**KMITL Sci. Tech. J. Vol. 15 No. 1 Jan. - Jun. 2015**

# Confidence Interval for The Coefficient of Variation in A Normal Distribution with A Known Population Mean after A Preliminary T Test

Wararit Panichkitkosolkul*

Department of Mathematics and Statistics, Faculty of Science and Technology, Thammasat University, Pathumthani 12120, Thailand

## Abstract

This paper proposes the confidence interval for the coefficient of variation in a normal distribution with a known population mean after a preliminary T test. This has applications when one knows the population mean of a control group. A Monte Carlo simulation study was conducted to evaluate the performance of the proposed confidence interval with two existing confidence intervals. Two performance measures were used to assess the confidence interval for the coefficient of variation, namely: coverage probability and expected length. Simulation results showed that the proposed confidence interval performs well in terms of coverage probability and expected length compared with two existing confidence intervals. A real data example presenting the melting point of beeswax from 59 sources was used for illustration and performing a comparison.

**Keywords:** measure of dispersion, preliminary test, coverage probability, expected length, simulation study.

## 1. Introduction

The coefficient of variation is one of the dispersion measures of data, defined as a ratio of the population standard deviation $\sigma$ to the population mean $\mu$, $\kappa = \sigma / \mu$, where $\mu \neq 0$. It is a unit free measure that quantifies the degree of variability relative to the mean [1]. The coefficient of variation has been widely preferred to the standard deviation for comparing the variations of several variables obtained by different units. The natural sample estimate of $\kappa$ is given as

$$\hat{\kappa} = S / \bar{X}, \tag{1}$$

where $\bar{X}$ and $S$ are the sample mean and sample standard deviation, respectively. The sample coefficient of variation, $\hat{\kappa}$ has been widely applied as a point estimate of $\kappa$ in many fields such as science, medical sciences, engineering, economics and others (see Nairy and Rao [2]). For example, the coefficient of variation was reported by Ahn [3] to analyze the uncertainty of fault trees. The coefficient of variation of the measured strength for ceramics was studied by Gong and Li [4]. The work of Faber and Korn [5] applied the coefficient of variation for analyzing the variation in the mean synaptic response of the central nervous system. Hamer *et al.* [6] used the coefficient of variation to study the homogeneity of bone test samples produced from a particular method to help determine the effect of external treatments on the properties of bones. Billings *et al.* [7] studied the impact of socioeconomic status on hospital use in New York City by using the coefficient of variation. Miller and Karson [8] tested the equality of the

*Corresponding author: E-mail: wararit@mathstat.sci.tu.ac.th

coefficient of variation for two stocks. The coefficient of variation was also used by Worthington and Higgs [9] to measure the degree of risk in relation to the mean return. Furthermore, Pyne *et al.* [10] applied the coefficient of variation to study the variability of the competitive performance of Olympic swimmers.

It is well known that the point estimate of $\kappa$ given in (1) is a useful statistical measure. However, its confidence interval is more useful than the point estimate because a confidence interval provides much more information about the population characteristic of interest than does a point estimate (e.g., Smithson [11], Thompson [12], Steiger [13]). Several researchers have constructed many confidence intervals for estimating the population coefficient of variation. A confidence interval for $\kappa$ based on the chi-square distribution was proposed by McKay [14] and his confidence interval works well when the value of $\kappa$ is less than 0.33 [15-19]. Vangel [20] developed a new confidence interval for $\kappa$, which is called a modified McKay's confidence interval. The modified McKay's confidence interval is closely related to McKay's confidence interval but it is usually more accurate and nearly exact under normality. The modified McKay's confidence interval was improved by replacing the typical sample estimate of $\kappa$ with the maximum likelihood estimate in the case of a normal distribution [21]. Sharma and Krishna [22] discussed the asymptotic distribution and confidence interval of the reciprocal of the coefficient of variation without making any assumptions about the population distribution. Miller [23] studied the approximate distribution of $\kappa$ and presented the approximation confidence interval for $\kappa$ in a normal distribution. Comparisons of many confidence intervals for $\kappa$, namely McKay's, Miller's and Sharma-Krishna's confidence intervals were conducted under the same simulation conditions by Ng [24].

Mahmoudvand and Hassani [25] derived an approximately unbiased estimator for $\kappa$ in a normal distribution and also used this estimator for estimating two approximation confidence intervals of $\kappa$. Albatineh *et al*. [26] proposed the confidence intervals for $\kappa$ using ranked set sampling. The confidence intervals for $\kappa$ in normal and lognormal distributions were developed by Koopmans *et al.* [27] and Verril [28]. Buntao and Niwitpong [29] also proposed the interval estimation for the difference of the coefficient of variation for lognormal and delta-lognormal distributions. Vangel [30] proposed a method based on an analysis of the distribution of a class of approximate pivotal quantities for the normal coefficient of variation. The confidence interval for $\kappa$ in the case of non-iid random variables was constructed by Curto and Pinto [31]. Gulhar *et al.* [32] compared many confidence intervals for the population coefficient of variation based on parametric, nonparametric and modified methods.

In many situations, the population mean may be known. For instance the population mean of a control group is known. The confidence intervals for $\kappa$ proposed by the aforementioned researchers have not been used for estimating the population coefficient of variation for a normal distribution with a known population mean. Therefore, the recent work of Panichkitkosolkul [33] proposed three confidence intervals for the coefficient of variation in a normal distribution with a known population mean, namely: the normal approximation confidence interval, the shortest-length confidence interval, and the equal-tailed confidence interval. Numerous research papers have indicated that the preliminary test was a useful tool for improving the accuracy of a confidence interval and prediction interval (see Paksaranuwat and Niwitpong [34], Kabaila and Farchione [35], Chiou and Han [36], Panichkitkosolkul and Niwitpong [37-38]). This paper extends the confidence interval for the coefficient of variation proposed by Panichkitkosolkul [33]. Namely, we apply a preliminary T test in order to improve the accuracy of the confidence interval for the coefficient of variation in a normal distribution with a known population mean. We use the T test as a preliminary test for hypothesis testing: $H_0 : \mu = \mu_0$ versus $H_1 : \mu \neq \mu_0$, where $\mu_0$ is a known population mean.

The rest of this paper is organized as follows: The theoretical background of the confidence intervals is discussed in Section 2. Section 3 presents the investigations of the performance of the confidence intervals through a Monte Carlo simulation study. A comparison of the confidence intervals is also illustrated by using a real data example in Section 4. Conclusions are provided in the final section.

## 2. Theoretical Results

In this section, we review the normal approximation confidence interval for the coefficient of variation in the normal distribution presented by Mahmoudvand and Hassani [25] in Section A. Later, the details of the normal approximation confidence interval for the coefficient of variation in a normal distribution with a known population mean, constructed by Panichkitkosolkul [33], are shown in Section B. Finally, Section C proposes the confidence interval for the coefficient of variation in a normal distribution with a known population mean after a preliminary T test.

### 2.1 Normal Approximation Confidence Interval for the Coefficient of Variation in a Normal Distribution

To find the unbiased estimator of $\kappa$ and the variance of the unbiased estimator of $\kappa$, we have to prove the following lemma. After that, the normal approximation confidence interval for $\kappa$ in a normal distribution is proposed.

**Lemma 1.** Let $X_1,...,X_n$ be a random sample from normal distribution with a mean $\mu$ and variance $\sigma^2$. An approximation of $E(\hat{\kappa})$ is given by

$$E(\hat{\kappa}) \approx (2-c_n)\kappa,$$

and an approximation of $\operatorname{var}(\hat{\kappa})$ is $\operatorname{var}(\hat{\kappa}) \approx (1-c_n^2)\kappa^2$,
where

$$c_n = \sqrt{\frac{2}{n-1}}\frac{\Gamma(n/2)}{\Gamma((n-1)/2)}$$

and $\Gamma(\cdot)$ is the gamma function.

**Proof of Lemma 1.** See Mahmoudvand and Hassani [25].

Note that $c_n \to 1$ as $n \to \infty$. Then $(2-c_n)\kappa \to \kappa$ as $n \to \infty$. Therefore, it follows that

$$\lim_{n\to\infty} E(\hat{\kappa}) = \kappa.$$

It means that $\hat{\kappa}$ is the asymptotically unbiased estimator for $\kappa$. From Lemma 1, the unbiased estimator of $\kappa$ is

$$\hat{\kappa}' = \frac{\hat{\kappa}}{2-c_n}.$$

Using Lemma 1, the mean and variance of $\hat{\kappa}'$ are estimated by

$$E(\hat{\kappa}') = E\left(\frac{\hat{\kappa}}{2-c_n}\right) \approx \kappa,$$

and

$$\operatorname{var}(\hat{\kappa}') = \operatorname{var}\left(\frac{\hat{\kappa}}{2-c_n}\right) = \frac{1}{(2-c_n)^2}\operatorname{var}(\hat{\kappa}) \approx \frac{(1-c_n^2)}{(2-c_n)^2}\kappa^2.$$

Thus,

$$\lim_{n\to\infty}\operatorname{var}(\hat{\kappa}') = 0.$$

Hence, $\hat{\kappa}'$ is also asymptotically consistent for $\kappa$. Using the normal approximate, we have

$$\begin{aligned} z &= \frac{\hat{\kappa}'-\kappa}{\sqrt{\operatorname{var}(\hat{\kappa}')}} = \frac{\hat{\kappa}/(2-c_n)-\kappa}{\sqrt{\kappa^2(1-c_n^2)/(2-c_n)^2}} \\ &= \frac{\hat{\kappa}-(2-c_n)\kappa}{\kappa\sqrt{1-c_n^2}} \to N(0,1). \end{aligned} \tag{2}$$

Therefore, the $100(1-\alpha)\%$ normal approximation confidence interval for $\kappa$ based on Equation (2) is

$$CI_{MH} = \left[\frac{\hat{\kappa}}{2-c_n+z_{1-\alpha/2}\sqrt{1-c_n^2}},\ \frac{\hat{\kappa}}{2-c_n-z_{1-\alpha/2}\sqrt{1-c_n^2}}\right], \tag{3}$$

where $z_{1-\alpha/2}$ is the $100(1-\alpha/2)$ percentile of the standard normal distribution.

### 2.2 Normal Approximation Confidence Interval for the Coefficient of Variation in a Normal Distribution with a Known Population Mean

If the population mean is known to be $\mu_0$, then the population coefficient of variation is given by $\kappa_0 = \sigma/\mu_0$. The sample estimate of $\kappa_0$ is

$$\hat{\kappa}_0 = \frac{S_0}{\mu_0}, \tag{4}$$

where

$$S_0 = \sqrt{n^{-1}\sum_{i=1}^{n}(X_i-\mu_0)^2}.$$

To find the expectation of Equation (4), we have to prove the following lemma.

**Lemma 2.** Let $X_1,\dots,X_n$ be a random sample from normal distribution with a known population mean $\mu_0$ and variance $\sigma^2$ and let

$$S_0 = \sqrt{n^{-1}\sum_{i=1}^{n}(X_i-\mu_0)^2}.$$

Then

$$E(S_0) = c_{n+1}\sigma,$$

and

$$\operatorname{var}(S_0) = (1 - c_{n+1}^2)\sigma^2,$$

where

$$c_{n+1} = \sqrt{\frac{2}{n}}\frac{\Gamma((n+1)/2)}{\Gamma(n/2)}$$

and $\Gamma(\cdot)$ is the gamma function.

**Proof of Lemma 2.** See Panichkitkosolkul [33].
By using Lemma 2, we can show that the mean and variance of $\hat{\kappa}_0$ are

$$E(\hat{\kappa}_0) = \frac{c_{n+1}\sigma}{\mu_0} = c_{n+1}\kappa_0, \tag{5}$$

and

$$\operatorname{var}(\hat{\kappa}_0) = \left(\frac{1 - c_{n+1}^2}{\mu_0^2}\right)\sigma^2 = (1 - c_{n+1}^2)\kappa_0^2.$$

Note that $c_{n+1} \to 1$ as $n \to \infty$. Therefore, it follows that

$$\lim_{n\to\infty} E(\hat{\kappa}_0) = \kappa_0.$$

It means that $\hat{\kappa}_0$ is asymptotically unbiased and asymptotically consistent for $\kappa_0$. From Equation (5), the unbiased estimator of $\kappa_0$ is

$$\hat{\kappa}_0' = \frac{\hat{\kappa}_0}{c_{n+1}}.$$

Using Lemma 2, the mean and variance of $\hat{\kappa}_0'$ are given by

$$E(\hat{\kappa}_0') = E\left(\frac{\hat{\kappa}_0}{c_{n+1}}\right) = \kappa_0,$$

and

$$\begin{aligned}\operatorname{var}(\hat{\kappa}_0') &= \operatorname{var}\left(\frac{\hat{\kappa}_0}{c_{n+1}}\right) = \frac{1}{c_{n+1}^2}\operatorname{var}\left(\frac{S_0}{\mu_0}\right)\\ &= \frac{1}{c_{n+1}^2\mu_0^2}(1 - c_{n+1}^2)\sigma^2 = \left(\frac{1 - c_{n+1}^2}{c_{n+1}^2}\right)\kappa_0^2.\end{aligned}$$

Thus,

$$\lim_{n\to\infty} \operatorname{var}(\hat{\kappa}_0') = 0.$$

Hence, $\hat{\kappa}_0'$ is also asymptotically consistent for $\kappa_0$. Using the normal approximate, we have

$$\begin{aligned} z &= \frac{\hat{\kappa}_0' - \kappa_0}{\sqrt{\operatorname{var}(\hat{\kappa}_0')}} = \frac{\hat{\kappa}_0 / c_{n+1} - \kappa_0}{\sqrt{(1-c_{n+1}^2)\kappa_0^2 / c_{n+1}^2}} \\ &= \frac{\hat{\kappa}_0 - c_{n+1}\kappa_0}{\kappa_0\sqrt{1-c_{n+1}^2}} \to N(0,1). \end{aligned} \tag{6}$$

Therefore, the $100(1-\alpha)\%$ normal approximation confidence interval for $\kappa_0$ based on Equation (6) is

$$CI_{PK} = \left[ \frac{\hat{\kappa}_0}{c_{n+1} + z_{1-\alpha/2}\sqrt{1-c_{n+1}^2}},\ \frac{\hat{\kappa}_0}{c_{n+1} - z_{1-\alpha/2}\sqrt{1-c_{n+1}^2}} \right], \tag{7}$$

where $z_{1-\alpha/2}$ is the $100(1-\alpha/2)$ percentile of the standard normal distribution.

### 2.3 Confidence Interval for the Coefficient of Variation in a Normal Distribution with a Known Population Mean after a Preliminary T Test

The null hypothesis and alternative hypotheses of the preliminary T test are

$$H_0 : \mu = \mu_0$$

$$H_1 : \mu \neq \mu_0.$$

The test statistic is $T = \sqrt{n}(\bar{X} - \mu_0)/S,$ where $\bar{X}$ and $S$ are the sample mean and sample standard deviation, respectively. The test calls to reject the null hypothesis at the $\gamma$ level of significance if $|T| \geq t_{1-\gamma/2}$ where $t_{1-\gamma/2}$ is the $1-\gamma/2$ percentile of the student's t distribution.

If $H_0$ is satisfied $(\mu = \mu_0)$, the $100(1-\alpha)\%$ confidence interval for the coefficient of variation is $CI_{PK}$ shown in Equation (7). If $H_0$ is rejected, the $100(1-\alpha)\%$ confidence interval for the coefficient of variation is $CI_{MH}$ given in Equation (3). Therefore, the confidence interval for the coefficient of variation in a normal distribution with a known population mean after a preliminary T test is

$$CI_{PRE} = \begin{cases} CI_{PK}, \text{ if } H_0 \text{ is not rejected,} \\ CI_{MH}, \text{ if } H_0 \text{ is rejected.} \end{cases} \tag{8}$$

## 3. Simulation Study

The estimated coverage probability is an important factor for judging the performance of a confidence interval. Generally, we prefer a confidence interval which has a coverage probability close to the nominal confidence level. Furthermore, the expected length has also been considered for comparing the performance of a confidence interval. Therefore, we prefer a short-length confidence interval which has a coverage probability close to the nominal confidence level. The estimated coverage probability and the expected length (based on $M$ replicates) is given by

$$1-\alpha = \frac{\#(L \le \kappa_0 \le U)}{M},$$

and
$$Length = \frac{\sum_{j=1}^{M}(U_j - L_j)}{M},$$

where $\#(L \le \kappa_0 \le U)$ denotes the number of simulation runs for which the population coefficient of variation $\kappa_0$ lies within the confidence interval.

A Monte Carlo simulation was conducted using the R statistical software [39-41] version 3.1.0 to evaluate the performance of a proposed confidence interval for the coefficient of variation in a normal distribution with a known population mean after a preliminary T test, $CI_{PRE}$, and two existing confidence intervals, $CI_{MH}$ and $CI_{PK}$. The data were generated from a normal distribution with a known population mean $\mu_0$ = 10 and $\kappa_0$ = 0.05, 0.10, 0.20, 0.33, 0.40 and 0.50, sample sizes; $n$ = 5, 10, 15, 25, 50 and 100. For each combination of $n$ and $\kappa_0$, 50,000 simulations were performed. The nominal confidence levels were set to 90% and 95%. For $CI_{PRE}$, the level of significance of a preliminary T test was set to 5%. Tables 1 and 2 show the estimated coverage probabilities of the $CI_{MH}$, $CI_{PK}$ and $CI_{PRE}$ and their expected lengths for 90% and 95% confidence levels, respectively. As can be seen from Tables 1 and 2, $CI_{PK}$ and $CI_{PRE}$ estimated coverage probabilities close to the nominal confidence level. On the other hand, $CI_{MH}$ is not acceptable as its coverage may drop below 90% and 95% in some situations, i.e., $\kappa_0 \ge 0.33$. Additionally, the estimated coverage probabilities of $CI_{MH}$ and $CI_{PRE}$ decrease as the values of $\kappa_0$ get larger (i.e. for 90% $CI_{MH}$, $n$ =10, 0.8893 for $\kappa_0$ = 0.05; 0.8600 for $\kappa_0$ = 0.33; 0.8168 for $\kappa_0$ = 0.50). However, the estimated coverage probabilities of the $CI_{PK}$ do not increase or decrease according to the values of $\kappa_0$.

Generally, if two or more confidence intervals have similar estimated coverage probabilities, we can compare their expected lengths, and we prefer the shorter one. Here, $CI_{PK}$ and $CI_{PRE}$ have similar estimated coverage probabilities but the proposed confidence interval, the expected length of $CI_{PRE}$, is shorter than that of $CI_{PK}$. Therefore, the preliminary T test can help to improve the accuracy of $CI_{PK}$. Additionally, we can see that as the sample size increases, expected lengths decrease (i.e., for 90% $CI_{PRE}$, $\kappa_0$ =0.20, 0.1704 for $n$ =10; 0.0984 for $n$ =25; 0.0676 for $n$ =50). The expected lengths of all confidence intervals are not significantly different for large samples.

**Table 1** Estimated Coverage Probabilities and Expected Lengths of 90% Confidence Intervals for the Coefficient of Variation in A Normal Distribution with A Known Population Mean

| $n$ | $\kappa_0$ | Coverage Probabilities | | | Expected Lengths | | |
|---|---|---|---|---|---|---|---|
| | | MH | PK | PRE | MH | PK | PRE |
| 5 | 0.05 | 0.8765 | 0.9036 | 0.8809 | 0.0652 | 0.0742 | 0.0722 |
| | 0.10 | 0.8732 | 0.9037 | 0.8812 | 0.1305 | 0.1482 | 0.1443 |
| | 0.20 | 0.8647 | 0.9041 | 0.8811 | 0.2617 | 0.2954 | 0.2876 |
| | 0.33 | 0.8455 | 0.9019 | 0.8792 | 0.4413 | 0.4901 | 0.4780 |
| | 0.40 | 0.8336 | 0.9045 | 0.8812 | 0.5408 | 0.5931 | 0.5804 |
| | 0.50 | 0.8054 | 0.9025 | 0.8754 | 0.6950 | 0.7435 | 0.7309 |
| 10 | 0.05 | 0.8893 | 0.9004 | 0.8899 | 0.0409 | 0.0432 | 0.0426 |
| | 0.10 | 0.8884 | 0.9019 | 0.8907 | 0.0816 | 0.0862 | 0.0850 |
| | 0.20 | 0.8774 | 0.9008 | 0.8879 | 0.1638 | 0.1727 | 0.1704 |
| | 0.33 | 0.8600 | 0.9024 | 0.8868 | 0.2726 | 0.2850 | 0.2818 |
| | 0.40 | 0.8397 | 0.9019 | 0.8822 | 0.3324 | 0.3457 | 0.3419 |
| | 0.50 | 0.8168 | 0.9002 | 0.8787 | 0.4197 | 0.4320 | 0.4286 |
| 15 | 0.05 | 0.8923 | 0.9012 | 0.8943 | 0.0322 | 0.0333 | 0.0330 |
| | 0.10 | 0.8916 | 0.9012 | 0.8934 | 0.0645 | 0.0667 | 0.0662 |
| | 0.20 | 0.8805 | 0.9007 | 0.8918 | 0.1289 | 0.1333 | 0.1321 |
| | 0.33 | 0.8600 | 0.9004 | 0.8874 | 0.2139 | 0.2199 | 0.2183 |
| | 0.40 | 0.8419 | 0.9008 | 0.8838 | 0.2601 | 0.2668 | 0.2649 |
| | 0.50 | 0.8175 | 0.9019 | 0.8805 | 0.3274 | 0.3334 | 0.3317 |
| 25 | 0.05 | 0.8957 | 0.9001 | 0.8961 | 0.0242 | 0.0247 | 0.0246 |
| | 0.10 | 0.8943 | 0.9007 | 0.8960 | 0.0484 | 0.0494 | 0.0492 |
| | 0.20 | 0.8838 | 0.9012 | 0.8943 | 0.0970 | 0.0989 | 0.0984 |
| | 0.33 | 0.8618 | 0.8996 | 0.8877 | 0.1605 | 0.1632 | 0.1624 |
| | 0.40 | 0.8446 | 0.9004 | 0.8846 | 0.1949 | 0.1977 | 0.1969 |
| | 0.50 | 0.8201 | 0.9006 | 0.8806 | 0.2447 | 0.2473 | 0.2466 |
| 50 | 0.05 | 0.8974 | 0.9017 | 0.8990 | 0.0168 | 0.0169 | 0.0169 |
| | 0.10 | 0.8951 | 0.9002 | 0.8971 | 0.0336 | 0.0339 | 0.0338 |
| | 0.20 | 0.8860 | 0.9012 | 0.8961 | 0.0671 | 0.0678 | 0.0676 |
| | 0.33 | 0.8662 | 0.9017 | 0.8914 | 0.1110 | 0.1120 | 0.1117 |
| | 0.40 | 0.8444 | 0.8980 | 0.8837 | 0.1347 | 0.1356 | 0.1353 |
| | 0.50 | 0.8201 | 0.9012 | 0.8811 | 0.1687 | 0.1696 | 0.1693 |
| 100 | 0.05 | 0.8973 | 0.9000 | 0.8986 | 0.0117 | 0.0118 | 0.0118 |
| | 0.10 | 0.8953 | 0.8995 | 0.8975 | 0.0235 | 0.0236 | 0.0236 |
| | 0.20 | 0.8859 | 0.9001 | 0.8961 | 0.0470 | 0.0472 | 0.0471 |
| | 0.33 | 0.8651 | 0.9023 | 0.8924 | 0.0776 | 0.0779 | 0.0778 |
| | 0.40 | 0.8471 | 0.9007 | 0.8866 | 0.0942 | 0.0945 | 0.0944 |
| | 0.50 | 0.8214 | 0.9006 | 0.8797 | 0.1177 | 0.1181 | 0.1180 |

**Table 2** Estimated Coverage Probabilities and Expected Lengths of 95% Confidence Intervals for the Coefficient of Variation in A Normal Distribution with A Known Population Mean

| $n$ | $\kappa_0$ | Coverage Probabilities | | | Expected Lengths | | |
|---|---|---|---|---|---|---|---|
| | | MH | PK | PRE | MH | PK | PRE |
| 5 | 0.05 | 0.9431 | 0.9548 | 0.9411 | 0.0929 | 0.1057 | 0.1029 |
| | 0.10 | 0.9420 | 0.9534 | 0.9398 | 0.1861 | 0.2115 | 0.2059 |
| | 0.20 | 0.9388 | 0.9561 | 0.9420 | 0.3746 | 0.4236 | 0.4125 |
| | 0.33 | 0.9205 | 0.9525 | 0.9380 | 0.6280 | 0.6983 | 0.6821 |
| | 0.40 | 0.9082 | 0.9544 | 0.9378 | 0.7695 | 0.8438 | 0.8257 |
| | 0.50 | 0.8871 | 0.9544 | 0.9349 | 0.9858 | 1.0567 | 1.0379 |
| 10 | 0.05 | 0.9466 | 0.9527 | 0.9464 | 0.0521 | 0.0551 | 0.0543 |
| | 0.10 | 0.9441 | 0.9521 | 0.9451 | 0.1046 | 0.1104 | 0.1090 |
| | 0.20 | 0.9390 | 0.9520 | 0.9447 | 0.2096 | 0.2207 | 0.2178 |
| | 0.33 | 0.9201 | 0.9505 | 0.9394 | 0.3484 | 0.3640 | 0.3598 |
| | 0.40 | 0.9094 | 0.9501 | 0.9374 | 0.4252 | 0.4420 | 0.4374 |
| | 0.50 | 0.8884 | 0.9524 | 0.9354 | 0.5364 | 0.5517 | 0.5473 |
| 15 | 0.05 | 0.9485 | 0.9521 | 0.9480 | 0.0400 | 0.0415 | 0.0411 |
| | 0.10 | 0.9445 | 0.9505 | 0.9456 | 0.0802 | 0.0831 | 0.0823 |
| | 0.20 | 0.9379 | 0.9503 | 0.9445 | 0.1604 | 0.1658 | 0.1644 |
| | 0.33 | 0.9239 | 0.9518 | 0.9424 | 0.2662 | 0.2738 | 0.2717 |
| | 0.40 | 0.9119 | 0.9514 | 0.9394 | 0.3235 | 0.3318 | 0.3295 |
| | 0.50 | 0.8890 | 0.9521 | 0.9355 | 0.4068 | 0.4144 | 0.4122 |
| 25 | 0.05 | 0.9471 | 0.9489 | 0.9462 | 0.0296 | 0.0302 | 0.0300 |
| | 0.10 | 0.9474 | 0.9512 | 0.9481 | 0.0592 | 0.0604 | 0.0601 |
| | 0.20 | 0.9417 | 0.9515 | 0.9476 | 0.1184 | 0.1207 | 0.1201 |
| | 0.33 | 0.9247 | 0.9506 | 0.9423 | 0.1962 | 0.1993 | 0.1984 |
| | 0.40 | 0.9115 | 0.9501 | 0.9392 | 0.2380 | 0.2414 | 0.2405 |
| | 0.50 | 0.8886 | 0.9508 | 0.9350 | 0.2989 | 0.3021 | 0.3011 |
| 50 | 0.05 | 0.9480 | 0.9497 | 0.9485 | 0.0202 | 0.0204 | 0.0204 |
| | 0.10 | 0.9476 | 0.9504 | 0.9481 | 0.0405 | 0.0409 | 0.0408 |
| | 0.20 | 0.9397 | 0.9497 | 0.9459 | 0.0810 | 0.0818 | 0.0816 |
| | 0.33 | 0.9233 | 0.9502 | 0.9422 | 0.1337 | 0.1347 | 0.1344 |
| | 0.40 | 0.9105 | 0.9498 | 0.9382 | 0.1623 | 0.1635 | 0.1632 |
| | 0.50 | 0.8904 | 0.9490 | 0.9331 | 0.2033 | 0.2044 | 0.2041 |
| 100 | 0.05 | 0.9488 | 0.9500 | 0.9490 | 0.0141 | 0.0142 | 0.0141 |
| | 0.10 | 0.9464 | 0.9493 | 0.9478 | 0.0282 | 0.0283 | 0.0283 |
| | 0.20 | 0.9421 | 0.9508 | 0.9483 | 0.0563 | 0.0566 | 0.0565 |
| | 0.33 | 0.9251 | 0.9499 | 0.9426 | 0.0930 | 0.0934 | 0.0933 |
| | 0.40 | 0.9115 | 0.9508 | 0.9396 | 0.1128 | 0.1132 | 0.1131 |
| | 0.50 | 0.8906 | 0.9492 | 0.9331 | 0.1412 | 0.1415 | 0.1414 |

## 4. A Real Date Example

To illustrate the application of the confidence interval proposed in the previous section, we used the melting point of beeswax obtained from 59 sources. The melting points (°C) are listed as follows:

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 63.78 | 63.45 | 63.58 | 63.08 | 63.40 | 64.42 | 63.27 | 63.10 | 63.34 | 63.50 |
| 63.83 | 63.63 | 63.27 | 63.30 | 63.83 | 63.50 | 63.36 | 63.86 | 63.34 | 63.92 |
| 63.88 | 63.36 | 63.36 | 63.51 | 63.51 | 63.84 | 64.27 | 63.50 | 63.56 | 63.39 |
| 63.78 | 63.92 | 63.92 | 63.56 | 63.43 | 64.21 | 64.24 | 64.12 | 63.92 | 63.53 |
| 63.50 | 63.30 | 63.86 | 63.93 | 63.43 | 64.40 | 63.61 | 63.03 | 63.68 | 63.13 |
| 63.41 | 63.60 | 63.13 | 63.69 | 63.05 | 62.85 | 63.31 | 63.66 | 63.60 | |

The data were taken from the study by White *et al*. [42] (cited in Rice [43], p.378). The average of the melting point of beeswax was 63.58881°C, with a standard deviation of 0.347221°C. From past experience, we know the population mean of the melting point of beeswax is about 63.58°C. An unbiased estimator of the coefficient of variation is $\hat{k}' \simeq 5.416 \times 10^{-3}$ The histogram, density plot, Box-and-Whisker plot and normal quantile-quantile plot are displayed in Figure 1. The Shapiro-Wilk test for normality ($W$ = 0.9748, p-value = 0.2579) and Figure 1 supported the assertion that the data follows a normal distribution.

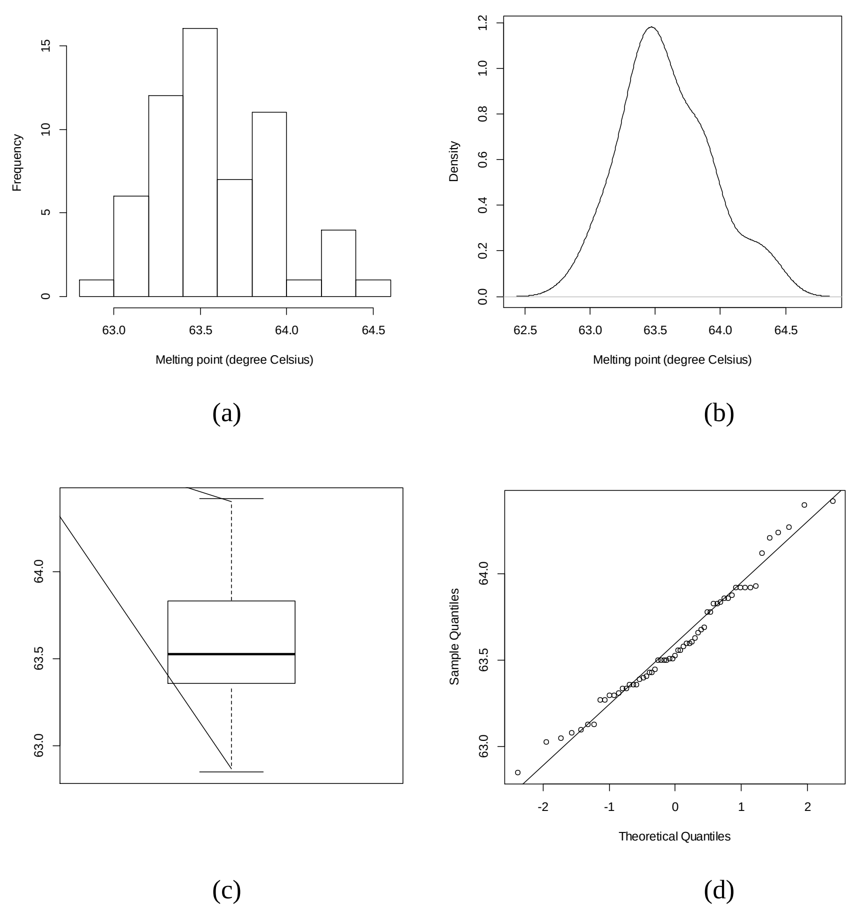


**Figure 1** (a) histogram, (b) density plot, (c) Box-and-Whisker plot and (d) normal quantile-quantile plot of the melting point of beeswax

The 95% proposed and existing confidence intervals for the coefficient of variation were calculated and reported in Table III. The null and alternative hypotheses of the preliminary T test are $H_0 : \mu = 63.58$ versus $H_1 : \mu \neq 63.58.$ The preliminary T test ($T$ =0.195, p-value = 0.8461) reported that the null hypothesis is not rejected. Thus, the proposed confidence interval was the same as Panichkitkosolkul's confidence interval. This result confirms that the confidence interval proposed in this paper is more efficient than the existing confidence intervals in terms of coverage probability.

**Table 3** The 95% Confidence Intervals for the Coefficient of Variation of the Melting Point of Beeswax

| Methods | Confidence Intervals | | Lengths |
|---|---|---|---|
| | Lower Limit | Upper Limit | |
| MH | $4.605\times10^{-3}$ | $6.637\times10^{-3}$ | $2.032\times10^{-3}$ |
| PK | $4.607\times10^{-3}$ | $6.640\times10^{-3}$ | $2.033\times10^{-3}$ |
| PRE | $4.607\times10^{-3}$ | $6.640\times10^{-3}$ | $2.033\times10^{-3}$ |

## 5. Conclusions

In this paper, the confidence interval for the coefficient of variation in a normal distribution with a known population mean after a preliminary T test was proposed. The proposed confidence interval was compared with Mahmoudvand-Hassani's [25] and Panichkitkosolkul's [33] confidence intervals by using simulated data generated from a normal distribution with a known population mean. Two main criteria for comparison were implemented, namely: coverage probability and expected lengths. Simulation results revealed that the proposed confidence interval tend to be better than two existing confidence intervals in terms of the coverage probability and the expected length. A real data example presenting 59 melting points of beeswax was analyzed. For this example, the proposed confidence interval had an expected length that was slightly wider than the confidence interval of Mahmoudvand-Hassani [25]. However, the proposed confidence interval is a good alternative in terms of coverage probability.

## 6. Acknowledgements

The author wishes to express sincere thanks to the referees for their constructive and valuable suggestions and comments that have greatly improved the presentation of this paper.

## References

[1] Kelley, K. **2007**. Sample size planning for the coefficient of variation from the accuracy in parameter estimation approach, *Behavior Research Methods*, 39, 755-766.

[2] Nairy, K.S., and Rao, K.A., **2003**. Tests of coefficients of variation of normal population, *Communication in Statistics-Simulation and Computation*, 32, 641-661.

[3] Ahn, K., **1995**. Use of coefficient of variation for uncertainty analysis in fault tree analysis, *Reliability Engineering and System Safety*, 47, 229-230.

[4] Gong, J., and Li, Y., **1999**. Relationship between the estimated Weibull modulus and the coefficient of variation of the measured strength for ceramics, *Journal of the American Ceramic Society*, 82, 449-452.

[5] Faber, D.S., and Korn, H., **1991**. Applicability of the coefficient of variation method for analyzing synaptic plasticity, *Biophysical Society*, 60, 1288-1294.
[6] Hamer, A.J., Strachan, J.R.., Black, M.M., Ibbotson, C., and Elson, R.A., **1995**. A new method of comparative bone strength measurement, *Journal of Medical Engineering and Technology*, 19, 1-5.
[7] Billings, J., Zeitel, L., Lukomink, J., Carey, T.S., Blank, A.E., and Newman, L., **1993**. Impact of socioeconomic status of hospital use in New York City, *Health Affairs*, 12, 162-173.
[8] Miller, E.G., and Karson, M.J., **1977**. Testing the equality of two coefficients of variation, American Statistical Association: *Proceedings of the Business and Economics Section*, Part I, 278-283.
[9] Worthington, A.C., and Higgs, H., **2003**. Risk, return and portfolio diversification in major painting markets: the application of conventional financial analysis to unconventional investments, *Discussion Papers* No.148 in Economics, Finance and International Competitiveness, School of Economics and Finance, Queensland University of Technology.
[10] Pyne, D.B., Trewin, C.B., and Hopkins, W.G., **2004**. Progression and variability of competitive performance of Olympic swimmers, *Journal of Sports Sciences*, 22, 613-620.
[11] Smithson, M., **2001**. Correct confidence intervals for various regression effect sizes and parameters: the importance of noncentral distributions in computing intervals, *Educational and Psychological Measurement*, 61, 605-632.
[12] Thompson, B., **2002**. What future quantitative social science research could look like: Confidence intervals for effect sizes, *Educational Researcher*, 31, 25-32.
[13] Steiger, J.H., **2004**. Beyond the F test: Effect size confidence intervals and tests of close fit in the analysis of variance and contrast analysis, *Psychological Methods*, 9, 164-182.
[14] McKay, A.T., **1932**. Distribution of the coefficient of variation and the extended t distribution, *Journal of the Royal Statistics Society*, 95, 695-698.
[15] Fieller, E.C. **1932**. A numerical test of the adequacy of A.T. McKay's approximation, *Journal of the Royal Statistical Society*, 95, 699-702.
[16] Iglewicz, B., **1967**. Some properties of the coefficient of variation, *Ph.D. thesis*, Virginia Polytechnic Institute.
[17] Iglewicz, B., and Myers, R.H., **1970**. Comparisons of approximations to the percentage points of the sample coefficient of variation, *Technometrics*, 12, 166-169.
[18] Pearson, E.S., **1932**. Comparison of A.T. McKay's approximation with experimental sampling results, *Journal of the Royal Statistics Society*, 95, 703-704.
[19] Umphrey, G.J., **1983**. A comment on McKay's approximation for the coefficient of variation, *Communications in Statistics-Simulation and Computation*, 12, 629-635.
[20] Vangel, M.G., **1996**. Confidence intervals for a normal coefficient of variation, *The American Statistician*, 15, 21-26.
[21] Panichkitkosolkul, W., **2009**. Improved confidence intervals for a coefficient of variation of a normal distribution, *Thailand Statistician*, 7, 193-199.
[22] Sharma, K.K., and Krishna,H., **1994** Asymptotic sampling distribution of inverse coefficient-of-variation and its applications, *IEEE Transactions on Reliability*, 43, 630-633.
[23] Miller, E.G., **1991**. Asymptotic test statistics for coefficient of variation, *Communications in Statistics Theory and Methods*, 20, 3351-3363.
[24] Ng, K.C., **2006**. Performance of three methods of interval estimation of the coefficient of variation, InterStat Available at http://interstat.statjournals.net/ YEAR/2006/ articles/ 0609002.
[25] Mahmoudvand, R., and Hassani, H., **2009**. Two new confidence intervals for the coefficient of variation in a normal distribution, *Journal of Applied Statistics*, 36, 429-442.

[26] Albatineh, A.N., Kibria, B.M.G., Wilcox, M.L., and Zogheib, B., **2014**. Confidence interval estimation for the population coefficient of variation using ranked set sampling, *Journal of Applied Statistics*, 41, 733-751.
[27] Koopmans, L.H., Owen, D.B., and Rosenblatt, J.I., **1964**. Confidence intervals for the coefficient of variation for the normal and lognormal distributions, *Biometrika*, 51, 25-32.
[28] Verrill, S., **2003**. Confidence bounds for normal and log-normal distribution coefficient of variation, *Research Paper EPL-RP-609*, U.S. Department of Agriculture Madison, Wisconsin, USA.
[29] Buntao, N., and Niwitpong, S., **2012**. Confidence intervals for the difference of coefficients of variation for lognormal distributions and delta-lognormal distributions, *Applied Mathematical Sciences*, 6, 6691-6704.
[30] Vangel, M.G., **1996**. Confidence intervals for a normal coefficient of variation, *The American Statistician*, 50, 21-26.
[31] Curto, J.D., and Pinto, J.C., **2009**. The coefficient of variation asymptotic distribution in the case of non-iid random variables, *Journal of Applied Statistics*, 36, 21-32.
[32] Gulhar, M., Kibria, B.M.G., Albatineh, A.N., and Ahmed, N.U., **2012**. A comparison of some confidence intervals for estimating the population coefficient of variation: a simulation study, *SORT*, 36, 45-68.
[33] Panichkitkosolkul, W., **2013**. Confidence intervals for the coefficient of variation in a normal distribution with a known population mean, *Journal of Probability and Statistics*, Article ID 324940, 11 pages, doi:10.1155/2013/324940
[34] Paksaranuwat, P., and Niwitpong, S., **2009**. Confidence intervals for the difference between two means with missing data following a preliminary test, *Science Asia*, 35, 310-313.
[35] Kabaila, P., and Farchione, D., **2012**. The minimum coverage probability of confidence intervals in regression after a preliminary test, *Journal of Statistical Planning and Inference*, 142, 956-964.
[36] Chiou, P., and Han, C.P., **1995**. Interval estimation of error variance following a preliminary test in one-way random model, *Communications in Statistics-Simulation and Computation*, 24, 817-824.
[37] Panichkitkosolkul, W., and Niwitpong, S., **2011**. On multistep-ahead prediction intervals following unit root tests for a Gaussian AR(1) process with additive outliers, *Applied Mathematical Sciences*, 5, 2297-2316.
[38] Panichkitkosolkul, W., and Niwitpong, S., **2012**. Prediction intervals for the Gaussian autoregressive processes following the unit root tests, *Model Assisted Statistics and Applications: An International Journal*, 7, 1-15.
[39] Ihaka, R., and Gentleman, R.R., **1996**. A Language for data analysis and graphics, *Journal of Computational and Graphical Statistics*, 5, 299-314.
[40] Development Core Team, R., **2014**. An Introduction to R, Vienna, R Foundation for Statistical Computing.
[41] Development Core Team, R., **2014**. A Language and Environment for Statistical Computing. Vienna, R Foundation for Statistical Computing.
[42] White, J., Riethof, M., and Kushnir, I., **1960**. Estimation of microcrystalline wax in beeswax, *Journal of the Association of Official Analytical Chemists*, 43, 781-790.
[43] Rice, J.A., **2007**. *Mathematical Statistics and Data Analysis*, California, Duxbury Press.