Confidence Intervals for Common Variance of Normal Distributions

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

Applied Statistics, King Mongkut's University of Technology North Bangkok,
Bangkok 10800, Thailand
Nsmithpreecha@gmail.com, {sa-aat.n,suparat.n}@sci.kmutnb.ac.th

Abstract. This paper presents a construction of confidence intervals for the common variance of normal distributions based on generalized confidence intervals, and then compares the results with a large sample approach. A Monte Carlo simulation was used to evaluate the coverage probability and average length of confidence intervals. Simulation studies showed that the generalized confidence interval approach provided much better confidence interval estimates than the large sample approach. Two real data examples are exhibited to illustrate our approaches.

Keywords: Common variance · Generalized confidence interval · Large sample approach · Normal distribution

1 Introduction

The construction of confidence intervals for a normal variance are well known and simple to apply and attracted a great deal of attention from researchers. An investigation of the history and development of constructing confidence intervals for a normal variance was given in Cohen [1], and 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]. 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 that 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 are used to compare several means. Under the assumption of analysis of variance are normality, homogeneity of variance, and independence of errors. If 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 of developing procedures for interval estimation of common variance based on several independent normal samples are important. Thus, the goal of this paper was to provide two approaches for the confidence interval estimation of common variance derived from several independent samples from normal distributions. The generalized confidence interval and large sample confidence interval concept will be used for the end evaluation. The approach is based on the concepts of generalized confidence intervals. The notions of generalized confidence intervals are proposed by Weerahandi [10]. The generalized confidence interval 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 example, Tian [11], Tian and Wu [12], Krishnamoorthy [13], and Ye et al. [14]. However, the generalized confidence interval approach used to construct these confidence interval estimations for the common variance are also interesting. To our knowledge, there are no previous works on inferences on common variance referring to normal distributions with a generalized confidence interval approach compared with the large sample approach.

The remainder of the paper is organized as follows. Section 2 introduces the basic properties of normal distribution. Section 3 presents the generalized variable approach developed and describes computational procedures. Section 4 presents simulation results to evaluate the performances of generalized confidence interval approach and the large sample approach on coverage probabilities and average lengths. Section 5 illustrates the proposed approaches with real examples. Finally, conclusions are given in Sect. 6.

2 Properties of Normal Distribution

If the random variable X follows the normal distribution, that is $X \sim N\left(\mu, \sigma^2\right)$. The probability density function of X is given by

$$f\left(x\right) = \frac{1}{\sigma\sqrt{2\pi}} exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), -\infty < x < \infty.$$

The maximum likelihood estimators (MLE) of μ and σ^2 are $\widehat{\mu}$ and $\widehat{\sigma}^2$ respectively,

where
$$\widehat{\mu} = \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i, \quad \widehat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2.$$

The estimator $\widehat{\sigma}^2$ is the sample variance of the sample $(X_1, X_2, ..., X_n)$. In practice, another estimator is often used instead of the $\widehat{\sigma}^2$. This other estimator is denoted S^2, and is also called the sample variance. The estimator S^2 differs from $\widehat{\sigma}^2$ by having $(n-1)$ instead of n in the denominator

Thus
$$S^2 = \frac{n}{n-1}\widehat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2.$$

The estimator S^2 is an unbiased estimator of the underlying parameter σ^2, whereas $\widehat{\sigma}^2$ is biased.

Theorem 1. *Suppose $X \sim N\left(\mu, \sigma^2\right)$, where μ, σ^2 are respectively population mean and population variance of X. Then the estimator of σ^2 is $\widehat{\sigma}^2 = S^2$, the variance of $\widehat{\sigma}^2$ is*

$$var\left(\widehat{\sigma}^2\right) = \frac{2\sigma^4}{n-1}.$$

Proof. Let $X_1, X_2, ..., X_n$ be an independent and identically distributed random variables with mean μ and variance σ^2, then $\overline{X}$ and S^2 are unbiased estimators of μ and σ^2 :

where
$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n}x_i, \qquad S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \overline{X}\right)^2.$$

Also, by the LehmannScheff theorem the estimator S^2 is uniformly minimum variance unbiased (UMVU). In finite samples both S^2 and $\widehat{\sigma}^2$ have scaled chi-squared distribution with $(n-1)$ degrees of freedom

so,
$$S^2 \sim \frac{\sigma^2}{n-1}\cdot\chi^2_{n-1}, \qquad \widehat{\sigma}^2 \sim \frac{\sigma^2}{n}\cdot\chi^2_{n-1},$$

then sampling distribution of $\dfrac{(n-1)S^2}{\sigma^2}$ is chi-square with $n-1$ degrees of freedom. For the chi-square distribution, it turns out that the mean and variance are $E\left(\chi^2_{n-1}\right) = n-1, var\left(\chi^2_{n-1}\right) = 2\left(n-1\right).$

We can use this to get the mean and variance of S^2

$$E\left(S^2\right) = E\left(\frac{\sigma^2\chi^2_{n-1}}{(n-1)}\right) = \frac{\sigma^2}{n-1}\left(n-1\right) = \sigma^2,$$

$$var\left(S^2\right) = var\left(\frac{\sigma^2\chi^2_{n-1}}{(n-1)}\right) = \frac{\sigma^4}{(n-1)^2}2\left(n-1\right) = \frac{2\sigma^4}{n-1}.$$

Hence,
$$var\left(\widehat{\sigma}^2\right) = var\left(S^2\right) = \frac{2\sigma^4}{n-1}.$$

3   The Confidence Interval Approaches of the Common Variance

3.1   The Generalized Confidence Interval 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 $X_{ij} \sim N\left(\mu_i, \sigma_i^2\right)$, for $i = 1, 2, ..., k, j = 1, 2, ..., n_i$ are a random samples from a distribution which depends on a vector of parameters $\underset{\sim}{\theta} = \left(\theta, \underset{\sim}{\nu}\right)$ where θ is the parameter of interest and $\underset{\sim}{\nu}$ is a vector of nuisance parameters. A generalized pivot $R\left(\underset{\sim}{X}, \underset{\sim}{x}, \theta, \underset{\sim}{\nu}\right)$ 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. $R\left(\underset{\sim}{X}, \underset{\sim}{x}, \theta, \underset{\sim}{\nu}\right)$ has a distribution free of the vector of nuisance parameters $\underset{\sim}{\nu}$.
2. The observed value of $R\left(\underset{\sim}{X}, \underset{\sim}{x}, \theta, \underset{\sim}{\nu}\right)$ is θ.

Let R_α be the 100α-th percentile of R. Then R_α becomes the $100\left(1 - \alpha\right)\%$ lower bound for θ and $\left(R_{\alpha/2}, R_{1-\alpha/2}\right)$ becomes a $100\left(1 - \alpha\right)\%$ 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 i-th normal population as follows:

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

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

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

$$\frac{\left(n_i - 1\right) S_i^2}{\sigma_i^2} = V_i \sim \chi_{n_i-1}^2,$$

so,
$$\sigma_i^2 = \frac{\left(n_i - 1\right) S_i^2}{V_i} \quad where\ \ V_i \sim \chi_{n_i-1}^2.$$

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

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

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

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

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

$$\widehat{\theta}^{(i)} = \widehat{\sigma}_i^2, \quad where \ \widehat{\sigma}_i^2 = S_i^2. \tag{3}$$

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

$$var\left(\widehat{\theta}^{(i)}\right) = var\left(\widehat{\sigma}_i^2\right) = var\left(S_i^2\right) = \frac{2\sigma_i^4}{n_i - 1}, \quad see \ Theorem \ 1. \tag{4}$$

The generalized pivotal quantity, we propose 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],

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

where

$$R_{w_i} = \frac{1}{R_{var\left(\widehat{\theta}^{(i)}\right)}}, \tag{6}$$

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

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

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

1. Compute $\bar{x}_i$ and s_i^2 for $i = 1, 2, \ldots, k$.
2. Generate $V_i \sim \chi_{n_i-1}^2$ and then calculate $R_{\sigma_i^2}$ from (1) for $i = 1, 2, \ldots, k$.
3. Calculate $R_\theta^{(i)}$ from (2) for $i = 1, 2, \ldots, k$.
4. Repeat steps 2, calculate R_{w_i} from (6) and (8) for $i = 1, 2, \ldots, k$.
5. Compute R_θ following (5).
6. Repeat step 2–5 a total m times and obtain an array of R_θ's.
7. Rank this array of R_θ's from small to large.

The 100α-th percentile of R_θ's, $R_\theta(\alpha)$, is an estimate of the lower bound of the one - sided $100(1-\alpha)\%$ confidence interval and $(R_\theta(\alpha/2), R_\theta(1 - \alpha/2))$ is a two - sided $100(1 - \alpha)\%$ confidence interval.

3.2   The Large Sample Approach

The large sample estimate of normal variance is a pooled estimate of the common normal variance defined as

$$\widehat{\theta} = \frac{\sum_{i=1}^{k} \dfrac{\widehat{\theta}^{(i)}}{var\left(\widehat{\theta}^{(i)}\right)}}{\sum_{i=1}^{k} \dfrac{1}{var\left(\widehat{\theta}^{(i)}\right)}}, \tag{8}$$

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

Hence, the large sample solution for confidence interval estimation is

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

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

1. Compute $\overline{x}_i$ and s_i^2 for $i = 1, 2, ..., k$.
2. Calculate $var\left(\widehat{\theta}^{(i)}\right)$ from (4) for $i = 1, 2, ..., k$.
3. Compute $\widehat{\theta}$ following (8).
4. Calculate confidence interval estimation from (9) for $i = 1, 2, ..., k$.

4   Simulation Studies

A simulation study was performed to estimate the coverage probabilities and average lengths of the common variance of the normal distributions for various combinations of the number of samples $k = 2$ and $k = 6$, the sample sizes $n_1 = ... = n_k = n$, the values used for sample sizes were 10, 30, 50,100 and 200 the population mean of normal data within each sample 1, and the population standard deviation $\sigma = 0.10, 0.20, 0.30, 0.40, 0.50, 0.60, 0.70, 0.80, 0.90, 1.00$ and 2.00. In this simulation study, we compared two methods, comprising of our proposed procedure generalized confidence interval approach and the large sample approach. For each parameter setting, 5000 random samples were generated, 2500 R_θ's were obtained for each of the random samples.

Tables 1 and 2 present the coverage probabilities and average lengths for 2 and 6 sample cases respectively. In 2 and 6 sample cases, the generalized

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

n	σ	GCI approach		Large sample approach	
		CP	Length	CP	Length
10	0.10	0.9156	0.0147	0.7602	0.0113
	0.20	0.9282	0.0588	0.7628	0.0453
	0.30	0.9198	0.1321	0.7486	0.1015
	0.40	0.9242	0.2368	0.7572	0.1820
	0.50	0.9306	0.3687	0.7588	0.2842
	0.60	0.9248	0.5330	0.7512	0.4103
	0.70	0.9248	0.7176	0.7526	0.5519
	0.80	0.9242	0.9473	0.7562	0.7311
	0.90	0.9230	1.1958	0.7524	0.9210
	1.00	0.9212	1.4668	0.7468	1.1298
	2.00	0.9268	5.8861	0.7530	4.5365
30	0.10	0.9440	0.0077	0.8762	0.0069
	0.20	0.9336	0.0311	0.8658	0.0277
	0.30	0.9312	0.0699	0.8532	0.0623
	0.40	0.9370	0.1241	0.8672	0.1106
	0.50	0.9328	0.1941	0.8632	0.1726
	0.60	0.9436	0.2806	0.8696	0.2497
	0.70	0.9364	0.3793	0.8626	0.3379
	0.80	0.9402	0.5007	0.8618	0.4453
	0.90	0.9430	0.6295	0.8712	0.5615
	1.00	0.9382	0.7800	0.8654	0.6949
	2.00	0.9360	3.1166	0.8630	2.7751
50	0.10	0.9418	0.0058	0.8936	0.0054
	0.20	0.9434	0.0235	0.9018	0.0217
	0.30	0.9438	0.0529	0.8988	0.0488
	0.40	0.9444	0.0941	0.9014	0.0870
	0.50	0.9360	0.1474	0.8946	0.1361
	0.60	0.9428	0.2127	0.9014	0.1963
	0.70	0.9438	0.2893	0.9020	0.2672
	0.80	0.9398	0.3768	0.8950	0.3473
	0.90	0.9494	0.4778	0.9006	0.4414
	1.00	0.9426	0.5897	0.9000	0.5441
	2.00	0.9464	2.3638	0.9000	2.1832

Table 1. (*continued*)

n	σ	GCI approach		Large sample approach	
		CP	Lenght	CP	Lenght
100	0.10	0.9416	0.0040	0.9224	0.0038
	0.20	0.9454	0.0162	0.9190	0.0155
	0.30	0.9436	0.0366	0.9230	0.0349
	0.40	0.9402	0.0650	0.9178	0.0620
	0.50	0.9456	0.1016	0.9242	0.0969
	0.60	0.9482	0.1468	0.9234	0.1399
	0.70	0.9508	0.1996	0.9272	0.1903
	0.80	0.9448	0.2604	0.9186	0.2484
	0.90	0.9424	0.3298	0.9212	0.3149
	1.00	0.9444	0.4062	0.9240	0.3876
	2.00	0.9504	1.6329	0.9286	1.5557
200	0.10	0.9448	0.0028	0.9326	0.0027
	0.20	0.9512	0.0113	0.9400	0.0110
	0.30	0.9472	0.0255	0.9366	0.0248
	0.40	0.9454	0.0452	0.9332	0.0440
	0.50	0.9504	0.0707	0.9394	0.0689
	0.60	0.9506	0.1020	0.9426	0.0994
	0.70	0.9500	0.1389	0.9376	0.1353
	0.80	0.9480	0.1811	0.9362	0.1764
	0.90	0.9496	0.2293	0.9386	0.2233
	1.00	0.9502	0.2836	0.9352	0.2761
	2.00	0.9480	1.1331	0.9346	1.1039

confidence interval approach and the large sample approach provide the underestimates coverage probabilities for most of the scenarios, especially when the sample size is small. Additionally, the coverage probabilities of the generalized confidence interval approach are better than the large sample approach for all sample sizes, especially when the sample size is small. In overall, the generalized confidence interval approach and the large sample approach have the coverage probabilities close to the nominal level when the sample size increases. In this case, there is no need to see the average lengths from two intervals since the large sample approach provide the coverage probability below the generalized confidence interval approach for almost cases. Finally, it was discovered that the generalized confidence interval approach provided much better results over the large sample approach in terms of coverage probabilities.

Table 2. Empirical coverage probabilities (CP) and length of approximate 95 % two side confidence bounds for common variance of normal distributions (based on 5000 simulations): 6 sample cases.

n	σ	GCI approach		Large sample approach	
		CP	Lenght	CP	Lenght
10	0.10	0.6156	0.0073	0.4614	0.0056
	0.20	0.6234	0.0296	0.4560	0.0227
	0.30	0.6196	0.0666	0.4672	0.0513
	0.40	0.6082	0.1178	0.4514	0.0903
	0.50	0.6086	0.1844	0.4528	0.1418
	0.60	0.6074	0.2653	0.4574	0.2044
	0.70	0.6082	0.3622	0.4488	0.2783
	0.80	0.6180	0.4746	0.4554	0.3648
	0.90	0.6170	0.6001	0.4580	0.4625
	1.00	0.6034	0.7393	0.4482	0.5680
	2.00	0.6098	2.9526	0.4588	2.2705
30	0.10	0.7840	0.0045	0.7238	0.0038
	0.20	0.7880	0.0180	0.7318	0.0154
	0.30	0.7818	0.0405	0.7230	0.0346
	0.40	0.7794	0.0722	0.7274	0.0616
	0.50	0.7824	0.1127	0.7190	0.0960
	0.60	0.7752	0.1624	0.7156	0.1385
	0.70	0.7850	0.2213	0.7210	0.1887
	0.80	0.7756	0.2888	0.7194	0.2468
	0.90	0.7848	0.3654	0.7254	0.3120
	1.00	0.7846	0.4509	0.7334	0.3855
	2.00	0.7730	1.8012	0.7196	1.5387
50	0.10	0.8344	0.0034	0.8080	0.0030
	0.20	0.8314	0.0137	0.8016	0.0122
	0.30	0.8344	0.0309	0.8074	0.0276
	0.40	0.8374	0.0550	0.8124	0.0491
	0.50	0.8274	0.0857	0.7984	0.0766
	0.60	0.8326	0.1236	0.8062	0.1106
	0.70	0.8366	0.1685	0.8004	0.1505
	0.80	0.8380	0.2199	0.8124	0.1966
	0.90	0.8408	0.2787	0.8110	0.2491
	1.00	0.8356	0.3435	0.8026	0.3070
	2.00	0.8446	1.3751	0.8146	1.2294

Table 2. (*continued*)

n	σ	GCI approach		Large sample approach	
		CP	Lenght	CP	Lenght
100	0.10	0.8894	0.0023	0.8806	0.0022
	0.20	0.8888	0.0094	0.8764	0.0088
	0.30	0.8930	0.0213	0.8846	0.0199
	0.40	0.8878	0.0377	0.8784	0.0354
	0.50	0.8784	0.0590	0.8712	0.0554
	0.60	0.8918	0.0851	0.8808	0.0798
	0.70	0.8876	0.1159	0.8798	0.1087
	0.80	0.8948	0.1514	0.8840	0.1420
	0.90	0.8902	0.1914	0.8814	0.1795
	1.00	0.8878	0.2366	0.8786	0.2218
	2.00	0.8844	0.9456	0.8778	0.8871
200	0.10	0.9088	0.0016	0.9072	0.0015
	0.20	0.9158	0.0065	0.9146	0.0063
	0.30	0.9188	0.0147	0.9184	0.0142
	0.40	0.9146	0.0262	0.9124	0.0253
	0.50	0.9132	0.0409	0.9120	0.0396
	0.60	0.9120	0.0590	0.9106	0.0570
	0.70	0.9220	0.0804	0.9210	0.0776
	0.80	0.9154	0.1049	0.9128	0.1014
	0.90	0.9172	0.1328	0.9134	0.1283
	1.00	0.9148	0.1638	0.9140	0.1583
	2.00	0.9110	0.6562	0.9066	0.6337

5 An Empirical Application

In this section, two real data examples are exhibited to illustrate the generalized confidence interval approach and the large sample 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 [15]. Using the data from Table 3, for test the hypothesis of equal mean treatment effects. Under the assumption are normality, homogeneity of variance, and independence of errors. The Shapiro - Wilk normality test indicate 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 generalized confidence interval approach, the generalized confidence interval for the overall variance was (0.0012, 0.0104) with the length of interval 0.0092. In

Table 3. The 18 observations of the shear strength for steel plate girders.

Karlsruhe method			Lehigh method		
1.186	1.151	1.322	1.061	0.992	1.063
1.339	1.200	1.402	1.062	1.065	1.178
1.365	1.537	1.559	1.037	1.086	1.052

Table 4. The 50 observations of blood Sugar Levels (mg/100 g) for 10 Animals from Each of five Breeds (A−E).

A		B		C		D		E	
124	120	111	129	117	148	104	119	142	149
116	110	101	122	142	141	128	106	139	150
101	127	130	103	121	122	130	107	133	149
118	106	108	122	123	139	103	107	120	120
118	130	127	127	121	125	121	115	127	116

comparison, the confidence interval by the large sample approach was (0.0003, 0.0050) with the length of interval 0.0047.

The second example was blood sugar levels (mg/100g) measured from ten animals of five different breeds [16]. The results are presented in Table 4, for test 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 generalized confidence interval approach, the generalized confidence interval for the overall variance was (56.9113, 156.9829) with the length of interval 100.0716. In comparison, the confidence interval by the large sample approach was (62.6062, 155.0373) with the length of interval 92.43106.

6   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 two approaches, the generalized confidence interval and large sample approaches. The generalized confidence interval approach provided coverage probability close to nominal level 0.95 and is better than the large sample approach for all sample sizes. The average lengths increased when the value of σ increased for both approaches. The results indicated that the confidence interval for the common variance of normal distributions based on the generalized confidence interval approach is better than confidence interval based on the large sample approach. In conclusion, the generalized confidence interval can

be successfully used to estimate the common variance of normal distributions. This conclusion supports the research papers of Tian [11], Tian and Wu [12], Krishnamoorthy [13] and Ye et al. [14].

References

1. Cohen, A.: Improved confidence intervals for the variance of a normal distribution. J. Am. Stat. Assoc. **67**, 382–387 (1972)
2. Shorrock, G.: A minimax generalized bayes confidence interval for a normal variance. Ph.D. dissertation, Dept. Statistics, Rutgers Univ (1982)
3. Shorrock, G.: Improved confidence intervals for a normal variance. Ann. Stat. **18**, 972–980 (1990)
4. Nagata, Y.: Improvements of interval estimations for the variance and the ratio of two variances. J. Jpn. Stat. Soc. **19**, 151–161 (1989)
5. Casella, G., Goutis, C.: Improved invariant confidence intervals for a normal variance. Ann. Statist. **19**, 2015–2031 (1991)
6. Kubokawa, T.: A unified approach to improving equivariant estimators. Ann. Stat. **22**, 290–299 (1994)
7. Shorrock, R.W., Zidek, J.V.: An improved estimator of the generalized variance. Ann Stat. **4**(3), 629–638 (1976)
8. Sarkar, S.K.: On improving the shortest length confidence interval for the generalized variance. J. Multivar. Anal. **31**, 136–147 (1989)
9. Iliopoulos, G., Kourouklis, S.: On improved interval estimation for the generalized variance. J. Stat. Plan Infer. **66**, 305–320 (1998)
10. Weerahandi, S.: Generalized confidence intervals. J. Am. Stat. Assoc. **88**, 899–905 (1993)
11. Tian, L.: Inferences on the common coefficient of variation. Stat. Med. **24**, 2213–2220 (2005)
12. Tian, L., Wu, J.: Inferences on the common mean of several log-normal populations: the generalized variable approach. Biometrical J. **49**, 944–951 (2007)
13. Krishnamoorthy, K., Lu, Y.: Inference on the common means of several normal populations based on the generalized variable method. Biometrics **59**, 237–247 (2003)
14. Ye, R.D., et al.: Inferences on the common mean of several inverse gaussian populations. Comput. Stat. Data Anal. **54**, 906–915 (2010)
15. Montgomery, C.D.: Design and Analysis of Experiments, p. 57. Wiley, New York (2001)
16. Rencher, A.C., Schaalje, B.G.: Linear Models in Statistics, p. 373. Wiley, New Jersey (2008)