Far East Journal of Mathematical Sciences (FJMS)

http://www.pphmj.com
http://dx.doi.org/10.17654/MS102040749
Volume 102, Number 4, 2017, Pages 749-767 **ISSN: 0972-0871**

CONFIDENCE INTERVALS FOR COMMON VARIANCE OF NORMAL DISTRIBUTIONS

Narudee Smithpreecha and Suparat Niwitpong

Department of Applied Statistics
Faculty of Applied Science
King Mongkut's University of Technology North Bangkok
Bangkok, 10800, Thailand

Abstract

This paper presents four approaches to construct confidence intervals for the common variance of normal distributions and compares the results based on the generalized confidence intervals approach (GCI), large sample approach, adjusted method of variance estimates recovery 1 approach (adjusted MOVER 1) and adjusted method of variance estimates recovery 2 approach (adjusted MOVER 2). A Monte Carlo simulation is used to evaluate the coverage probability and average length of confidence intervals. Simulation studies show that the adjusted MOVER 1 approach provided the best confidence interval estimates. Two real data examples are exhibited to illustrate our approaches.

Received: April 10, 2017; Accepted: June 15, 2017

2010 Mathematics Subject Classification: 62F25.

Keywords and phrases: common variance, normal distribution, generalized confidence interval, large sample approach, adjusted method of variance estimates recovery.

1. Introduction

The construction of confidence intervals for a normal variance is well known and simple to apply which attracted a great deal of attention from researchers. An investigation of the history and development of constructing confidence intervals for a normal variance is given in Cohen [1]. He constructed confidence intervals for the variance that had the same length as the usual minimum length interval but greater coverage probability. Analogously, Shorrock [2, 3] presented an improved interval based on Stein's technique and a smooth version of Cohen's interval using Brewster and Zidek's technique. Stein-type improvements of confidence intervals for the normal variance with unknown mean were also obtained by Nagata [4]. Goutis and Casella [5] constructed a class of intervals each of which improved both coverage probability and size over the usual interval. Lastly, Kubokawa [6] presented a unified approach to the variance estimation problem. There are many researchers who are also interested in the estimation of variance; see, e.g., Shorrock and Zidek [7]. Sarkar [8] constructed the shortest confidence interval and Iliopoulos and Kourouklis [9] presented a stein-type interval for generalized variances.

The motivation of this paper comes from an analysis of variance (ANOVA) which is used to compare several means. Under the assumption of analysis of variance are normality, homogeneity of variance, and independence of errors. The quantitative data of the sample n observations from k populations come from a different time or space and in experimental situations have repeated many times. In this case, if the variances are homogeneous, what is the best way for construction the confidence interval estimation of common variance to obtain a single estimation? Therefore, interval estimation procedures regarding common variance of normal distributions are interesting.

The practical and theoretical developing procedures for interval estimation of common variance based on several independent normal samples are important. Thus, the goal of this paper is to provide four approaches for the confidence interval estimation of common variance derived from several independent samples from normal distributions. The generalized confidence interval (the GCI), the large sample, the adjusted method of variance estimates recovery 1 (the adjusted MOVER 1) and the adjusted method of variance estimates recovery 2 (the adjusted MOVER 2) confidence interval concept are used for the end evaluation. The GCI approach is based on the concepts of generalized confidence intervals. The notions of generalized confidence intervals were proposed by Weerahandi [10]. The GCI approach has been successfully used to construct the confidence interval for many common parameters and since then these ideas have been applied to solve many statistical problems, for examples, Tian [11], Tian and Wu [12], Krishnamoorthy and Lu [13], and Ye et al. [14]. The adjusted MOVER 1 approach and the adjusted MOVER 2 approach are applied from the method of variance of estimates recovery (MOVER) and the large sample approach. Zou and Donner [15] introduced a detailed outline of the method of variance estimates recovery. The main idea was to recover the variance estimates needed for interval estimation obtained from the limits for parameter separately. Several researchers have used the MOVER approach to construct confidence intervals in previous publications; e.g., see Donner and Zou [16], Suwan and Niwitpong [17], Li et al. [18] and Wongkhao [19]. However, the adjusted MOVER 1 approach and the adjusted MOVER 2 approach used to construct these confidence interval estimations for the common variance are also interesting. To our knowledge, there is no previous work on inferences on common variance referring to normal distributions with the adjusted MOVER 1 approach and the adjusted MOVER 2 approach.

The remainder of the paper is organized as follows: Section 2 presents four approaches developed. Section 3 presents simulation results to evaluate the performances of the GCI approach, the large sample approach, the adjusted MOVER 1 approach and the adjusted MOVER 2 approach on coverage probabilities and average lengths. Section 4 illustrates the proposed approaches with real examples. Finally, conclusions are given in Section 5.

2. The Confidence Interval Approaches

2.1. The generalized confidence intervals approach

The generalized confidence intervals (GCI) are based on the simulation of a known generalized pivotal quantity (GPQ). Weerahandi [10] introduced the concept of a generalized pivotal quantity for a parameter θ as follows:

Suppose that $X_{ij} \sim N(\mu_i, \sigma_i^2)$, for $i = 1, ..., k$, $j = 1, ..., n_i$ are random samples from a distribution which depends on a vector of parameters $\underset{\sim}{\theta} = (\theta, \underset{\sim}{\nu})$, where θ is the parameter of interest and $\underset{\sim}{\nu}$ is a vector of nuisance parameters. A generalized pivot $R(\underset{\sim}{X}, \underset{\sim}{x}, \theta, \underset{\sim}{\nu})$ for interval estimation, where $\underset{\sim}{x}$ is an observed value of $\underset{\sim}{X}$, as a random variable having the following two properties:

(1) $R(\underset{\sim}{X}, \underset{\sim}{x}, \theta, \underset{\sim}{\nu})$ has a distribution free of the vector of nuisance parameters $\underset{\sim}{\nu}$.

(2) The observed value of $R(\underset{\sim}{X}, \underset{\sim}{x}, \theta, \underset{\sim}{\nu})$ is θ.

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

Generalized variable approach

Consider k independent normal populations with a common variance θ. Let $X_{i1}, X_{i2}, ..., X_{in_i}$ be a random sample from the ith normal population as follows:

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

Thus,

$$\theta = \sigma_i^2.$$

Let S_i^2 denote the sample variance for data X_{ij} for the ith sample and let s_i^2 denote the observed sample variance, respectively. From

$$\frac{(n_i - 1)S_i^2}{\sigma_i^2} = V_i \sim \chi^2_{n_i - 1},$$

so

$$\sigma_i^2 = \frac{(n_i - 1)S_i^2}{V_i}, \text{ where } V_i \sim \chi^2_{n_i - 1},$$

where V_i is χ^2 which variates with degree of freedom and $n_i - 1,$ we have the generalized pivot

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

The generalized pivotal quantity for estimating θ based on the ith sample is

$$R_\theta^{(i)} = R_{\sigma_i^2}. \tag{2}$$

From the ith sample, the maximum likelihood estimator of θ is

$$\hat{\theta}^{(i)} = \hat{\sigma}_i^2, \text{ where } \hat{\sigma}_i^2 = s_i^2. \tag{3}$$

The large sample variance for $\hat{\theta}^{(i)}$ is

$$\text{var}(\hat{\theta}^{(i)}) = \text{var}(\hat{\sigma}_i^2) = \text{var}(s_i^2) = \frac{2\sigma_i^4}{n_i - 1}. \tag{4}$$

The generalized pivotal quantity proposed for the common variance θ is a weighted average of the generalized pivot $R_\theta^{(i)}$ based on k individual samples as; see Ye et al. [14]:

$$R_\theta = \frac{\sum_{i=1}^{k} R_w R_\theta^{(i)}}{\sum_{i=1}^{k} R_{w_i}}, \tag{5}$$

where

$$R_{w_i} = \frac{1}{R_{\text{var}(\hat{\theta}^{(i)})}}, \tag{6}$$

$$R_{\text{var}(\hat{\theta}^{(i)})} = \frac{2(R_{\sigma_i^2})^2}{n_i - 1}. \tag{7}$$

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

2.2. The large sample approach

The large sample estimate of normal variance is a pooled estimate of the common normal variance defined as, see Tian [11],

$$\hat{\theta} = \frac{\sum_{i=1}^{k} \frac{\hat{\theta}^{(i)}}{\text{var}(\hat{\theta}^{(i)})}}{\sum_{i=1}^{k} \frac{1}{\text{var}(\hat{\theta}^{(i)})}}, \tag{8}$$

where $\hat{\theta}^{(i)}$ is defined in (3) and $\text{var}(\hat{\theta}^{(i)})$ is an estimate of $\text{var}(\hat{\theta}^{(i)})$ in (4) with σ_i^2 replaced by s_i^2, respectively.

Hence, the large sample solution for confidence interval estimation is

$$\left(\hat{\theta} - z_{1-\alpha/2} \sqrt{\frac{1}{\sum_{i=1}^{k} \frac{1}{\text{var}(\hat{\theta}^{(i)})}}},\ \hat{\theta} + z_{1-\alpha/2} \sqrt{\frac{1}{\sum_{i=1}^{k} \frac{1}{\text{var}(\hat{\theta}^{(i)})}}} \right). \tag{9}$$

2.3. The adjusted method of variance estimates recovery 1 approach

The adjusted method of variance estimates recovery 1 (the adjusted MOVER 1 approach) uses the concepts of the method of variance estimates recovery (the MOVER approach) and the large sample approach.

Zou and Donner [15] introduced the concept of the method of variance estimates recovery (the MOVER approach). The main idea was to recover the variance estimates needed for interval estimation obtained from the limits for parameter separately. This method considers two parameters $\theta_1 + \theta_2$ which have $100(1-\alpha)\%$ confidence limits (l_1, u_1) and (l_2, u_2), respectively. Under the assumption of independence between the point estimates $\hat{\theta}_1$ and $\hat{\theta}_2$ and the application of the central limit theorem, the lower limit L and the upper limit U are given by

$$[L, U] = (\hat{\theta}_1 + \hat{\theta}_2) \pm z_{\alpha/2}\sqrt{\hat{\text{var}}(\hat{\theta}_1) + \hat{\text{var}}(\hat{\theta}_2)}, \tag{10}$$

where the variance terms may be estimated near the upper and lower confidence limits of $\theta_1 + \theta_2$ using the information already available in the confidence intervals of the individual parameters θ_i, where $i = 1, 2$. The Slutsky's theorem and the central limit theorem $(z \sim (\hat{\theta}_i - \theta_i)/\sqrt{\text{var}(\hat{\theta}_i)})$ are applied to estimates variances near the lower $\hat{\text{var}}(\hat{\theta}_i)$ and upper $\hat{\text{var}}(\hat{\theta}_i)$ limits of θ_i as

$$\hat{\text{var}}(\hat{\theta}_i) = \frac{(\hat{\theta}_i - l_i)^2}{z_{\alpha/2}^2}, \quad \hat{\text{var}}(\hat{\theta}_i) = \frac{(u_i - \hat{\theta}_i)^2}{z_{\alpha/2}^2}.$$

Using these estimates with from (10), two-side $100(1-\alpha)\%$ confidence limits for $\theta_1 + \theta_2$ given as

$$L = (\hat{\theta}_1 + \hat{\theta}_2) - \sqrt{(\hat{\theta}_1 - l_1)^2 + (\hat{\theta}_2 - l_2)^2},$$

$$U = (\hat{\theta}_1 + \hat{\theta}_2) + \sqrt{(u_1 - \hat{\theta}_1)^2 + (u_2 - \hat{\theta}_2)^2}.$$

Let $\theta^{(1)}, \theta^{(2)}, ..., \theta^{(k)}$ be k parameters of interest, where the estimates $\hat{\theta}^{(1)}, \hat{\theta}^{(2)}, ..., \hat{\theta}^{(k)}$ are independent. Construction of a $100(1-\alpha)\%$ two-sided confidence interval (L, U) for $\theta^{(1)}, \theta^{(2)}, ..., \theta^{(k)}$ is of interest. The concept of Donner and Zou [16] was used to find the confidence interval for $\hat{\theta}^{(1)} + \hat{\theta}^{(2)} + \cdots + \hat{\theta}^{(k)}$, which is the estimator for $\theta^{(1)} + \theta^{(2)} + \cdots + \theta^{(k)}$.

Thus,

$$[L, U] = (\hat{\theta}^{(1)} + \cdots + \hat{\theta}^{(k)}) \pm z_{\alpha/2}\sqrt{\hat{\text{var}}(\hat{\theta}^{(1)}) + \cdots + \hat{\text{var}}(\hat{\theta}^{(k)})},$$

where $z_{\alpha/2}$ is the upper $\alpha/2$ quintile of the standard normal distribution.

Suppose the $100(1-\alpha)\%$ two-sided confidence interval for $\theta^{(i)}$ is given by (l_i, u_i), where $i = 1, 2, ..., k$. The lower limit L is in the neighborhood of $l_1 + \cdots + l_k$. The spirit of the score-type confidence interval and the central limit theorem were used to estimate variances at $\theta^{(i)} = l_i$.

Thus, the variance estimate for $\hat{\theta}^{(i)}$ at $\theta^{(i)} = l_i$ is equal to

$$\hat{\text{var}}(\hat{\theta}^{(i)}) = \frac{(\hat{\theta}^{(i)} - l_i)^2}{z_{\alpha/2}^2}.$$

Analogous steps with the notion that $u_1 + u_2 + \cdots + u_k$ is close to upper limit U, and the variance estimate at $\theta^{(i)} = u_i$ is

$$\hat{\text{var}}(\hat{\theta}^{(i)}) = \frac{(u_i - \hat{\theta}^{(i)})^2}{z_{\alpha/2}^2}.$$

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

$$L = \hat{\theta}^{(1)} + \cdots + \hat{\theta}^{(k)} - \sqrt{(\hat{\theta}^{(1)} - l_1)^2 + \cdots + (\hat{\theta}^{(k)} - l_k)^2},$$

$$U = \hat{\theta}^{(1)} + \cdots + \hat{\theta}^{(k)} + \sqrt{(u_1 - \hat{\theta}^{(1)})^2 + \cdots + (u_k - \hat{\theta}^{(k)})^2}.$$

From the ith sample, where $i = 1, 2, ..., k$. The maximum likelihood estimator of common variance θ is

$$\hat{\theta}^{(i)} = s_i^2.$$

The method of large sample for normal variance with pooled estimate was used. The common normal variance θ is weighted average of variance $\hat{\theta}^{(i)}$ based on k individual samples as

$$\hat{\theta} = \sum_{i=1}^{k} \frac{\hat{\theta}^{(i)}}{\operatorname{var}(\hat{\theta}^{(i)})} \Bigg/ \sum_{i=1}^{k} \frac{1}{\operatorname{var}(\hat{\theta}^{(i)})} \tag{11}$$

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

$$\operatorname{var}(\hat{\theta}^{(i)}) = \frac{1}{2}\left(\frac{(\hat{\theta}^{(i)} - l_i)^2}{z_{\alpha/2}^2} + \frac{(u_i - \hat{\theta}^{(i)})^2}{z_{\alpha/2}^2} \right). \tag{12}$$

We have

$$l_i = \frac{(n_i - 1)s_i^2}{V}, \text{ where } i = 1, 2, ..., k, V \sim \chi^2_{1-\alpha/2}, \tag{13}$$

$$u_i = \frac{(n_i - 1)s_i^2}{U}, \text{ where } i = 1, 2, ..., k, U \sim \chi^2_{\alpha/2}. \tag{14}$$

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

$$L = \hat{\theta} - z_{1-\alpha/2} \sqrt{1 \Bigg/ \sum_{i=1}^{k} 1 \Bigg/ \frac{(\hat{\theta}^{(i)} - l_i)^2}{z_{\alpha/2}^2}}.$$

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

$$U = \hat{\theta} + z_{1-\alpha/2} \sqrt{1 \Bigg/ \sum_{i=1}^{k} 1 \Bigg/ \frac{(u_i - \hat{\theta}^{(i)})^2}{z_{\alpha/2}^2}}.$$

Hence, the adjusted MOVER 1 solution for confidence interval estimation is

$$\left(\hat{\theta} - z_{1-\alpha/2}\sqrt{\frac{1}{\sum_{i=1}^{k}\frac{z_{\alpha/2}^2}{(\hat{\theta}^{(i)} - l_i)^2}}},\ \hat{\theta} + z_{1-\alpha/2}\sqrt{\frac{1}{\sum_{i=1}^{k}\frac{z_{\alpha/2}^2}{(u_i - \hat{\theta}^{(i)})^2}}}\right). \tag{15}$$

2.4. The adjusted method of variance estimates recovery 2 approach

The adjusted method of variance estimates recovery 2 (the adjusted MOVER 2 approach) uses the concepts same as the adjusted MOVER 1 approach.

From the ith sample, we are of interest to construct a $100(1-\alpha)\%$ two-sided confidence interval (L, U) for common variance $\theta = \sigma_i^2$.

Let $\theta^{(i)} = \theta^{(1)}, \theta^{(2)}, ..., \theta^{(k)} = \sigma_i^2,\ i = 1, 2, ..., k,$ where the estimates $\hat{\theta}^{(1)}, \hat{\theta}^{(2)}, ..., \hat{\theta}^{(k)}$ are independent. The common normal variance θ is weighted average of variance $\hat{\theta}^{(i)}$ based on k individual samples in (11). Therefore, the lower limit L and upper limit U for the common normal variance θ are given by

$$L = \hat{\theta} - \sqrt{(\hat{\theta} - l)^2}, \quad U = \hat{\theta} + \sqrt{(u - \hat{\theta})^2}.$$

The maximum likelihood estimator of common variance θ is $\hat{\theta}^{(i)} = s_i^2$. Thus $\theta^{(i)} = \sigma_i^2,\ i = 1, 2, ..., k$ is contained in confidence interval (l_i, u_i) in (13) and (14). We have

$$\mathrm{var}(l_i) = \frac{2s_i^4(n_i - 1)}{V^2}, \text{ where } i = 1, 2, ..., k, V \sim \chi^2_{1-\alpha/2}, \tag{16}$$

$$\mathrm{var}(u_i) = \frac{2s_i^4(n_i - 1)}{U^2}, \text{ where } i = 1, 2, ..., k, U \sim \chi^2_{\alpha/2}. \tag{17}$$

The common l, u which is a weighted average of l_i, u_i based on k individual samples defined as

$$l = \sum_{i=1}^{k} \frac{l_i}{\mathrm{var}(l_i)} \Bigg/ \sum_{i=1}^{k} \frac{1}{\mathrm{var}(l_i)}, \tag{18}$$

$$u = \sum_{i=1}^{k} \frac{u_i}{\mathrm{var}(u_i)} \Bigg/ \sum_{i=1}^{k} \frac{1}{\mathrm{var}(u_i)}, \tag{19}$$

where l_i, u_i are defined in (13), (14) and $\mathrm{var}(l_i)$, $\mathrm{var}(u_i)$ are defined in (16), (17), respectively.

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

$$L = \hat{\theta} - \sqrt{(\hat{\theta} - l)^2}.$$

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

$$U = \hat{\theta} + \sqrt{(u - \hat{\theta})^2}.$$

Hence, the adjusted MOVER 2 solution for confidence interval estimation is

$$\left(\hat{\theta} - \sqrt{(\hat{\theta} - l)^2},\ \hat{\theta} + \sqrt{(u - \hat{\theta})^2}\right). \tag{20}$$

3. Simulation Studies

A simulation study was performed with the coverage probabilities and average lengths of the common variance of the normal distributions for various combinations of the number of samples $k = 2$, $k = 4$ and $k = 6$, the sample sizes $n_1 = \cdots = n_k = n$, the values used for sample sizes were 10, 30, 50, 100 and 200 the population mean of normal data within each sample $\mu_1 = \cdots = \mu_{[k/2]}$, $\mu_{[k/2]+1} = \cdots = \mu_k$, $\mu_1 / \mu_{[k/2]+1} = 1$ and $5/3$, and the population variance $\sigma^2 = 1, 2$ and 4. In this simulation study, four methods were compared, comprising of our proposed procedure the generalized confidence interval approach (the GCI approach), the large sample approach, the adjusted method of variance estimates recovery 1 approach (the adjusted MOVER 1 approach) and the adjusted method of variance estimates recovery 2 approach (the adjusted MOVER 2 approach). For each parameter setting, 5000 random samples were generated, 2500 R_θ's were obtained for each of the random samples.

Tables 1-3 present the coverage probabilities and average lengths for 2, 4 and 6 sample cases, respectively. In 2, 4 and 6 sample cases, the adjusted MOVER 2 approach overestimated the coverage probabilities for all of the scenarios. The GCI approach, the large sample approach and the adjusted MOVER 1 approach provide the underestimates coverage probabilities for most of the scenarios, especially when the sample size is small. Additionally, the coverage probabilities of the adjusted MOVER 1 approach are better than the GCI approach and the large sample approach for all sample sizes, especially when k increases and the sample size is small.

In this case, there is no need to see the average lengths from four intervals since the approach provides the coverage probability below another approach for all cases and needs the shortest average length. Finally, it was discovered that the adjusted MOVER 2 approach is a conservative confidence interval leading to the large length. Overall, the adjusted MOVER 1 approach has the coverage probabilities close to the nominal confidence level at 0.95 when the sample size increases and the average length is the shortest.

Table 1. Empirical coverage probabilities (CP) and average length (AL) of approximate 95% two-side confidence bounds for common variance (based on 5000 simulations): 2 sample cases

n	Ratio*	σ^2	GCI approach		Large sample approach		The adjusted MOVER 1 approach		The adjusted MOVER 2 approach	
			CP	AL	CP	AL	CP	AL	CP	AL
10	1	1	0.9228	1.4689	0.7554	1.1303	0.9264	1.7492	0.9678	2.3886
		2	0.9222	2.9377	0.7494	2.2615	0.9228	3.4999	0.9694	4.7801
		4	0.9256	5.9178	0.7524	4.5433	0.9260	7.0312	0.9708	9.5918
	5/3	1	0.9266	1.4884	0.7668	1.1457	0.9270	1.7731	0.9710	2.4215
		2	0.9194	2.9265	0.7440	2.2468	0.9270	3.4771	0.9682	4.7469
		4	0.9222	5.8772	0.7536	4.5216	0.9256	6.9976	0.9714	9.5574
30	1	1	0.9340	0.7777	0.8580	0.6926	0.9310	0.7891	0.9854	1.1000
		2	0.9394	1.5610	0.8718	1.3918	0.9342	1.5858	0.9832	2.2109
		4	0.9386	3.1022	0.8694	2.7679	0.9340	3.1538	0.9834	4.3976
	5/3	1	0.9378	0.7771	0.8650	0.6919	0.9404	0.7884	0.9860	1.0989
		2	0.9414	1.5563	0.8684	1.3853	0.9370	1.5784	0.9854	2.2002
		4	0.9328	3.1059	0.8610	2.7622	0.9302	3.1473	0.9848	4.3851
50	1	1	0.9408	0.5900	0.8942	0.5443	0.9384	0.5876	0.9866	0.8233
		2	0.9456	1.1786	0.8994	1.0874	0.9446	1.1741	0.9894	1.6452
		4	0.9430	2.3589	0.8998	2.1781	0.9396	2.3517	0.9876	3.2958
	5/3	1	0.9430	0.5903	0.9028	0.5453	0.9404	0.5888	0.9884	0.8252
		2	0.9416	1.1811	0.8974	1.0898	0.9394	1.1767	0.9880	1.6490
		4	0.9430	2.3510	0.9002	2.1724	0.9454	2.3455	0.9890	3.2876
100	1	1	0.9404	0.4054	0.9140	0.3869	0.9424	0.4017	0.9920	0.5655
		2	0.9400	0.8134	0.9138	0.7759	0.9390	0.8058	0.9914	1.1340
		4	0.9434	1.6260	0.9168	1.5501	0.9438	1.6097	0.9936	2.2654
	5/3	1	0.9440	0.4075	0.9196	0.3888	0.9406	0.4038	0.9898	0.5683
		2	0.9478	0.8146	0.9228	0.7766	0.9456	0.8065	0.9930	1.1350
		4	0.9410	1.6214	0.9160	1.5487	0.9428	1.6083	0.9904	2.2639
200	1	1	0.9486	0.2826	0.9356	0.2753	0.9486	0.2805	0.9938	0.3957
		2	0.9494	0.5653	0.9394	0.5507	0.9478	0.5611	0.9934	0.7915
		4	0.9512	1.1334	0.9428	1.1037	0.9518	1.1245	0.9944	1.5864
	5/3	1	0.9408	0.2832	0.9312	0.2759	0.9410	0.2811	0.9918	0.3966
		2	0.9414	0.5661	0.9270	0.5513	0.9434	0.5617	0.9888	0.7924
		4	0.9466	1.1331	0.9354	1.1027	0.9452	1.1235	0.9914	1.5850

*Ratio is defined as μ_1/μ_2

Table 2. Empirical coverage probabilities (CP) and average length (AL) of approximate 95% two-side confidence bounds for common variance (based on 5000 simulations): 4 sample cases

n	Ratio*	σ^2	GCI approach		Large sample approach		The adjusted MOVER 1 approach		The adjusted MOVER 2 approach	
			CP	AL	CP	AL	CP	AL	CP	AL
10	1	1	0.7838	0.9199	0.5872	0.7256	0.8884	1.1230	0.9722	2.0996
		2	0.7824	1.8455	0.5812	1.4564	0.8908	2.2540	0.9752	4.2168
		4	0.7822	3.6864	0.5880	2.9101	0.8874	4.5037	0.9690	8.4260
	5/3	1	0.7828	0.9209	0.5884	0.7276	0.8894	1.1261	0.9716	2.1073
		2	0.7842	1.8400	0.5886	1.4493	0.8880	2.2429	0.9692	4.1929
		4	0.7970	3.7223	0.6080	2.9383	0.8978	4.5474	0.9740	8.5095
30	1	1	0.8748	0.5483	0.7962	0.4769	0.9198	0.5434	0.9948	1.0604
		2	0.8690	1.0961	0.7936	0.9552	0.9096	1.0884	0.9944	2.1247
		4	0.8690	2.1921	0.7924	1.9102	0.9164	2.1765	0.9968	4.2491
	5/3	1	0.8700	0.5469	0.7926	0.4764	0.9214	0.5428	0.9946	1.0597
		2	0.8660	1.0914	0.7826	0.9517	0.9120	1.0844	0.9956	2.1172
		4	0.8628	2.1836	0.7818	1.9016	0.9134	2.1667	0.9968	4.2292
50	1	1	0.8888	0.4178	0.8412	0.3781	0.9204	0.4082	0.9976	0.8044
		2	0.8832	0.8330	0.8388	0.7548	0.9256	0.8149	0.9980	1.6063
		4	0.8956	1.6748	0.8514	1.5133	0.9280	1.6339	0.9988	3.2188
	5/3	1	0.8902	0.4173	0.8454	0.3775	0.9266	0.4076	0.9984	0.8033
		2	0.8882	0.8357	0.8428	0.7568	0.9212	0.8172	0.9980	1.6105
		4	0.8886	1.6712	0.8404	1.5105	0.9234	1.6309	0.9986	3.2132
100	1	1	0.9122	0.2886	0.8894	0.2721	0.9360	0.2826	0.9996	0.5609
		2	0.9256	0.5774	0.9042	0.5449	0.9472	0.5659	0.9996	1.1235
		4	0.9164	1.1538	0.8972	1.0894	0.9386	1.1313	0.9990	2.2461
	5/3	1	0.9152	0.2885	0.8944	0.2722	0.9350	0.2826	0.9996	0.5611
		2	0.9180	0.5772	0.8996	0.5448	0.9394	0.5658	1.0000	1.1234
		4	0.9136	1.1556	0.8938	1.0886	0.9366	1.1305	0.9998	2.2439
200	1	1	0.9314	0.2002	0.9208	0.1939	0.9466	0.1976	1.0000	0.3938
		2	0.9348	0.4009	0.9228	0.3882	0.9444	0.3956	1.0000	0.7882
		4	0.9348	0.8025	0.9270	0.7777	0.9452	0.7924	1.0000	1.5790
	5/3	1	0.9298	0.2004	0.9226	0.1941	0.9438	0.1978	1.0000	0.3941
		2	0.9380	0.4011	0.9272	0.3885	0.9472	0.3958	1.0000	0.7888
		4	0.9302	0.8015	0.9240	0.7770	0.9440	0.7917	0.9998	1.5778

*We set $\mu_1 = \mu_2$, $\mu_3 = \mu_4$ and ratio is defined as μ_1/μ_3

Table 3. Empirical coverage probabilities (CP) and average length (AL) of approximate 95% two-side confidence bounds for common variance (based on 5000 simulations): 6 sample cases

n	Ratio*	σ^2	GCI approach		Large sample approach		The adjusted MOVER 1 approach		The adjusted MOVER 2 approach	
			CP	AL	CP	AL	CP	AL	CP	AL
10	1	1	0.6252	0.7460	0.4604	0.5723	0.8494	0.8858	0.9740	1.9932
		2	0.6028	1.4719	0.4448	1.1317	0.8384	1.7514	0.9726	3.9406
		4	0.6140	2.9531	0.4554	2.2731	0.8426	3.5179	0.9734	7.9216
	5/3	1	0.6202	0.7423	0.4612	0.5699	0.8388	0.8820	0.9750	1.9838
		2	0.6152	1.4826	0.4632	1.1424	0.8464	1.7680	0.9778	3.9829
		4	0.6226	2.9641	0.4704	2.2884	0.8510	3.5415	0.9746	7.9842
30	1	1	0.7766	0.4509	0.7196	0.3845	0.8904	0.4381	0.9976	1.0430
		2	0.7694	0.9022	0.7152	0.7692	0.8868	0.8764	0.9976	2.0862
		4	0.7820	1.8043	0.7232	1.5419	0.8950	1.7568	0.9976	4.1837
	5/3	1	0.7770	0.4507	0.7166	0.3842	0.8884	0.4377	0.9978	1.0421
		2	0.7762	0.9011	0.7192	0.7693	0.8912	0.8765	0.9980	2.0874
		4	0.7696	1.8021	0.7140	1.5387	0.8842	1.7532	0.9978	4.1744
50	1	1	0.8398	0.3437	0.8108	0.3077	0.9172	0.3322	0.9998	0.8004
		2	0.8342	0.6880	0.8086	0.6149	0.9164	0.6639	0.9992	1.5992
		4	0.8364	1.3753	0.8096	1.2281	0.9166	1.3260	0.9994	3.1933
	5/3	1	0.8280	0.3432	0.8016	0.3068	0.9086	0.3313	0.9996	0.7980
		2	0.8320	0.6862	0.8022	0.6138	0.9068	0.6628	0.9994	1.5965
		4	0.8354	1.3761	0.8016	1.2300	0.9096	1.3280	0.9996	3.1983
100	1	1	0.8866	0.2365	0.8756	0.2218	0.9282	0.2303	0.9998	0.5596
		2	0.8880	0.4733	0.8808	0.4436	0.9330	0.4607	1.0000	1.1190
		4	0.8956	0.9476	0.8882	0.8882	0.9344	0.9224	1.0000	2.2402
	5/3	1	0.8770	0.2363	0.8668	0.2215	0.9288	0.2300	0.9998	0.5587
		2	0.8836	0.4732	0.8736	0.4436	0.9276	0.4606	1.0000	1.1189
		4	0.8852	0.9454	0.8792	0.8870	0.9274	0.9212	1.0000	2.2377
200	1	1	0.9166	0.1641	0.9150	0.1585	0.9402	0.1614	1.0000	0.3939
		2	0.9180	0.3278	0.9166	0.3167	0.9446	0.3227	1.0000	0.7872
		4	0.9088	0.6554	0.9066	0.6332	0.9346	0.6452	1.0000	1.5739
	5/3	1	0.9136	0.1638	0.9108	0.1583	0.9370	0.1612	1.0000	0.3934
		2	0.9126	0.3279	0.9120	0.3169	0.9390	0.3229	1.0000	0.7876
		4	0.9112	0.6554	0.9090	0.6338	0.9414	0.6458	1.0000	1.5754

*We set $\mu_1 = \mu_2 = \mu_3$, $\mu_4 = \mu_5 = \mu_6$ and ratio is defined as μ_1/μ_4

4. An Empirical Application

In this section, two real data examples are exhibited to illustrate the GCI approach, the large sample approach, the adjusted MOVER 1 approach and the adjusted MOVER 2 approach. The first data set compares two different procedures for the shear strength for steel plate girders. Data for nine girders for two of these procedures, Karlsruhe method and Lehigh method [20], testing the hypothesis of equal mean treatment effects. Under the assumption are normality, homogeneity of variance, and independence of errors. The Shapiro-Wilk normality test indicates that the two sets of data come from normal populations and the variances were homogeneous by Levene’s test. The sample variances of the normal data were 0.0213 and 0.0024 for Karlsruhe method and Lehigh method, respectively. Using the GCI approach, the generalized confidence interval for the overall variance was (0.0012, 0.0104) with the length of interval 0.0092. The confidence interval by the large sample approach was (0.0003, 0.0050) with the length of interval 0.0047. In comparison, the confidence interval by the adjusted MOVER 1 approach was (0.0014, 0.0092) with the length of interval 0.0078 and the confidence interval by the adjusted MOVER 2 approach was (0.0012, 0.0098) with the length of interval 0.0086.

The second example was blood sugar levels (mg/100g) measured from ten animals of five different breeds [21] for testing the hypothesis of equality of means for the five breeds. The data on the five set were tested from normal populations by Shapiro-Wilk normality test and the variances were homogeneous by Levene’s test. The sample variances of the normal data were 84.0000, 124.6667, 126.5444, 101.1111 and 173.1667 for breeds A, B, C, D and E, respectively. Using the GCI approach, the generalized confidence interval for the overall variance was (56.9113, 156.9829) with the length of interval 100.0716. The confidence interval by the large sample approach was (62.6062, 155.0373) with the length of interval 92.43106. In comparison, the confidence interval by the adjusted MOVER 1 approach was

(82.4669, 225.5117) with the length of interval 143.0447 and the confidence interval by the adjusted MOVER 2 approach was (51.4854, 362.6868) with the length of interval 311.2013. Hence, the results from above two examples show that the adjusted MOVER 1 approach is shorter than the adjusted MOVER 2 which supports the simulation results.

5. Discussion and Conclusions

This paper has presented a simple approach to construct confidence intervals for the common variance of normal distributions. The proposed confidence intervals were constructed by four approaches, the GCI approach, the large sample approach, the adjusted MOVER 1 approach and the adjusted MOVER 2 approach. The adjusted MOVER 2 approach provided coverage probabilities more than the confidence level at 0.95 and is better than the adjusted MOVER 1 approach, the GCI approach and the large sample approach for all sample sizes. For sample sizes are large, the adjusted MOVER 1 approach provided coverage probabilities close to nominal level 0.95 and is better than the GCI approach and the large sample approach.

The average lengths increased when the value of σ^2 increased for four approaches. In this study, the researchers need an approach to provide coverage probability close to nominal level 0.95 and have the shortest average length. The results indicated that the confidence interval for the common variance of normal distributions based on the adjusted MOVER 1 approach outperforms other approaches.

Acknowledgments

The first author gratefully acknowledges the financial support from Rajamangala University of Technology Phra Nakhon of Thailand. The research is also funded by the Faculty of Applied Science, King Mongkut's University of Technology North Bangkok. Contract Number 6045105.

References

[1] C. A. Cohen, Improved confidence intervals for the variance of a normal distribution, J. Amer. Statist. Assoc. 67 (1972), 382-387.

[2] G. Shorrock, A minimax generalized Bayes confidence interval for a normal variance, Ph.D. Dissertation, Dept. Statistics, Rutgers Univ., 1982.

[3] G. Shorrock, Improved confidence intervals for a normal variance, Ann. Statist. 18 (1990), 972-980.

[4] Y. Nagata, Improvements of interval estimations for the variance and the ratio of two variances, J. Japan Statist. Soc. 19 (1989), 151-161.

[5] C. Goutis and G. Casella, Improved invariant confidence intervals for a normal variance, Ann. Statist. 19 (1991), 2015-2031.

[6] T. Kubokawa, A unified approach to improving equivariant estimators, Ann. Statist. 22 (1994), 290-299.

[7] R. W. Shorrock and J. V. Zidek, An improved estimator of the generalized variance, Ann Statist. 4(3) (1976), 629-638.

[8] S. K. Sarkar, On improving the shortest length confidence interval for the generalized variance, J. Multivariate Anal. 31 (1989), 136-147.

[9] G. Iliopoulos and S. Kourouklis, On improved interval estimation for the generalized variance, J. Stat. Plan. Inference 66 (1998), 305-320.

[10] S. Weerahandi, Generalized confidence intervals, J. Amer. Statist. Assoc. 88 (1993), 899-905.

[11] L. Tian, Inferences on the common coefficient of variation, Statistics in Medicine 24 (2005), 2213-2220.

[12] L. Tian and J. Wu, Inferences on the common mean of several log-normal populations: the generalized variable approach, Biom. J. 49 (2007), 944-951.

[13] K. Krishnamoorthy and Y. Lu, Inference on the common means of several normal populations based on the generalized variable method, Biometrics 59 (2003), 237-247.

[14] R. D. Ye, T. F. Ma and S. G. Wang, Inferences on the common mean of several inverse Gaussian populations, Comput. Statist. Data Anal. 54 (2010), 906-915.

[15] G. Y. Zou and A. Donner, Construction of confidence limits about effect measures: a general approach, Stat. Med. 27 (2008), 1693-1702.

[16] A. Donner and G. Y. Zou, Closed-form confidence intervals for function of the normal standard deviation, Stat. Methods Med. Res. (2010), 86-89.

[17] S. Suwan and S. Niwitpong, Estimated variance ratio confidence interval of non-normal distributions, Far East J. Math. Sci. (FJMS) Special Volume, Issue 4 (2013), 339-350.

[18] H. Q. Li, M. L. Tang and W. K. Wong, Confidence intervals for ratio of two Poisson rates using the method of variance estimates recovery, Comput. Statist. 29 (2014), 869-889.

[19] A. Wongkhao, Confidence intervals for parameters of normal distribution, Ph.D. thesis, King Mongkut's University of Technology, North Bangkok, 2014.

[20] C. Douglas Montgomery, Design and Analysis of Experiments, John Wiley & Sons, New York, 2001, 57 pp.

[21] Alvin C. Rencher and G. Bruce Schaalje, Linear Models in Statistics, John Wiley and Sons, Inc., New Jersey, 2008, 373 pp.