



Bayesian Confidence Interval for Ratio of the Coefficients of Variation of Normal Distributions: A Practical Approach in Civil Engineering

Warisa Thangjai ¹, Sa-Aat Niwitpong ², Suparat Niwitpong ^{2*}

¹ Department of Statistics, Faculty of Science, Ramkhamhaeng University, Bangkok, 10240, Thailand.

² Department of Applied Statistics, Faculty of Applied Science, King Mongkut's University of Technology North Bangkok, Bangkok, 10800, Thailand.

Received 21 January 2022; Revised 15 April 2022; Accepted 23 April 2022; Published 10 May 2022

Abstract

The coefficient of variation (CV) is a useful statistical tool for measuring the relative variability between multiple populations, while the ratio of CVs can be used to compare the dispersion. In statistics, the Bayesian approach is fundamentally different from the classical approach. For the Bayesian approach, the parameter is a quantity whose variation is described by a probability distribution. The probability distribution is called the prior distribution, which is based on the experimenter's belief. The prior distribution is updated with sample information. This updating is done with the use of Bayes' rule. For the classical approach, the parameter is quantity and an unknown value, but the parameter is fixed. Moreover, the parameter is based on the observed values in the sample. Herein, we develop a Bayesian approach to construct the confidence interval for the ratio of CVs of two normal distributions. Moreover, the efficacy of the Bayesian approach is compared with two existing classical approaches: the generalised confidence interval (GCI) and the method of variance estimates recovery (MOVER) approaches. A Monte Carlo simulation was used to compute the coverage probability (CP) and average length (AL) of three confidence intervals. The results of a simulation study indicate that the Bayesian approach performed better in terms of the CP and AL. Finally, the Bayesian and two classical approaches were applied to analyse real data to illustrate their efficacy. In this study, the application of these approaches for use in classical civil engineering topics is targeted. Two real data, which are used in the present study, are the compressive strength data for the investigated mixes at 7 and 28 days, as well as the PM_{2.5} air quality data of two stations in Chiang Mai province, Thailand. The Bayesian confidence intervals are better than the other confidence intervals for the ratio of CVs of normal distributions.

Keywords: Air Quality; Compressive Strength; Bayesian Approach; Confidence Interval; Coefficient of Variation; Highest Posterior Density; Variance Estimates Recovery (MOVER).

1. Introduction

In 2018, the air quality in Thailand was worse than in previous years because smog, exacerbated by human activities and comprising particulate matter (PM), a complex mixture of particles and liquid droplets, has steadily increased. PM is defined by size, and PM_{2.5} (2.5 μm) is of particular concern. The main causes of PM are the exhaust from diesel engines, the burning of biomass, and industrial activity. The World Health Organization has stated that the level of PM_{2.5} should not exceed 25 $\mu\text{g}/\text{m}^3$ over a 24-hour mean period [1], whereas the Pollution Control Department in Thailand has set 50 $\mu\text{g}/\text{m}^3$ over a 24-hour mean period as acceptable [2]. However, some air quality monitoring stations

* Corresponding author: suparat.n@sci.kmutnb.ac.th

<http://dx.doi.org/10.28991/CEJ-SP2021-07-010>



© 2021 by the authors. Licensee C.E.J, Tehran, Iran. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).

in Thailand have found levels over three times higher than this in some areas of Thailand. Excessive exposure to PM_{2.5} can lead to health problems such as heart and respiratory diseases, and so people should avoid outdoor activities and wear a facemask during high PM_{2.5} days. The PM_{2.5} level dispersion in different areas can be compared using the coefficient of variation (CV). Also, in civil engineering, compressive strength refers to the ability of a certain material or structural element to withstand loads tending to reduce size. The compressive strength is often tested to evaluate the actual mix meets the requirements of the design specification. In order to conduct the compressive strength test, the compressive strength is used at the age of 7, 14, 28, and 56 days. The coefficient of variation (CV) has been used to describe the compressive strength and can be used to compare the compressive strength variability in two or more different days for the investigated mixes. In this study, the Bayesian confidence interval for the ratio of the coefficients of variation (CV) of normal distributions for these two applied issues in civil engineering will be evaluated.

The ratio of the standard deviation and the mean of a population is called the CV, which is a metric that is free of the unit of measurement. The larger the CV value, the greater the dispersion, with the lowest value indicating the lowest risk. The CV has been widely used in many fields, such as atmospheric and medical sciences. For instance, the CV has been used to compare the dispersion of PM_{2.5} in many areas [3] and blood sample measurements taken from the various laboratories [4]. Several methods for estimating the confidence interval of the CV have been suggested. Vangel [5] analyzed the small-sample distribution of a class of approximate pivotal quantities for the CV of a normal distribution. Tian [6] presented a new approach for making inferences about the common CV of several independent normal distributions using the concepts of generalized variables. Mahmoudvand and Hassani [7] introduced an approximately unbiased estimator for the population CV of a normal distribution. Panichkitkosolkul [8] improved the confidence interval for the CV of a normal distribution by replacing the sample CV in Vangel's confidence interval with the maximum likelihood estimator. Moreover, Wongkhao et al. [9] proposed confidence intervals for the ratio of two independent CVs of normal distributions based on the generalized confidence interval (GCI) and method of variance estimates recovery (*MOVER*) approaches. Kalkur and Rao [10] provided a Bayesian estimator for the CV and inverse CV of a normal distribution. Thangjai et al. [11] proposed adjusted GCIs for the common CV of several normal distributions. Thangjai et al. [12] presented Bayesian confidence intervals for the means of normal distributions with unknown coefficients of variation. Thangjai et al. [13] presented confidence intervals for the single CV of a normal distribution and the difference between the CVs of two normal distributions.

Statistics can be divided into two different techniques, which are the Bayesian approach and the classical approach [14]. There are differences in Bayesian and classical approaches in methods and analysis. For the Bayesian approach, the parameter is a variable that has a probability distribution. Prior probability is the probability assigned to an event before the arrival of evidence information. After receiving the information, the prior probability is updated with the sample information using the Bayes' rule. The updated prior probability is called the posterior probability. The posterior distribution is used for constructing the confidence interval of the parameter. The Bayesian approach has been successfully utilised to establish the confidence interval for parameters of interest. For example, Thangjai et al. [12] presented the Bayesian confidence interval for the mean of a normal distribution with an unknown CV and the Bayesian confidence interval for the difference between two means of normal distributions with unknown coefficients of variation. Maneerat and Niwitpong [15] compared medical care costs using Bayesian credible intervals for the ratio of means of delta-lognormal distributions. Maneerat et al. [16] proposed a Bayesian approach to construct the interval estimation for the difference between variances of delta-lognormal distributions. Moreover, Thangjai et al. [13] proposed the Bayesian confidence interval for the CV of a normal distribution and the Bayesian confidence interval for the difference between the CVs of normal distributions. In addition, Thangjai et al. [17] constructed the Bayesian confidence interval for the CV of a log-normal distribution and the Bayesian confidence interval for the difference between CVs of log-normal distributions. For the classical approach, the parameter is an unknown constant. The best interval estimation of the parameter was obtained from the results of experiments. The classical approach includes many methods, such as the GCI method and the *MOVER* method. The GCI approach has been widely used to estimate the confidence interval for parameters. For instance, Tian [6] introduced the GCI for constructing the confidence interval for the common CV of normal distributions.

Tian and Wu [18] constructed the confidence interval for the common mean of log-normal distributions using the GCI approach. Ye et al. [19] presented the confidence interval for the common mean of inverse Gaussian distributions based on the *GCI* approach. Niwitpong and Wongkhao [20] constructed the confidence interval for the difference between the inverse means of normal distributions. Thangjai et al. [21] presented the GCI approach to construct the confidence interval for the mean and difference between means of normal distributions with unknown coefficients of variation. Thangjai et al. [22] proposed the GCI approach to construct the confidence interval for the difference between variances of one-parameter exponential distributions. Thangjai and Niwitpong [23] used the GCI approach to estimate the confidence intervals for the signal-to-noise ratio and difference between signal-to-noise ratios of log-normal distributions. Moreover, the *MOVER* approach has been widely used to construct the confidence interval for the parameter. For example, Donner and Zou [24] introduced the *MOVER* approach for constructing the confidence interval for a function of the normal standard deviation. Niwitpong [25] proposed a confidence interval for the difference

between CVs of normal distribution with bounded parameters based on the MOVER approach. Wongkhao et al. [26] constructed the confidence intervals for the ratio of CVs of normal distributions. Thangjai et al. [27] proposed the MOVER approach to construct the simultaneous confidence intervals for all differences in CVs of log-normal distributions.

It is of theoretical importance to develop approaches for confidence interval estimation for the ratio of the CVs of two independent normal distributions. In this study, we developed an approach using the concept of Bayesian statistics. The problem of constructing the confidence interval for the CV of a normal distribution based on the Bayesian approach has been considered by many authors (for example, see Kalkur and Rao [10] and Thangjai et al. [13]). In the present study, the Bayesian approach was used to construct the confidence interval for the ratio of CVs of normal distributions, followed by comparing its performance with that of the GCI and MOVER approaches proposed by Wongkhao et al. [9].

This study is organised as follows: In Section 2, the Bayesian confidence interval for the ratio of CVs of normal distributions is presented. In Section 3, simulation results of estimated coverage probabilities (CPs) and average lengths (ALs) are obtained by using Monte Carlo studies. In addition, a Bayesian approach is applied to a real-life dataset. Concluding remarks are summarised in Section 4.

2. Bayesian Confidence Interval for Ratio of the CVs

Let $X = (X_1, X_2, \dots, X_n)$ be a random sample of size n from a normal distribution with mean μ_x and variance σ_x^2 . Also, let $Y = (Y_1, Y_2, \dots, Y_m)$ be a random sample of size m from a normal distribution with mean μ_y and variance σ_y^2 . The CV is the ratio of standard deviation to the mean. The CVs of X and Y are defined as:

$$\theta_x = \frac{\sigma_x}{\mu_x} \tag{1}$$

And;

$$\theta_y = \frac{\sigma_y}{\mu_y} \tag{2}$$

The ratio of CVs is defined as:

$$\delta = \frac{\theta_x}{\theta_y} = \frac{\sigma_x / \mu_x}{\sigma_y / \mu_y} = \frac{\sigma_x \mu_y}{\sigma_y \mu_x} \tag{3}$$

Let \bar{X} , \bar{Y} , S_x , and S_y be the sample means and sample standard deviations of X and Y , respectively. The maximum likelihood estimator of δ is obtained by:

$$\hat{\delta} = \frac{\hat{\theta}_x}{\hat{\theta}_y} = \frac{S_x / \bar{X}}{S_y / \bar{Y}} = \frac{S_x \bar{Y}}{S_y \bar{X}} \tag{4}$$

The confidence interval for ratio of CVs of normal distribution is considered using the Bayesian approach. The Bayesian approach derives the posterior probability which is based on the likelihood function and the prior probability [28]. In this paper, the highest posterior density (HPD) interval is used to construct the Bayesian confidence interval.

Definition: Suppose that $X = (X_1, X_2, \dots, X_n)$ denotes a random sample. Let $x = (x_1, x_2, \dots, x_n)$ be observed value of $X = (X_1, X_2, \dots, X_n)$. Also, let $p(\theta|x)$ be a posterior density function. Box and Tiao [29] defines that a region \mathfrak{R} in the parameter space of θ is called the HPD region of content $(1 - \alpha)$ if the following two conditions are satisfied:

- (i) $\Pr(\theta \in \mathfrak{R} | x) = 1 - \alpha$
- (ii) For $\theta_1 \in \mathfrak{R}$ and $\theta_2 \notin \mathfrak{R}$, $p(\theta_1 | x) \geq p(\theta_2 | x)$

The HPD interval is based on various priors such as Jeffreys prior and reference prior. In this paper, the independence Jeffreys prior is used to construct the HPD interval based on Bayesian framework. The independence Jeffreys priors are:

$$P(\mu_x, \sigma_x^2) \propto \frac{1}{\sigma_x^2} \tag{5}$$

And;

$$P(\mu_y, \sigma_y^2) \propto \frac{1}{\sigma_y^2} \tag{6}$$

Considering the mean μ_x , the posterior distribution of μ_x given σ_x^2 and x is normal distribution which is defined as:

$$\mu_x | \sigma_x^2, x \sim N(\hat{\mu}_x, \frac{\sigma_x^2}{n}) \tag{7}$$

Similarly, the posterior distribution of μ_y given σ_y^2 and y is normal distribution which is defined as:

$$\mu_y | \sigma_y^2, y \sim N(\hat{\mu}_y, \frac{\sigma_y^2}{m}) \tag{8}$$

Considering the variance σ_x^2 , the posterior distribution for σ_x^2 given x is inverse gamma distribution which is defined as:

$$\sigma_x^2 | x \sim IG(\frac{n-1}{2}, \frac{(n-1)s_x^2}{2}) \tag{9}$$

The posterior distribution for σ_y^2 given y is inverse gamma distribution which is defined as:

$$\sigma_y^2 | y \sim IG(\frac{m-1}{2}, \frac{(m-1)s_y^2}{2}) \tag{10}$$

Bayesian approach uses posterior distribution of ratio of CVs to construct the confidence interval through Monte Carlo simulation. Since the posterior distribution of ratio of CVs of normal distributions is defined as:

$$\delta_{BS} = \frac{\sqrt{\sigma_x^2 / \mu_x}}{\sqrt{\sigma_y^2 / \mu_y}} = \frac{\sigma_x / \mu_x}{\sigma_y / \mu_y} = \frac{\sigma_x \mu_y}{\sigma_y \mu_x} \tag{11}$$

where μ_x and μ_y are simulated from the posterior distributions as defined in Equations 7 and 8, respectively. Moreover, σ_x^2 and σ_y^2 are simulated from the posterior distributions as defined in Equations 9 and 10, respectively. Therefore, the 100(1 - α)% two-sided confidence interval for the ratio of CVs based on the Bayesian approach is obtained as:

$$CI_{\delta_{BS}} = [L_{\delta_{BS}}, U_{\delta_{BS}}] \tag{12}$$

Where $L_{\delta_{BS}}$ and $U_{\delta_{BS}}$ are the lower limit and the upper limit of the shortest 100(1 - α)% highest posterior density interval of δ_{BS} , respectively.

The following algorithm is used to construct the Bayesian confidence interval for the ratio of CVs of normal distributions:

Algorithm 1.

- Step 1:** Generate $\mu_x | \sigma_x^2, x$ form Equation 7 and $\mu_y | \sigma_y^2, y$ form Equation 8;
- Step 2:** Generate $\sigma_x^2 | x$ form Equation 9 and $\sigma_y^2 | y$ form Equation 10;
- Step 3:** Compute δ_{BS} from Equation 11;
- Step 4:** Repeat step 1 - step 3, a total q times and obtain an array of δ_{BS} 's;
- Step 5:** Compute $L_{\delta_{BS}}$ and $U_{\delta_{BS}}$ from Equation 12.

Here, two approaches of Wongkhao et al. [9] are briefly discussed to construct the confidence intervals for the ratio of CVs of normal distributions. The approaches are the *GCI* approach and *MOVER* approach.

Firstly, the *GCI* approach uses the generalized pivotal quantity to construct the confidence interval. The generalized pivotal quantities for the CVs of X and Y are defined as:

$$R_{\theta_x} = \frac{\sqrt{(n-1)s_x^2 / \chi_{n-1}^2}}{\bar{x} - t_{n-1} s_x / \sqrt{n}} \tag{13}$$

And;

$$R_{\theta_y} = \frac{\sqrt{(m-1)s_y^2 / \chi_{m-1}^2}}{\bar{y} - t_{m-1} s_y / \sqrt{m}} \tag{14}$$

where \bar{x} , \bar{y} , s_x , and s_y are the observed values of \bar{X} , \bar{Y} , S_X , and S_Y , respectively. Also, t_{n-1} and t_{m-1} are the student's t distributions with $n-1$ and $m-1$ degrees of freedom, respectively. Moreover, χ_{n-1}^2 and χ_{m-1}^2 are the chi-squared distributions with $n-1$ and $m-1$ degrees of freedom, respectively.

The generalized pivotal quantity for the ratio of CVs is defined as:

$$R_\delta = \frac{R_{\theta_x}}{R_{\theta_y}} \tag{15}$$

where R_{θ_x} and R_{θ_y} are defined in Equations 13 and 14, respectively.

Therefore, the $100(1-\alpha)\%$ two-sided confidence interval for the ratio of CVs based on the GCI approach is defined as:

$$CI_{\delta,GCI} = [L_{\delta,GCI}, U_{\delta,GCI}] \tag{16}$$

where $L_{\delta,GCI}$ and $U_{\delta,GCI}$ are the $100(\alpha/2)$ -th and the $100(1-\alpha/2)$ -th percentiles of R_δ as defined in Equation 15, respectively.

Next, the lower and upper limits of confidence interval for CV of X are:

$$l_x = \frac{S_x}{d_x} (\bar{X} - \sqrt{\max(0, \bar{X}^2 + a_x d_x (a_x - 2))}) \tag{17}$$

And;

$$u_x = \frac{S_x}{d_x} (\bar{X} + \sqrt{\max(0, \bar{X}^2 + b_x d_x (b_x - 2))}) \tag{18}$$

where $a_x = \sqrt{(n-1) / \chi_{1-\alpha/2, n-1}^2}$, $b_x = \sqrt{(n-1) / \chi_{\alpha/2, n-1}^2}$, and $d_x = \bar{X}^2 - z_{\alpha/2}^2 S_x^2 / n$.

Similarly, the lower and upper limits of confidence interval for CV of Y are:

$$l_y = \frac{S_y}{d_y} (\bar{Y} - \sqrt{\max(0, \bar{Y}^2 + a_y d_y (a_y - 2))}) \tag{19}$$

And;

$$u_y = \frac{S_y}{d_y} (\bar{Y} + \sqrt{\max(0, \bar{Y}^2 + b_y d_y (b_y - 2))}) \tag{20}$$

where $a_y = \sqrt{(m-1) / \chi_{1-\alpha/2, m-1}^2}$, $b_y = \sqrt{(m-1) / \chi_{\alpha/2, m-1}^2}$, and $d_y = \bar{Y}^2 - z_{\alpha/2}^2 S_y^2 / m$.

The lower and upper limits of confidence interval for the ratio of CVs are:

$$L_{\delta,MOVER} = \frac{\hat{\theta}_x \hat{\theta}_y - \sqrt{(\hat{\theta}_x \hat{\theta}_y)^2 - l_x u_y (2\hat{\theta}_x - l_x)(2\hat{\theta}_y - u_y)}}{u_y (2\hat{\theta}_y - u_y)} \tag{21}$$

And;

$$U_{\delta,MOVER} = \frac{\hat{\theta}_x \hat{\theta}_y + \sqrt{(\hat{\theta}_x \hat{\theta}_y)^2 - u_x l_y (2\hat{\theta}_x - u_x)(2\hat{\theta}_y - l_y)}}{l_y (2\hat{\theta}_y - l_y)} \tag{22}$$

where $\hat{\theta}_x = S_x / \bar{X}$, $\hat{\theta}_y = S_y / \bar{Y}$, and l_x , u_x , l_y , and u_y are defined in Equations 17 to 20, respectively.

Therefore, the $100(1-\alpha)\%$ two-sided confidence interval for the ratio of CVs based on MOVER approach is defined as:

$$CI_{\delta,MOVER} = [L_{\delta,MOVER}, U_{\delta,MOVER}] \tag{23}$$

Where $L_{\delta,MOVER}$ and $U_{\delta,MOVER}$ are defined in Equations 21 and 22, respectively.

3. Results and Discussions

3.1. Simulation Study

A Monte Carlo simulation study was carried out to analyze the performance of the Bayes estimator and to compare it with those of two classical estimators, GCI and MOVER estimators for constructing confidence intervals for the ratio of CVs of two normal distributions. In the study, we set $\mu_X = \mu_Y = 1.0$, $\sigma_X = 1.0$, and $\sigma_Y = 0.3, 0.5, 0.7, 0.9, 1.0, 1.1, 1.3, 1.5, 1.7, 2.0$. For each set of parameters, 5000 random samples were generated and 2500 simulation replications were used for each of the random samples. The following algorithm is used to estimate the CP and average length:

Algorithm 2.

For a given $m, \mu_X, \mu_Y, \sigma_X, \sigma_Y$, and δ

For $h = 1$ to M

Generate x from $N(\mu_X, \sigma_X^2)$

Generate y from $N(\mu_Y, \sigma_Y^2)$

Calculate $\bar{x}, \bar{y}, S_X, S_Y, \theta_X, \theta_Y$, and δ

Construct $CI_{\delta,BS(h)} = [L_{\delta,BS(h)}, U_{\delta,BS(h)}]$

Construct $CI_{\delta,GCI(h)} = [L_{\delta,GCI(h)}, U_{\delta,GCI(h)}]$

Construct $CI_{\delta,MOVER(h)} = [L_{\delta,MOVER(h)}, U_{\delta,MOVER(h)}]$

Record whether or not all the values of δ fall in their corresponding confidence intervals

Compute $U_{(h)} - L_{(h)}$

End h loop

Compute the CP and the AL for each confidence interval.

Table 1 summarizes the empirical CPs and ALs of the *GCI*, *MOVER*, and Bayesian confidence intervals. According to Thangjai et al. [30], the best confidence interval will have the CP in the range [0.9440, 0.9560] at the 95% confidence level and the confidence interval has the shortest average length. The results show that the CPs of three confidence intervals are in the range [0.9440, 0.9560]. Moreover, only the Bayesian confidence interval provided CPs greater than the nominal confidence level of 0.95 for all cases. In addition, the ALs of the Bayesian confidence interval were shorter than the others. Therefore, the results clearly indicate that the Bayesian approach performed satisfactorily in small, moderate, and large sample sizes.

Table 1. The CPs and ALs of 95% two-sided confidence intervals for ratio of CVs of normal distributions

n	m	$\frac{\sigma_Y}{\sigma_X}$	CI _{δ,GCI}		CI _{δ,MOVER}		CI _{δ,BS}	
			CP	AL	CP	AL	CP	AL
30	30	0.3	0.9466	2.8284	0.9432	3.6672	0.9554	2.5226
		0.5	0.9504	2.2750	0.9474	2.1768	0.9578	2.0425
		0.7	0.9494	1.9886	0.9442	2.1365	0.9544	1.7955
		0.9	0.9508	1.7886	0.9458	1.7114	0.9596	1.6182
		1.0	0.9516	1.7364	0.9462	1.8487	0.9548	1.5758
		1.1	0.9524	1.6573	0.9480	1.5829	0.9580	1.5066
		1.3	0.9460	1.5965	0.9426	1.5839	0.9538	1.4554
		1.5	0.9594	1.5047	0.9534	1.5399	0.9638	1.3750
		1.7	0.9498	1.4705	0.9422	1.4253	0.9588	1.3502
		2.0	0.9516	1.3836	0.9474	1.3706	0.9608	1.2772

n	m	$\frac{\sigma_y}{\sigma_x}$	CI _{δ.GCI}		CI _{δ.MOVER}		CI _{δ.BS}	
			CP	AL	CP	AL	CP	AL
50	50	0.3	0.9506	1.8471	0.9486	1.7940	0.9544	1.7418
		0.5	0.9462	1.4792	0.9460	1.4376	0.9552	1.3976
		0.7	0.9524	1.2838	0.9512	1.2509	0.9536	1.2190
		0.9	0.9484	1.1788	0.9466	1.1488	0.9552	1.1215
		1.0	0.9458	1.1286	0.9436	1.1004	0.9524	1.0749
		1.1	0.9532	1.0996	0.9490	1.0712	0.9562	1.0472
		1.3	0.9428	1.0400	0.9406	1.0143	0.9458	0.9934
		1.5	0.9414	1.0007	0.9386	0.9760	0.9472	0.9568
		1.7	0.9480	0.9658	0.9458	0.9422	0.9540	0.9250
		2.0	0.9500	0.9147	0.9460	0.8925	0.9586	0.8780
30	50	0.3	0.9502	2.6646	0.9446	2.5608	0.9562	2.3574
		0.5	0.9508	2.1207	0.9448	2.0204	0.9558	1.8757
		0.7	0.9556	1.8284	0.9484	1.7505	0.9568	1.6272
		0.9	0.9534	1.6420	0.9502	1.5807	0.9604	1.4658
		1.0	0.9520	1.6043	0.9468	1.5702	0.9558	1.4347
		1.1	0.9524	1.5237	0.9464	1.4562	0.9540	1.3693
		1.3	0.9536	1.4344	0.9488	1.3507	0.9580	1.2899
		1.5	0.9524	1.3335	0.9476	1.3102	0.9552	1.2032
		1.7	0.9572	1.2943	0.9508	1.2536	0.9594	1.1702
		2.0	0.9526	1.2078	0.9484	1.1445	0.9568	1.0979
100	100	0.3	0.9518	1.1757	0.9518	1.1620	0.9532	1.1418
		0.5	0.9474	0.9512	0.9476	0.9397	0.9520	0.9245
		0.7	0.9486	0.8329	0.9464	0.8233	0.9492	0.8104
		0.9	0.9492	0.7595	0.9498	0.7511	0.9506	0.7402
		1.0	0.9480	0.7315	0.9462	0.7233	0.9488	0.7130
		1.1	0.9500	0.7058	0.9486	0.6980	0.9504	0.6886
		1.3	0.9526	0.6729	0.9514	0.6653	0.9544	0.6566
		1.5	0.9516	0.6423	0.9502	0.6354	0.9514	0.6279
		1.7	0.9540	0.6231	0.9516	0.6165	0.9534	0.6096
		2.0	0.9550	0.5937	0.9520	0.5871	0.9596	0.5808
50	100	0.3	0.9488	1.6807	0.9470	1.6195	0.9520	1.5710
		0.5	0.9460	1.3405	0.9446	1.2943	0.9540	1.2572
		0.7	0.9454	1.1589	0.9436	1.1205	0.9520	1.0888
		0.9	0.9482	1.0414	0.9474	1.0074	0.9540	0.9792
		1.0	0.9508	1.0027	0.9484	0.9714	0.9546	0.9446
		1.1	0.9524	0.9534	0.9512	0.9234	0.9578	0.8990
		1.3	0.9460	0.8999	0.9438	0.8732	0.9528	0.8500
		1.5	0.9428	0.8557	0.9408	0.8306	0.9482	0.8088
		1.7	0.9484	0.8127	0.9468	0.7896	0.9554	0.7695
		2.0	0.9534	0.7667	0.9526	0.7455	0.9564	0.7282

In this study, the Bayesian approach performs satisfactorily for constructing the confidence interval for the ratio of CVs of normal distributions. It was similar to the results of Thangjai et al. [12, 13, 17], Maneerat and Niwitpong [15], and Maneerat et al. [16]. Furthermore, the GCI and MOVER approaches do not perform well for constructing the confidence intervals for the ratio of CVs of normal distributions. However, the GCI and MOVER approaches are recommended to construct the simultaneous confidence intervals for all differences in means of normal distributions with unknown CVs [31].

3.2. Empirical Application

Example 1: Datasets of compressive strength for the investigated mixes at 7 and 28 days obtained from Ali et al. [32] are reported in Table 2. The sample statistics of the compressive strength values at 7 days are $n = 9$, $\bar{x} = 41.6444$, $s_x^2 = 623.1003$, and $\hat{\theta}_x = 0.5994$, whereas the sample statistics of the compressive strength values at 28 days are $m = 9$, $\bar{y} = 44.5667$, $s_y^2 = 662.805$, and $\hat{\theta}_y = 0.5777$. The histogram and normal QQ-plots of the data are presented in Figures 1 and 2, respectively. The Shapiro-Wilk normality test with p -values 0.8344 and 0.6857 for 7 days and 28 days, respectively. According to the p -values, the data sets are fitted the normal distribution. The true ratio of CVs was 1.0376. The 95% two-sided confidence intervals were [0.3529, 3.0812] with an interval length of 2.7283 using GCI, [0.3938, 2.7566] with an interval length of 2.3628 using MOVER, and [0.2096, 2.4502] with an interval length of 2.2406 using the Bayesian approach. Thus, it can be seen that all of them covered the ratio of CVs for the two populations. The results are in agreement with the simulation results in that the Bayesian confidence interval had the shortest length.

Table 2. Compressive strength of the investigated mixes 7 days and 28 days

Mix	7 days	28 days
1	2.7	3.0
2	16.0	17.5
3	21.3	22.7
4	40.5	45.5
5	44.1	46.0
6	45.3	50.9
7	56.0	64.8
8	71.5	72.5
9	77.4	78.2

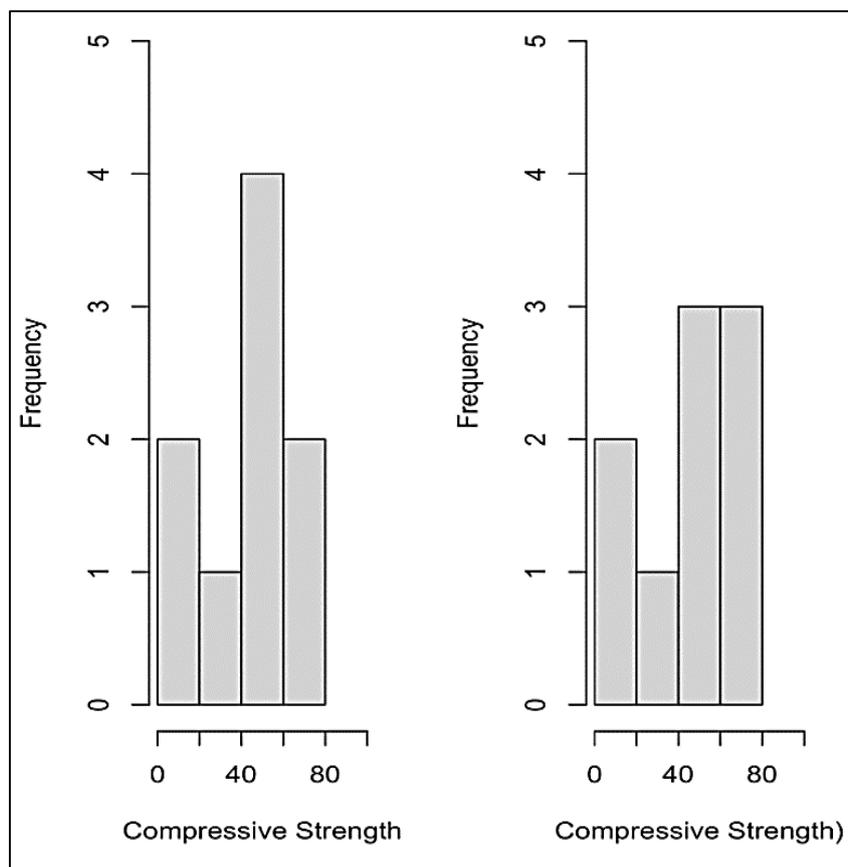


Figure 1. Histogram plots of compressive strength values at 7 days and at 28 days

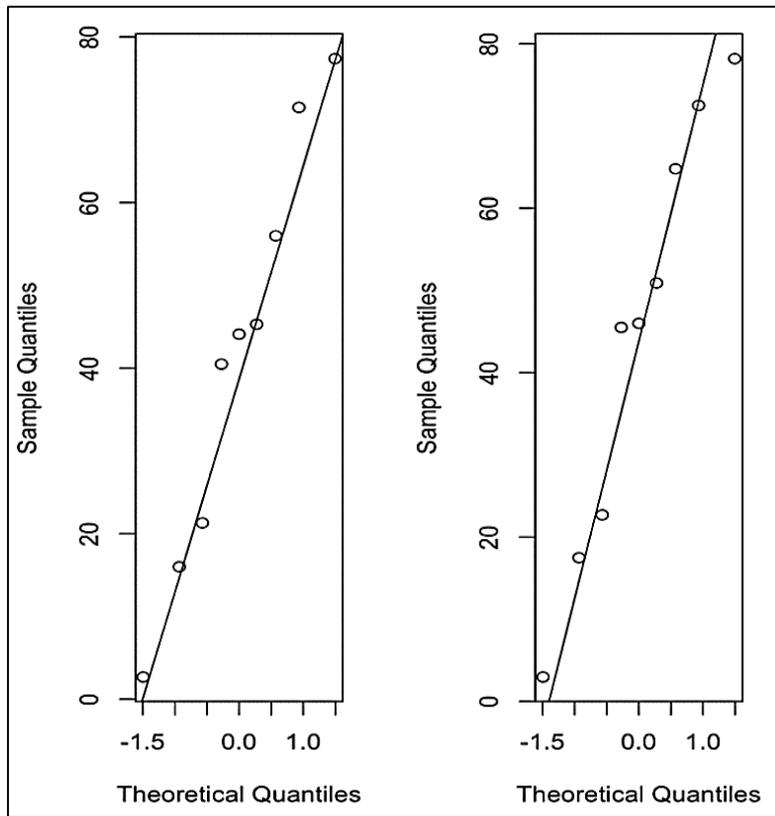


Figure 2. The normal QQ-plots of compressive strength values at 7 days and at 28 days

Example 2: Thailand is divided into six geographical regions. Figure 3 presents geographical regions of Thailand [33]. There are north region, northeast region, central region, east region, west region, and south region. The northern Thailand is region with frequent air pollution problems.



Figure 3. Geographical regions of Thailand [33]

Chiang Mai province is the hub of the northern Thailand. This example is interested in Mueang Chiang Mai district which is the capital district of Chiang Mai province in northern Thailand. This is because the PM2.5 has been the seasonal problem in the North of Thailand likes Chiang Mai province. The PM2.5 appears from January to April. However, the extremely dry conditions increase the magnitude of forest fires in March. Real data examples of PM2.5 data are used to illustrate the GCI, MOVER, and Bayesian approaches. The data were reported by the *Regional Environment Office 1* (<http://www.reo01.mnre.go.th>).

The PM2.5 data of two stations in Mueang Chiang Mai district, Chiang Mai province from 1 March 2018 to 30 April 2018 are reported in Table 3. The data sets consist of 54 measurements in Tambon Chang Phueak station and 56 measurements in Tambon Sri Phum station. The sample statistics of Tambon Chang Phueak station are $\bar{x} = 59.3889$, $s_x^2 = 528.0161$, and $\hat{\theta}_x = 0.3869$, whereas the sample statistics of Tambon Sri Phum station are $\bar{y} = 59.6964$, $s_y^2 = 521.0902$, and $\hat{\theta}_y = 0.3824$. The histogram and normal QQ-plots of the data are presented in Figures 4 and 5, respectively.

Table 3. PM2.5 data of Tambon Chang Phueak and Tambon Sri Phum Stations ($\mu\text{g}/\text{m}^3$)

Tambon Chang Phueak Station					Tambon Sri Phum Station				
82	74	58	46	66	79	72	57	59	67
89	68	100	72	38	94	65	102	73	39
59	67	60	82	79	56	73	59	83	74
72	68	78	70	57	69	64	73	71	60
77	104	105	71	47	74	108	107	72	45
52	34	41	33	26	52	34	43	37	28
23	18	25	41	72	20	14	27	43	76
69	71	89	94	64	69	76	69	76	86
33	27	35	46	56	76	61	33	25	34
76	75	85	69	44	50	58	80	73	82
38	32	25	25		71	45	37	34	22
					17				

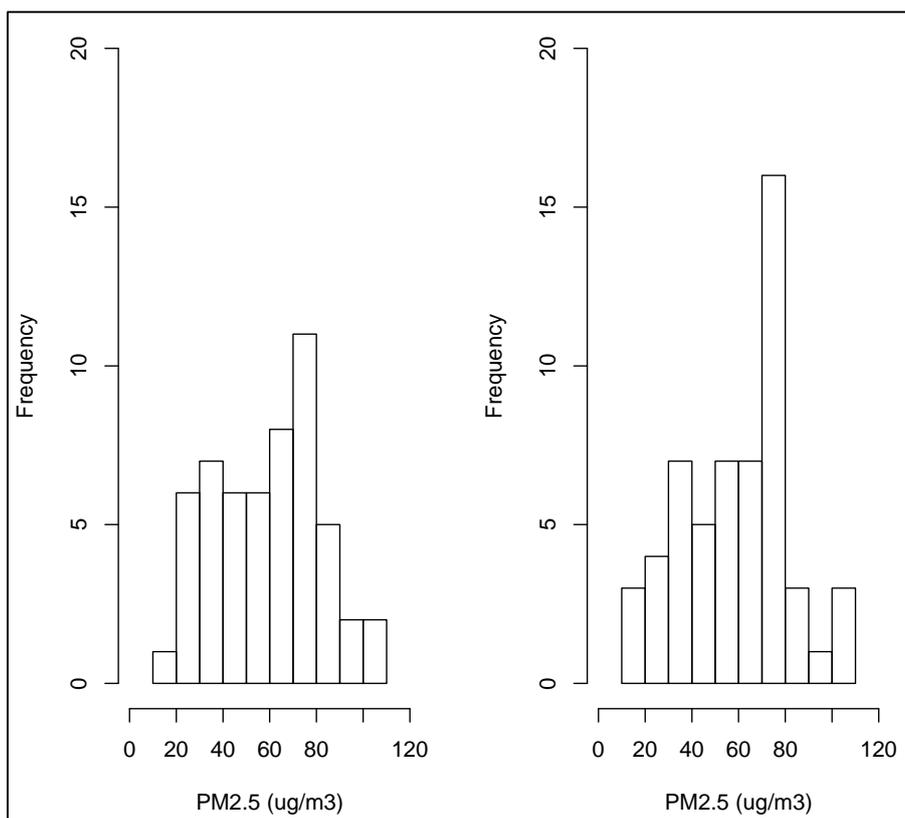


Figure 4. Histogram plots of PM2.5 data of Tambon Chang Phueak and Tambon Sri Phum stations

The Shapiro-Wilk normality test with p-values 0.08361 and 0.1077 for Tambon Chang Phueak and Tambon Sri Phum stations, respectively. According to the p-values, the data sets are fitted the normal distribution. The true ratio of CVs was 1.0118. The data were used to establish the confidence interval for ratio of CVs using GCI, MOVER, and Bayesian approaches. Firstly, the 95% two-sided GCI was [0.7471, 1.3850] with an interval length of 0.6379. In addition, the 95% two-sided MOVER confidence interval was [0.7447, 1.3771] with an interval length of 0.6324. Finally, the 95% two-sided Bayesian confidence interval was [0.7207, 1.3306] with an interval length of 0.6099. It is also true to claim that all confidence intervals for ratio of CVs cover the population ratio of CVs. The results confirm the simulation results in the previous section that the Bayesian confidence interval has the shortest length. Therefore, the Bayesian approach is recommended to establish the confidence interval for ratio of CVs of normal distributions.

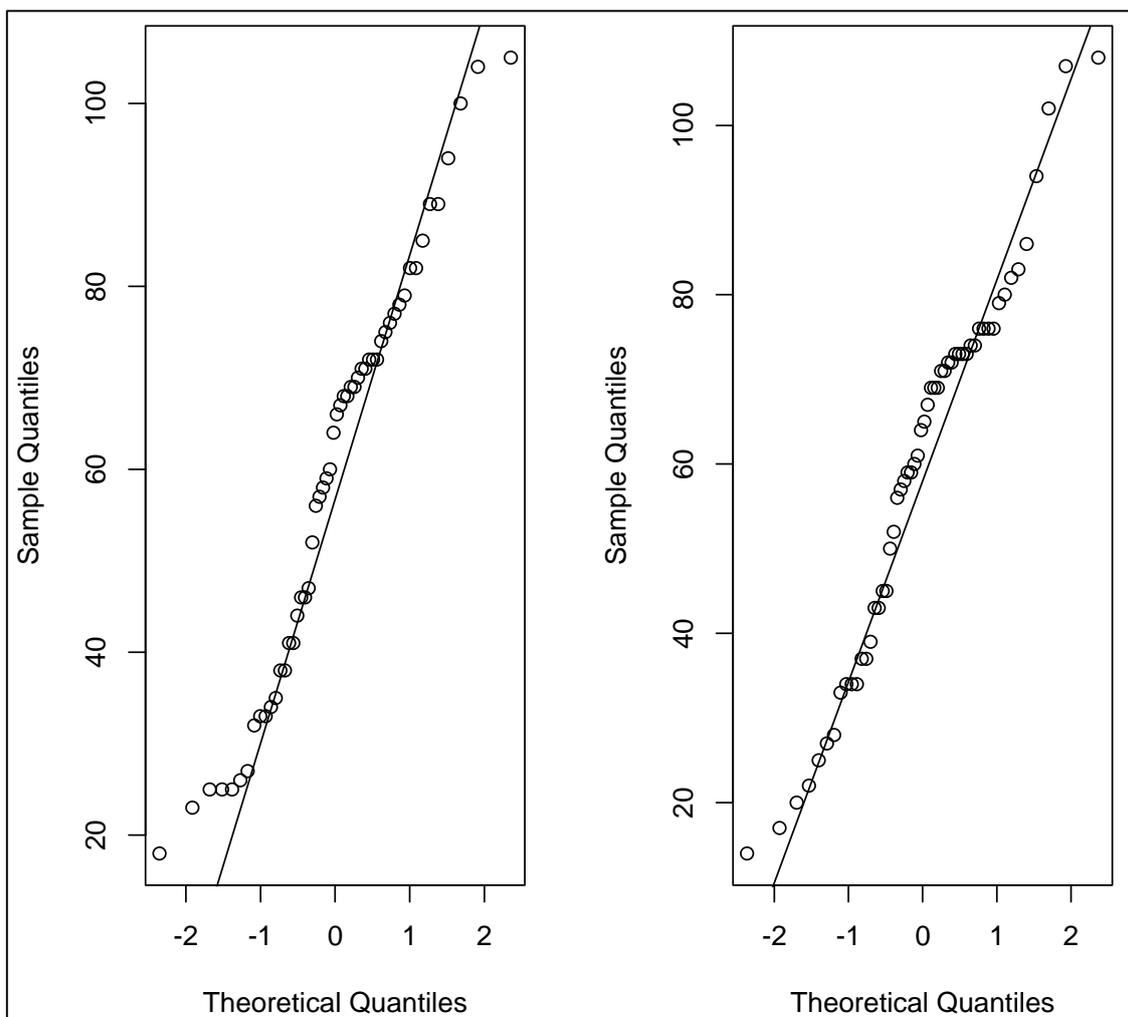


Figure 5. The normal QQ-plots of PM2.5 data of Tambon Chang Phueak and Tambon Sri Phum stations

4. Conclusion

The CV is used to measure the dispersion of a probability distribution. The lower CV describes the lower dispersion or lower risk, whereas the larger CV describes the greater dispersion or greater risk. It is commonly used in many fields, such as engineering and environmental data. The ratio of CVs is used to compare the dispersion of two populations. This paper is interested in the ratio of CVs of two normal populations using engineering and environmental data. For civil engineering data, the ratio of the CVs is used to compare the dispersion of the compressive strength at the ages of 7 and 28 days. Moreover, the ratio of the CVs is evaluated to describe the spread of PM2.5 data for Tambon Chang Phueak and Tambon Sri Phum Stations in Mueang Chiang Mai district, Chiang Mai province. In practice, behavioural models have more often been derived by using the classical approach rather than the Bayesian approach. This paper is an extension of previous works by Wongkhao et al. [9] and Thangjai et al. [13]. Wongkhao et al. [9] proposed the classical GCI and MOVER approaches for constructing the confidence intervals for the ratio of CVs of two normal distributions. Moreover, Thangjai et al. [13] proposed the Bayesian approach for the CV and difference of CVs of normal distributions. Therefore, we proposed a Bayesian approach for confidence interval estimation for the ratio of the CVs of two normal distributions. The Bayesian approach is compared with the existing classical approaches: the GCI and the

MOVER approach. The Bayesian and GCI approaches use software packages to estimate the confidence intervals, whereas the MOVER approach uses a formula to construct the confidence interval. The simulation results indicate that the Bayesian approach performed better than the classical approaches and is thus recommended for constructing the confidence interval for the ratio of the CVs of two normal populations. The results of this investigation were similar to those of Thangjai and Niwitpong [3] and Thangjai et al. [12, 13]. Further research will be proposed with other approaches for comparison.

5. Declarations

5.1. Author Contributions

Conceptualization, W.T. and S.N.; methodology, S-A.N.; software, W.T.; validation, W.T., S-A.N. and S.N.; formal analysis, W.T.; investigation, S-A.N.; resources, S.N.; data curation, S-A.N.; writing—original draft preparation, W.T.; writing—review and editing, W.T.; visualization, W.T.; supervision, S-A.N.; project administration, S.N.; funding acquisition, S.N. All authors have read and agreed to the published version of the manuscript.

5.2. Data Availability Statement

The data presented in this study are available on request from the corresponding author.

5.3. Funding

This research was funded by King Mongkut's University of Technology North Bangkok. Grant Number: KMUTNB-65-BASIC-35.

5.4. Conflicts of Interest

The authors declare no conflict of interest.

6. References

- [1] World Health Organization (WHO). (2016). WHO air quality guidelines for particulate matter, ozone, nitrogen dioxide and sulfur dioxide: global update 2005: summary of risk assessment. WHO Press, Geneva, Switzerland. Available online: http://whqlibdoc.who.int/hq/2006/WHO_SDE_PHE_OEH_06.02_eng.pdf (accessed on January 2022).
- [2] Pollution Control Department, (2019). Booklet on Thailand state of pollution 2018. Pollution Control Department, Ministry of Natural Resources and Environment. S. Mongkol Press Limited Partnership, Bangkok, Thailand. Available online: <https://www.matchlink.asia/business/page/Thailand/BANGKOK/63791-s-mongkol-press-limited-partnership.html> (accessed on January 2022).
- [3] Thangjai, W., & Niwitpong, S. (2020). Comparing particulate matter dispersion in Thailand using the Bayesian Confidence Intervals for ratio of coefficients of variation. In *Statistics in Transition*, 21(5), 41–60. doi:10.21307/STATTRANS-2020-054.
- [4] Mallard, A. R., Hollekim-Strand, S. M., Ingul, C. B., & Coombes, J. S. (2020). High day-to-day and diurnal variability of oxidative stress and inflammation biomarkers in people with type 2 diabetes mellitus and healthy individuals. *Redox Report*, 25(1), 64–69. doi:10.1080/13510002.2020.1795587
- [5] Vangel, M. G. (1996). Confidence Intervals for a Normal Coefficient of Variation. *The American Statistician*, 50(1), 21. doi:10.2307/2685039
- [6] Tian, L. (2005). Inferences on the common coefficient of variation. *Statistics in Medicine*, 24(14), 2213–2220. doi:10.1002/sim.2088.
- [7] Mahmoudvand, R., & Hassani, H. (2009). Two new confidence intervals for the coefficient of variation in a normal distribution. *Journal of Applied Statistics*, 36(4), 429–442. doi:10.1080/02664760802474249.
- [8] Panichkitkosolkul, W. (2009). Improved confidence intervals for a coefficient of variation of a normal distribution. *Thailand statistician*, 7(2), 193-199.
- [9] Wongkhao, A., Niwitpong, S. A., & Niwitpong, S. (2015). Confidence intervals for the ratio of two independent coefficients of variation of normal distribution. *Far East Journal of Mathematical Sciences*, 98(6), 741–757. doi:10.17654/FJMSNov2015_741_757.
- [10] Kalkur, T. A., & Rao, A. Bayes estimator for coefficient of variation and inverse coefficient of variation for the normal distribution. *International Journal of Statistics and Systems*, 12(4), 721–732.
- [11] Thangjai, W., Niwitpong, S. A., & Niwitpong, S. (2020). Adjusted generalized confidence intervals for the common coefficient of variation of several normal populations. *Communications in Statistics: Simulation and Computation*, 49(1), 194–206. doi:10.1080/03610918.2018.1484138.

- [12] Thangjai, W., Niwitpong, S.-A., & Niwitpong, S. (2020). Bayesian Confidence Intervals for Means of Normal Distributions with Unknown Coefficients of Variation. *Integrated Uncertainty in Knowledge Modelling and Decision Making*, 361–371. doi:10.1007/978-3-030-62509-2_30.
- [13] Thangjai, W., Niwitpong, S. A., & Niwitpong, S. (2021). A Bayesian Approach for Estimation of Coefficients of Variation of Normal Distributions. *Sains Malaysiana*, 50(1), 261–278. doi:10.17576/jsm-2021-5001-25.
- [14] Casella G, & Berger R.L. (2002). *Statistical inference* (2nd Ed.). Duxbury Press, Pacific Grove, United States.
- [15] Maneerat, P., Niwitpong, S. A., & Niwitpong, S. (2020). Bayesian confidence intervals for the difference between variances of delta-lognormal distributions. *Biometrical Journal*, 62(7), 1769–1790. doi:10.1002/bimj.201900079.
- [16] Maneerat, P., & Niwitpong, S. A. (2020). Comparing medical care costs using bayesian credible intervals for the ratio of means of delta-lognormal distributions. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 28(Supp01), 51–68. doi:10.1142/S021848852040005X.
- [17] Thangjai, W., Niwitpong, S. A., & Niwitpong, S. (2021). Bayesian confidence intervals for coefficients of variation of pm10 dispersion. *Emerging Science Journal*, 5(2), 139–154. doi:10.28991/esj-2021-01264.
- [18] Tian, L., & Wu, J. (2007). Inferences on the common mean of several log-normal populations: The generalized variable approach. *Biometrical Journal*, 49(6), 944–951. doi:10.1002/bimj.200710391.
- [19] Ye, R. D., Ma, T. F., & Wang, S. G. (2010). Inferences on the common mean of several inverse Gaussian populations. *Computational Statistics and Data Analysis*, 54(4), 906–915. doi:10.1016/j.csda.2009.09.039.
- [20] Niwitpong, S., & Wongkhao, A. (2016). Confidence Intervals for the Difference between Inverse of Normal Means. *Advances and Applications in Statistics*, 48(5), 337–347. doi:10.17654/as048050337
- [21] Thangjai, W., Niwitpong, S., & Niwitpong, S. A. (2017). Confidence intervals for mean and difference between means of normal distributions with unknown coefficients of variation. *Mathematics*, 5(3), 39. doi:10.3390/math5030039.
- [22] Thangjai, W., Niwitpong, S.-A., & Niwitpong, S. (2018). Confidence Intervals for Variance and Difference between Variances of One-Parameter Exponential Distributions. *Advances and Applications in Statistics*, 53(3), 259–283. doi:10.17654/as053030259.
- [23] Thangjai, W., & Niwitpong, S.-A. (2019). Confidence Intervals for the Signal-to-Noise Ratio and Difference of Signal-to-Noise Ratios of Log-Normal Distributions. *Stats*, 2(1), 164–173. doi:10.3390/stats2010012.
- [24] Donner, A., & Zou, G. Y. (2012). Closed-form confidence intervals for functions of the normal mean and standard deviation. *Statistical Methods in Medical Research*, 21(4), 347–359. doi:10.1177/0962280210383082.
- [25] Niwitpong, S. A. (2015). Confidence intervals for the difference between coefficients of variation of normal distribution with bounded parameters. *Far East Journal of Mathematical Sciences*, 98(5), 649–663. doi:10.17654/FJMSNov2015_649_663.
- [26] Wongkhao, A., Niwitpong, S. A., & Niwitpong, S. (2015). Confidence intervals for the ratio of two independent coefficients of variation of normal distribution. *Far East Journal of Mathematical Sciences*, 98(6), 741–757. doi:10.17654/FJMSNov2015_741_757.
- [27] Thangjai, W., Niwitpong, S., & Niwitpong, S. (2019). Simultaneous confidence intervals for all differences of coefficients of variation of log-normal distributions. *Hacettepe Journal of Mathematics and Statistics*, 48(5), 1505–1521. doi:10.15672/hujms.454804.
- [28] Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, & Rubin DB. (2013). *Bayesian data analysis* (3rd Ed.). Taylor & Francis Group, New York, United States. doi:10.1201/b16018.
- [29] Gokhale, D. V., Box, G. E. P., & Tiao, G. C. (1974). *Bayesian Inference in Statistical Analysis*. International Biometric Society, 30(1), 211. doi:10.2307/2529631
- [30] Thangjai, W., Niwitpong, S.-A., & Niwitpong, S. (2022). Analysis of Medical Data Using Interval Estimators for Common Mean of Gaussian Distributions with Unknown Coefficients of Variation. *Integrated Uncertainty in Knowledge Modelling and Decision Making*, 155–166. doi:10.1007/978-3-030-98018-4_13.
- [31] Thangjai, W., Niwitpong, S. A., & Niwitpong, S. (2018). Simultaneous confidence intervals for all differences of means of normal distributions with unknown coefficients of variation. *Studies in Computational Intelligence*, 753, 670–682. doi:10.1007/978-3-319-70942-0_48.
- [32] Ali, A. A., Al-Attar, T. S., & Abbas, W. A. (2022). A Statistical Model to Predict the Strength Development of Geopolymer Concrete Based on SiO₂/Al₂O₃ Ratio Variation. *Civil Engineering Journal*, 8(3), 454–471. doi:10.28991/cej-2022-08-03-04.
- [33] Martin, S. A., & Ritchie, R. J. (2020). Sourcing Thai geography literature for ASEAN and international education. *Singapore Journal of Tropical Geography*, 41(1), 61–85. doi:10.1111/sjtg.12296.