell(\hat{\theta}) - \ell(\theta)]$ that are asymptotically chi-square distributed with p degrees of freedom, χ_p^2 , where $\hat{\theta}$ is the overall MLE, which is the value of θ that maximized $\ell(\theta)$, and $\text{var}(\hat{\theta})$ is approximately the inverse of the Fisher's expected information. When the parameter of interest is $\psi = \psi(\theta)$, Barndorff-Nielsen and Cox [6] showed that similar statistics can be obtained. The standardized MLE statistic becomes

$$Q(\psi) = (\hat{\psi} - \psi)' [\text{var}(\hat{\psi})]^{-1} (\hat{\psi} - \psi) \quad (11)$$

where $\hat{\psi} = \psi(\hat{\theta})$, and $\text{var}(\hat{\psi})$ can be approximated by the delta method, which takes the form

$$\text{var}(\hat{\psi}) \approx \left\{ \frac{\partial \psi(\theta)}{\partial \theta} \right\}' \text{var}(\hat{\theta}) \left\{ \frac{\partial \psi(\theta)}{\partial \theta} \right\}. \quad (12)$$

The likelihood ratio statistic is

$$W(\psi) = 2[\ell(\hat{\theta}) - \ell(\tilde{\theta})], \quad (13)$$

where $\tilde{\theta}$ is the constrained MLE, which is obtained by maximizing $\ell(\theta)$ for the given ψ value. Both $Q(\psi)$ and $W(\psi)$ are asymptotically χ_d^2 . As defined in Fraser [7], the significance function for ψ is defined as $p(\psi) = P(\chi_d^2 \leq q(\psi))$ or $p(\psi) = P(\chi_d^2 \leq w(\psi))$ can be used to obtain inference concerning ψ where $q(\psi)$ and $w(\psi)$ are the observed values of $Q(\psi)$ and $W(\psi)$, respectively. In particular, the $(1-\alpha)100\%$ confidence region of ψ is

$$\begin{aligned} & \{\psi : q(\psi) \leq \chi_{d,1-\alpha}^2\} \\ & \text{and } \{\psi : w(\psi) \leq \chi_{d,1-\alpha}^2\}, \end{aligned} \quad (14)$$

respectively, where $\chi_{d,1-\alpha}^2$ is the $(1-\alpha)100^{\text{th}}$ percentile of χ_d^2 .

It is well-known that these two asymptotic methods have rate of convergence $O(n^{-1/2})$, and they are referred to as the first-order methods. In statistics literature, there exists various adjustments to improve the accuracy of the above methods. In particular, Barndorff-Nielsen [8, 9] introduced the modified signed log-likelihood ratio statistics, a third-order method. However, this method is restricted to scalar parameter of interest only. On the other hand, Bartlett [10] proposed a transformation of the likelihood ratio statistic

such that the mean of the transformed statistic matched the mean of the asymptotic distribution. More specifically,

$$W^*(\psi) = \frac{W(\psi)}{B} \quad (15)$$

where B is the Bartlett adjustment such that $E[W^*(\psi)] = d$. And $W^*(\cdot)$ is known as the Bartlett corrected likelihood ratio statistic. An obvious choice of B is

$$B = \frac{E[W^*(\psi)]}{d}. \quad (16)$$

Bartlett [10] showed that the Bartlett corrected likelihood ratio statistic is also asymptotically χ_d^2 distributed and it has rate of convergence $O(n^{-2})$. Therefore, it is an extremely accurate method. Nevertheless, except in a few well-defined problem, $E[W^*(\psi)]$ is very difficult to obtain which hinders the use of this method in applied statistics. A review of the Bartlett corrected likelihood ratio method can be found in Barndorff-Nielsen and Cox [6].

Although, mathematically, the explicit closed form of B , or even an asymptotic expansion of B , is difficult to obtain, we propose the following algorithmic way to obtain $E[W(\psi)]$ numerically, and hence, an estimated B .

Given: (x_1, \dots, x_n) is a sample of size n from a distribution with known probability density function $f(\cdot; \theta)$.

Interest: Inference concerning $\psi = \psi(\theta)$.

Have: Overall maximum likelihood estimate $\hat{\theta}$, the constrained maximum likelihood estimate $\tilde{\theta}$, and the observed likelihood ratio statistic $w(\psi)$.

Step 1: Simulate M samples of data of size n from $f(\cdot; \tilde{\theta})$.

Step 2: For each set of simulated data, obtain the simulated observed likelihood ratio statistic. As a result, we have $w_1(\psi), \dots, w_M(\psi)$.

Step 3: Calculate

$$\bar{w}(\psi) = \frac{\sum_{i=1}^M w_i(\psi)}{M}, \quad (17)$$

which is an estimate of the mean of the likelihood ratio statistic. Hence, we have $\hat{B} = \bar{w}(\psi)/d$.

Step 4: The observed simulated Bartlett corrected likelihood ratio statistic is

$$w^*(\psi) = \frac{w(\psi)}{\hat{B}}, \quad (18)$$

is asymptotically distributed as χ_d^2 with fourth order rate of convergence. Thus, the significance function is $p(\psi) = P(\chi_d^2 \leq w^*(\psi))$, and the $(1-\alpha)100\%$ confidence region of ψ is

$$\{\psi : w^*(\psi) \leq \chi_{d,1-\alpha}^2\}. \quad (19)$$

As a final note on the proposed algorithm, theoretically, the choice of M should be as large as possible. However, the larger M is, the more calculations are required to obtain $\bar{w}(\psi)$. Moreover, the more nuisance parameters exist in the model, the larger M has to be. We recommend to use trial-by-error of M until $\bar{w}(\psi)$ is stabilized.

3.2. *Applying Likelihood Based Method to Obtain Inference on the Ratio of Two Coefficients of Variation of Two Independent Log Normal Distribution.* Let $\bar{Y}_i = \sum_{j=1}^{n_i} Y_{ij}/n_i$ and $(n_i-1)S_i^2 = \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_i)^2$. Then \bar{Y}_i is normally distributed with mean μ_i and variance σ_i^2/n , and $(n_i-1)S_i^2/\sigma_i^2$ is $\chi_{n_i-1}^2$. Moreover, \bar{Y}_i and $(n_i-1)S_i^2/\sigma_i^2$ are independent. Hence, inference concerning σ_i^2 will be based on $\chi_{n_i-1}^2$. Since τ_i is a function of σ_i^2 only, inference concerning τ_i will be based on $\chi_{n_i-1}^2$. Let $\theta = (\sigma_1^2, \sigma_2^2)$. Then the likelihood function for θ can be written as

$$\begin{aligned} \ell(\theta) &= \ell(\sigma_1^2, \sigma_2^2; s_1^2, s_2^2) \\ &= -\frac{n_1-1}{2} \log \sigma_1^2 - \frac{(n_1-1)s_1^2}{2\sigma_1^2} - \frac{n_2-1}{2} \log \sigma_2^2 \\ &\quad - \frac{(n_2-1)s_2^2}{2\sigma_2^2}. \end{aligned} \quad (20)$$

It is easy to show that the overall MLE

$$\hat{\theta} = (\hat{\sigma}_1^2, \hat{\sigma}_2^2) = (s_1^2, s_2^2). \quad (21)$$

Since our parameter of interest is $\psi = \psi(\theta) = \tau_1/\tau_2$, where $\tau_i = \sqrt{\exp\{\sigma_i^2\} - 1}$, we have

$$\sigma_1^2 = \log(\psi^2 \exp\{\sigma_2^2\} - \psi^2 + 1). \quad (22)$$

For a given ψ value, the log-likelihood function in (20) can be expressed as a function of σ_2^2 only, and is

$$\begin{aligned} \ell(\sigma_2^2) &= \ell(\sigma_2^2; s_1^2, s_2^2) \\ &= -\frac{n_1-1}{2} \log[\log(\psi^2 \exp\{\sigma_2^2\} - \psi^2 + 1)] \\ &\quad - \frac{(n_1-1)s_1^2}{2[\log(\psi^2 \exp\{\sigma_2^2\} - \psi^2 + 1)]} \\ &\quad - \frac{n_2-1}{2} \log \sigma_2^2 - \frac{(n_2-1)s_2^2}{2\sigma_2^2}. \end{aligned} \quad (23)$$

Hence, to solve for the constrained MLE $\tilde{\theta} = (\tilde{\sigma}_1^2, \tilde{\sigma}_2^2)$, we have to find $\tilde{\sigma}_2^2$ that maximized (23), and then $\tilde{\sigma}_1^2 = \log(\psi^2 \exp\{\tilde{\sigma}_2^2\} - \psi^2 + 1)$. Once we have both the overall and constrained MLEs, we can obtain the observed likelihood ratio statistic $w(\psi)$ as given in (13). Therefore, the significance function is $p(\psi) = P(\chi_1^2 \leq w(\psi))$. Moreover, by applying the algorithm given in the previous section, we can also obtain the observed simulated modified likelihood ratio statistic $w^*(\psi)$ and the corresponding significance function is $p(\psi) = P(\chi_1^2 \leq w^*(\psi))$.

4. Real Data Example

To illustrate the application of the methods discussed in this paper, we revisit the example discussed in Nam and Kwon [4].

TABLE 1: 95% confidence interval for ψ for the experiment by Faupel-Badger et al. [11].

Method	95% confidence interval for ψ
Wald type	(2.1798, 3.8620)
Fieller type	(2.2705, 3.9992)
Log method	(2.2867, 3.9907)
Likelihood ratio	(2.2776, 4.0126)
Bartlett correction	(2.2770, 4.0138)

Faupel-Badger et al. [11] compare concentrations of estrogen metabolites by RIA with the concentrations obtained using a novel and high-performance liquid chromatography-tandem mass spectrometry (LC-MS/MS). The 10% blinded quality control samples were used for assessment of quality control of the laboratory assay. Partial summary of data were presented in Nam and Kwon [4] and we have

$$\begin{aligned} n_1 &= 48, \\ \hat{\sigma}_1^2 &= (0.1687)^2, \\ n_2 &= 53, \\ \hat{\sigma}_2^2 &= (0.0562)^2 \end{aligned} \quad (24)$$

where the first sample is taken from RIA, and the second sample is taken from LC-MS/MS. Table 1 records the 95% confidence interval for the ratio of the two coefficients of variation assuming that the data are obtained from independent lognormal distributions obtained by the methods discussed in this paper. Note that the MOVER method is identical to the Log method and Hasan and Krishnamoorthy [5] showed that results from the improved version of the MOVER method are still similar to those obtained by the Log method. Hence, both the MOVER method and its improved version are not included in the calculations. Except for the Wald type, the intervals obtained in Nam and Kwon [4] seem to be close to each other. Notice that the results from the Fieller type are different from that reported in Nam and Kwon [4]. Moreover, we observed that the likelihood ratio method and the proposed Bartlett correction method seem to be different from the other methods by having a larger upper confidence limit.

With the above observation, it is of interest to compare the accuracy of the methods discussed in this paper, especially when the sample size is small.

5. Simulation Studies

To compare the accuracy of the methods discussed in this paper, simulation studies are performed. The parameters settings are given in Table 2. Other settings have also been calculated but not reported because the results are very similar to those presented. However, they are available upon request. Since we are interested in developing a method that is accurate even for small sample sizes, hence the chosen sample sizes in the simulation studies are relatively small.

TABLE 2: Parameters settings for simulation studies.

Study	μ_1	σ_1	μ_2	σ_2	n_1	n_2
1	0.6	0.1	3.0	0.5	5	5
2					10	20
3					15	25
4					20	10
5	1.1	0.2	0.8	0.4	5	5
6					10	20
7					15	25
8					20	10
9	2.5	1.2	3.0	0.7	5	5
10					10	20
11					15	25
12					20	10
13	5.0	0.7	6.0	1.4	5	5
14					10	20
15					15	25
16					20	10

For each study, we obtain 10,000 simulated samples. Theoretically, M should be as large as possible because we want to use $\bar{w}(\psi)$ to be the estimate of $E[W(\psi)]$. However, numerically, we have N simulated samples, and for each simulated sample, we have to do M simulations to obtain $\bar{w}(\psi)$. For these simulation studies, we use $M = 500$. For each simulated sample, we compute the 95% confidence interval obtained by the methods discussed in this paper. Table 3 reported the percentage of samples where the true ψ is less than the lower 95% confidence limit (le), is within the 95% confidence interval (cc), and is greater than the upper 95% confidence limit (ue). The nominal values are 2.5%, 95% and 2.5%, respectively.

From Table 3, the three methods discussed in Nam and Kwon [4] do not give satisfactory coverage, especially when the sample sizes are small. The coverage of the likelihood ratio method is improving when the sample sizes increase and, in general, it has asymmetric errors. Nevertheless the proposed simulated Bartlett corrected likelihood ratio method is extremely accurate even when the sample sizes are as small as 5.

6. Testing Homogeneity of Coefficients of Variation from k Independent Lognormal Distributions

For k samples from independent lognormal(μ_i, σ_i) distribution, the required log-likelihood function can be written as

$$\begin{aligned} \ell(\theta) &= \ell(\sigma_1^2, \dots, \sigma_k^2; s_1^2, \dots, s_k^2) \\ &= \sum_{i=1}^k \left[-\frac{n_i - 1}{2} \log \sigma_i^2 - \frac{(n_i - 1) s_i^2}{2\sigma_i^2} \right] \end{aligned} \tag{25}$$

TABLE 3

(a) Empirical coverage rate for the simulation studies 1 to 8

Study	Method	le	cc	ue
1	Wald type	0.08	87.78	12.14
	Fieller type	3.48	95.50	1.02
	Log method	6.28	88.51	5.21
	Likelihood ratio	3.12	93.39	3.49
	Bartlett correction	2.40	94.90	2.70
2	Wald type	0.20	89.31	10.49
	Fieller type	0.82	94.01	5.17
	Log method	1.94	92.71	5.35
	Likelihood ratio	2.22	94.78	3.00
	Bartlett correction	1.92	95.29	2.79
3	Wald type	0.68	92.92	6.40
	Fieller type	3.48	95.50	1.02
	Log method	3.93	93.25	2.82
	Likelihood ratio	2.77	94.57	2.66
	Bartlett correction	2.53	94.85	2.27
4	Wald type	1.44	94.14	4.42
	Fieller type	7.81	91.99	0.20
	Log method	7.04	91.46	1.50
	Likelihood ratio	3.15	94.23	2.62
	Bartlett correction	2.88	94.85	2.27
5	Wald type	0.05	87.47	12.48
	Fieller type	2.99	95.28	1.73
	Log method	5.88	88.45	5.67
	Likelihood ratio	3.13	93.39	3.48
	Bartlett correction	2.42	94.90	2.68
6	Wald type	0.13	89.18	10.69
	Fieller type	0.67	93.59	5.74
	Log method	1.84	92.58	5.58
	Likelihood ratio	2.25	94.75	3.01
	Bartlett correction	2.00	95.23	2.77
7	Wald type	0.57	92.73	6.70
	Fieller type	2.87	95.26	1.87
	Log method	3.68	93.20	3.12
	Likelihood ratio	2.70	94.66	2.64
	Bartlett correction	2.52	94.90	2.58
8	Wald type	1.34	94.14	4.52
	Fieller type	7.36	92.28	0.36
	Log method	6.68	91.63	1.69
	Likelihood ratio	3.15	94.34	2.49
	Bartlett correction	2.88	94.83	2.29

(b) Empirical coverage rate for the simulation studies 9 to 16

Study	Method	le	cc	ue
9	Wald type	0.00	82.10	17.90
	Fieller type	0.01	98.88	1.11
	Log method	1.27	90.26	8.47
	Likelihood ratio	2.99	93.61	3.40
	Bartlett correction	2.40	94.94	2.66

(b) Continued.

Study	Method	le	cc	ue
10	Wald type	0.00	83.91	16.09
	Fieller type	0.00	90.18	9.82
	Log method	0.07	90.86	9.07
	Likelihood ratio	2.04	94.82	3.14
	Bartlett correction	1.92	95.24	2.84
11	Wald type	0.00	89.36	10.64
	Fieller type	0.00	96.70	3.30
	Log method	0.96	93.93	5.11
	Likelihood ratio	2.75	94.59	2.66
	Bartlett correction	2.48	94.97	2.55
12	Wald type	0.02	92.69	7.29
	Fieller type	0.47	99.16	0.37
	Log method	3.62	93.86	2.52
	Likelihood ratio	3.17	94.34	2.49
	Bartlett correction	2.96	94.71	2.33
13	Wald type	0.00	79.33	20.67
	Fieller type	2.01	97.99	0.00
	Log method	10.56	88.87	0.57
	Likelihood ratio	3.52	94.54	2.94
	Bartlett correction	2.68	95.03	2.29
14	Wald type	0.01	91.26	8.73
	Fieller type	0.64	99.32	0.04
	Log method	3.15	94.73	2.12
	Likelihood ratio	2.44	94.69	2.87
	Bartlett correction	2.10	95.19	2.71
15	Wald type	0.59	93.88	5.53
	Fieller type	5.44	94.56	0.00
	Log method	6.29	93.43	0.28
	Likelihood ratio	3.01	94.48	2.51
	Bartlett correction	2.84	94.83	2.33
16	Wald type	1.60	94.18	4.22
	Fieller type	12.16	87.84	0.00
	Log method	10.64	89.36	0.00
	Likelihood ratio	3.01	94.48	2.51
	Bartlett correction	2.95	94.71	2.34

where s_i^2 is the unbiased sample variance estimate of the i^{th} sample given in Section 3. It is well-known that the overall MLE is

$$\hat{\theta} = (\hat{\sigma}_1^2, \dots, \hat{\sigma}_k^2) = (s_1^2, \dots, s_k^2). \tag{26}$$

The aim is to test

$$H_0: \tau_1 = \dots = \tau_k = \tau \tag{27}$$

vs H_a : not all coefficients of variation are the same,

which, in this case, is the same as testing

$$H_0: \sigma_1^2 = \dots = \sigma_k^2 = \sigma^2 \tag{28}$$

vs H_a : not all variances are the same.

Therefore, when H_0 is true, the log-likelihood function can be re-written in terms of σ^2 and is

$$\begin{aligned} \ell(\sigma^2) &= \ell(\sigma^2; s_1^2, \dots, s_k^2) \\ &= \sum_{i=1}^k \left[-\frac{n_i - 1}{2} \log \sigma^2 - \frac{(n_i - 1) s_i^2}{2\sigma^2} \right], \end{aligned} \tag{29}$$

and the constrained MLE is

$$\tilde{\sigma}^2 = \frac{\sum_{i=1}^k (n_i - 1) s_i^2}{\sum_{i=1}^k (n_i - 1)}, \tag{30}$$

which is the usual pooled variance estimate. The observed likelihood ratio statistic is

$$w = 2 \left[\ell(\hat{\theta}) - \ell(\tilde{\sigma}^2) \right], \tag{31}$$

which is asymptotically distributed as χ_{k-1}^2 . Hence, the observed simulated Bartlett corrected likelihood ratio statistic is

$$w^* = \frac{w}{\bar{w}/(k-1)}, \tag{32}$$

where \bar{w} is obtained by the algorithm given in Section 2.

Simulation studies are performed to compare the accuracy of the likelihood ratio method and the simulated Bartlett corrected likelihood ratio method. In particular, three samples of data from lognormal(μ_i, σ) distribution are generated. w is calculated and w^* is also the calculation with $M = 1000$. We repeat this process $N = 10,000$. The proportion of samples that have p -values less than 5% is reported in Table 4 for various sample sizes. The choice of μ_i is not important because it does not involve in any of the calculations and, hence, we take it to be 0. Different choices of σ result in similar results and are not reported, but they are available upon request. Table 4 reported the cases $\mu_i = 0$ and $\sigma = 1$. When sample sizes are small, the likelihood ratio method does not give satisfactory results, but it is improving when the sample sizes increase. The simulated Bartlett corrected likelihood ratio method consistently gives extremely accurate result even when the sample sizes are small.

7. Conclusion

The lognormal distribution has been frequently used for modeling positive valued right skewed data, which commonly arise in health and medical research. In this paper, we proposed a simulated Bartlett corrected likelihood ratio approach to obtain inference concerning the ratio of two coefficients of variation for lognormal distribution. Simulation studies show that the proposed Bartlett correction method is extremely accurate even when the sample size is small. Moreover, the proposed proposed Bartlett correction method is extended to test homogeneity of k coefficients of variation from independent lognormal distributions.

TABLE 4: The proportion of samples rejected at 5% level of significance for testing $H_0 : \tau_1 = \tau_2 = \tau_3 = \tau$ when data are generated from lognormal(0, 1) distribution.

(n_1, n_2, n_3)	Proportion of samples rejected at $\alpha = 0.05$	
	Likelihood ratio	Bartlett correction
(5, 5, 5)	0.0658	0.0512
(5, 10, 15)	0.0646	0.0524
(10, 10, 10)	0.0565	0.0495
(10, 15, 20)	0.0529	0.0487
(50, 50, 50)	0.0544	0.0523

Data Availability

The data set for compared concentrations of estrogen metabolites by RIA with the concentrations obtained using a novel and high-performance liquid chromatography-tandem mass spectrometry (LC-MS/MS) is from previously reported in Faupel-Badger et al. [11], which has been cited. This data set was further analyzed in Nam and Kwon [4], which was also cited in the manuscript. The other numerical examples in the submitted paper are based on simulation studies, which is available from the corresponding author upon request.

- [11] J. M. Faupel-Badger, B. J. Fuhrman, X. Xu et al., "Comparison of liquid chromatography-tandem mass spectrometry, RIA, and ELISA methods for measurement of urinary estrogens," *Cancer Epidemiology, Biomarkers & Prevention*, vol. 19, no. 1, pp. 292–300, 2010.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] J. Aitchison and J. A. C. Brown, *The Lognormal Distribution*, Cambridge University Press, Cambridge, UK, 1957.
- [2] E. L. Crow and K. Simizu, *Lognormal Distributions. Theory and Application*, Marcel Dekker, New York, NY, USA, 1988.
- [3] T. R. Fears, R. G. Ziegler, J. L. Donaldson et al., "Reproducibility studies and interlaboratory concordance for androgen assays of male plasma hormone levels," *Cancer Epidemiology, Biomarkers & Prevention*, vol. 11, no. 8, pp. 785–789, 2002.
- [4] J. Nam and D. Kwon, "Inference on the ratio of two coefficients of variation of two lognormal distributions," *Communications in Statistics—Theory and Methods*, vol. 46, no. 17, pp. 8575–8587, 2017.
- [5] M. S. Hasan and K. Krishnamoorthy, "Improved confidence intervals for the ratio of coefficients of variation of two lognormal distributions," *Journal of Statistical Theory and Applications*, vol. 16, no. 3, pp. 345–353, 2017.
- [6] O. E. Barndorff-Nielsen and D. R. Cox, *Inference and Asymptotics*, Chapman and Hall, New York, NY, USA, 1994.
- [7] D. A. S. Fraser, "P-values: The insight to modern statistical inference," *Annual Review of Statistics and Its Application*, vol. 4, pp. 1–14, 2017.
- [8] O. E. Barndorff-Nielsen, "Inference on full or partial parameters based on the standardized signed log likelihood ratio," *Biometrika*, vol. 73, no. 2, pp. 307–322, 1986.
- [9] O. E. Barndorff-Nielsen, "Modified signed log likelihood ratio," *Biometrika*, vol. 78, no. 3, pp. 557–563, 1991.
- [10] M. S. Bartlett, "Properties of sufficiency and statistical tests," *Proceedings of the Royal Society A Mathematical, Physical and Engineering Sciences*, vol. 160, no. 901, pp. 268–282, 1937.

