STATISTICS IN TRANSITION new series, September 2023
Vol. 24, No. 4, pp. 109–122, https://doi.org/10.59170/stattrans-2023-054
Received – 12.04.2022; accepted – 10.02.2023

Testing the annual rainfall dispersion in Chaiyaphum, Thailand, by using confidence intervals for the coefficient of variation of an inverse gamma distribution

Wararit Panichkitkosolkul[1]

Abstract

In Thailand, droughts are regular natural disasters that happen nearly every year due to several factors such as precipitation deficiency, human activity, and the global warming. Since annual rainfall amount fits an inverse gamma (IG) distribution, we wanted to try testing annual rainfall dispersion via the coefficient of variation (CV). Herein, we propose two statistics for testing the CV of an IG distribution based on the Score and Wald methods. We evaluated their performances by means of the Monte Carlo simulations conducted under several shape parameter values for an IG distribution based on empirical type I error rates and powers of the tests. The simulation results reveal that the Wald-method test statistic performed better than the Score-method one in terms of the attained nominal significance level, and is thus recommended for analysis in similar scenarios. Furthermore, the efficacy of the proposed test statistics was illustrated by applying them to the annual rainfall amounts in Chaiyaphum, Thailand.

Key words: statistical test, measure of dispersion, continuous distribution, simulation, meteorology.

1. Introduction

Since damage from natural disasters has increased due to anomalous global climate changes, researchers have become interested in studying their occurrences. Thailand has been divided into six geographical regions by the National Research Council: north, northeast, central, east, west, and south; many of them are prone to droughts but they most often occur in the central northeastern part of Thailand. Thailand is one of the most drought-affected countries in the Asia-Pacific region and is marred by frequent droughts (Khadka et al., 2021). Drought in Thailand directly affects agriculture and water resources, which has a significant impact on the country's economy since most of the country is agrarian.

The north-eastern of Thailand is one of the highly drought-prone regions of the country (Prabnakorn et al., 2018). Chaiyaphum, one of the north-eastern provinces of Thailand, is faced with drought every year due to long periods of little rain causing a severe shortage of water for both consumption and farming (Srichaiwong et al., 2020). Figure 1 shows the map of Chaiyaphum from the Google Maps (2023). In July 2019, parts of Chaiyaphum were faced with a severe drought, and the water volume in the Chulabhorn Dam decreased

[1]Department of Mathematics and Statistics, Faculty of Science and Technology, Thammasat University, Pathumthani 12120, Thailand. E-mail: wararit@mathstat.sci.tu.ac.th. ORCID: https://orcid.org/0000-0001-8315-8185.

to its lowest level in 30 years (only 25% of its capacity) (Pattayamail, 2022). Moreover, in January 2020, eight hospitals in Chaiyaphum were impacted by the drought, leading to the Chaiyaphum Provincial Public Health Office drilling artesian wells to reserve water for medical services and sufficient staff consumption for at least three days while also requesting citizens to help by saving water (Nationthailand, 2022).

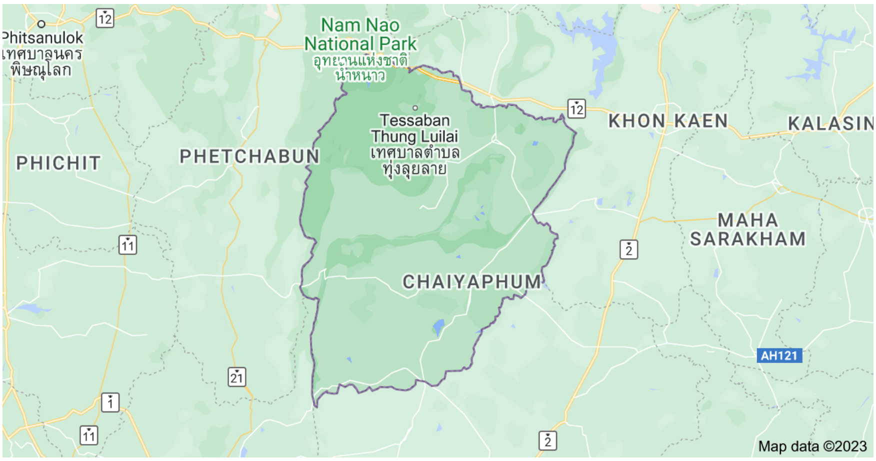

Figure 1: The map of Chaiyaphum, Thailand

Droughts take place whenever there are prolonged periods of rainfall deficiency for one season or more (Eartheclipse, 2022). The major cause of meteorological drought is a deficit of rainfall (Wichitarapongsakun et al., 2016). Since the rainfall amount varies greatly depending on the region and season, the coefficient of variation (CV) can be used to represent rainfall dispersion in different regions. The CV is a unit-free measure of variability relative to the population mean (Albatineh et al., 2017). It is defined as the ratio of the population standard deviation σ to the population mean μ namely $\theta = \sigma/\mu$, where $\mu \neq 0$. It has been more widely used than the standard deviation for comparing the variations of several variables obtained by different units.

The estimator of the CV has been widely applied in many fields of science, including the medical sciences, engineering, economics and others. For example, the applicability of the CV method for analyzing synaptic plasticity was studied by Faber and Korn (1991). Reed et al. (2002) used the CV in assessing the variability of quantitative assays. Kang et al. (2007) applied the CV for monitoring variability in statistical process control. Pang et al. (2008) proposed a simulation-based approach to the study of CV of dividend yields. The improved estimators of CV in a finite population were introduced by Archana and Rao (2011). Calif and Soubdhan (2016) used the CV to measure the spatial and temporal correlation of global solar radiation. Singh and Mishra (2019) proposed an improved estimation method for the population coefficient of variation, which uses information on a single auxiliary variable. Thangjai and Niwitpong (2020) proposed confidence interval estimation for the ratio of CV of two log-normal distributions constructed using the Bayesian approach.

The inverse gamma (IG) distribution is a two-parameter family of continuous distributions on the positive real line based on the reciprocal of a variable (Abid and Al-Hassany, 2016). Milevsky and Posner (1998) studied the IG distribution and pointed out estimation by method of moments. It is often used as a conjugate prior distribution in Bayesian statistics (Zhang and Zhang, 2022). There have been several research papers published on applying the IG distribution. For example, Gelman (2006) applied the IG distribution as a prior distribution for variance parameters in hierarchical models. Rasheed and Sultan (2015) proposed the Bayesian estimator for the scale parameter of IG distribution using Linex loss function and squared error loss function with non-informative prior. Abid and Al-Hassany (2016) studied some issues related to the inverted gamma distribution, which is the reciprocal of the gamma distribution. Llera and Beckmann (2016) introduced five different algorithms based on the method of moments, maximum likelihood, and Bayesian methodology to estimate the parameters of an IG distribution. Glen and Leemis (2017) applied the IG distribution to survival studies. Ramírez-Espinosa and Lopez-Martinez (2019) proposed the utility of the IG distribution in modeling composite fading channels. Yoo et al. (2019) provided empirical evidence that the IG distribution is an excellent alternative for the lognormal and gamma distributions which are often used to model shadowing. Furthermore, the confidence intervals for the ratio of the CVs of the IG distributions were introduced by Kaewprasert et al. (2023).

The literature on testing the CV for the IG distribution is limited. However, there are many methods available for estimating the confidence interval for a population CV for the IG distribution. Kaewprasert et al. (2020) presented three confidence intervals for the CV of an IG distribution using the Score method, the Wald method and the percentile bootstrap confidence interval. These confidence intervals for the CV can be used to test the hypothesis for the CV.

The objective of this paper is to propose some methods for testing the CV for the IG distribution and identify the appropriate methods for practitioners. Two confidence intervals proposed by Kaewprasert et al. (2020) are considered in order to test the CV. A simulation study was conducted to compare the performance of these methods. Based on the simulation results, test statistic with high power that attained a nominal significance level is recommended for practitioners.

The rest of this paper is organized as follows. The point estimation of parameters in an IG distribution is reviewed in Section 2. In Section 3, we present the proposed methods for testing the CV of the IG distribution. The simulation study and results are discussed in Section 4. Section 5 shows the application of the proposed statistical tests to real data is shown using the annual rainfall amounts in Chaiyaphum, Thailand. Discussion and conclusions are presented in the final section.

2. Point estimation of parameters in an inverse gamma distribution

In this section, we explain the point estimation of parameters in an IG distribution. Let $X = (x_1, \ldots, x_n)$ be a random sample from the IG distribution with the shape parameter α and scale parameter β, denoted as $IG(\alpha, \beta)$. The probability density function of X (Rivera et al., 2021) is given by

$$f(x;\alpha,\beta)=\frac{\beta^{\alpha}}{\Gamma(\alpha)}x^{-\alpha-1}\exp\left(-\frac{\beta}{x}\right),\ x>0,\ \alpha>0,\ \beta>0. \tag{1}$$

The population mean and variance of X are defined as $E(X)=\beta/(\alpha-1)$, for $\alpha>1$ and $Var(X)=\beta^2/\left[(\alpha-1)^2(\alpha-2)\right]$, for $\alpha>2$. Therefore, the CV of X can be expressed as $CV(X)=\theta=1/\sqrt{\alpha-2}$.

Since α is an unknown parameter, it is required to be estimated. We consider the maximum likelihood estimators (MLEs) for α and β. Thus, the log-likelihood function of α and β is given by

$$\ln L(\alpha,\beta)=-\sum_{i=1}^{n}\left(\frac{\beta}{X_i}\right)-(\alpha+1)\sum_{i=1}^{n}\ln(X_i)-n\ln\Gamma(\alpha)+n\alpha\ln(\beta).$$

Taking partial derivatives of the above equation with respect to α and β, respectively, the Score function is derived as

$$U(\alpha,\beta)=\begin{bmatrix}\sum_{i=1}^{n}\ln(X_i)-nln(\alpha)+\frac{n}{2\alpha}-n\ln(\beta)\\ -\sum_{i=1}^{n}X_i^{-1}+\frac{n\alpha}{\beta}\end{bmatrix}.$$

Then, the MLEs can be conducted for α and β, respectively,

$$\hat{\alpha}=\frac{1}{2\left[\frac{\sum_{i=1}^{n}\ln(X_i)}{n}+\ln\left(\frac{\sum_{i=1}^{n}X_i^{-1}}{n}\right)\right]},\ \text{and}\ \hat{\beta}=\frac{n\hat{\alpha}}{\sum_{i=1}^{n}X_i^{-1}}.$$

Also, the estimator of CV is given by $\hat{\theta}=1/\sqrt{\hat{\alpha}-2}$.

3. Methods for testing the coefficient of variation of the inverse gamma distribution

Let $X_1,\ldots,X_n$ be an independent and identically distributed random sample of size n from the IG distribution with the shape parameter α and scale parameter β. We want to test for the population CV. The null and alternative hypotheses are defined as follows:

$$H_0:\theta=\theta_0\ \text{ versus }\ H_1:\theta\neq\theta_0.$$

In this section, we discuss two test statistics for the CV based on the Score method and the Wald method.

3.1. Score method

Let α and β be the parameter of interest and the nuisance parameters, respectively. In general, the Score statistic (Rao, 1948, 2005) is denoted as

$$W_1 = U^T(\alpha_0, \hat{\beta}_0) I^{-1}(\alpha_0, \hat{\beta}_0) U(\alpha_0, \hat{\beta}_0),$$

where $\hat{\beta}_0$ is the MLE for β, under the null hypothesis $H'_0 : \alpha = \alpha_0$, $U(\alpha_0, \hat{\beta}_0)$ is the vector of the Score function and $I(\alpha_0, \hat{\beta}_0)$ is the matrix of the Fisher information; see e.g., Kay (1993). Here, it is easy to derive that the Score function under H'_0 is

$$U(\alpha_0, \hat{\beta}_0) = \begin{bmatrix} -\sum_{i=1}^{n} \ln(X_i) + \frac{n}{2\alpha_0} - n\ln\left(n/\sum_{i=1}^{n} X_i^{-1}\right) \\ 0 \end{bmatrix}.$$

The inverse of the matrix of the Fisher information can be derived as follows:

$$I^{-1}(\alpha_0, \hat{\beta}_0) = \begin{bmatrix} \frac{2\alpha_0^2}{n} & -\frac{2\alpha_0^2}{\sum_{i=1}^{n} X_i^{-1}} \\ -\frac{2\alpha_0^2}{\sum_{i=1}^{n} X_i^{-1}} & \frac{n\alpha_0(2\alpha_0-1)}{\left(\sum_{i=1}^{n} X_i^{-1}\right)^2} \end{bmatrix}.$$

Using the property of the Score function, we can see that the pivotal

$$Z_{score} = \sqrt{\frac{2\alpha_0^2}{n}} \left[-\sum_{i=1}^{n} \ln(X_i) + \frac{n}{2\alpha_0} + n\ln\left(n/\sum_{i=1}^{n} X_i^{-1}\right) \right] \tag{2}$$

converges in distribution to the standard normal distribution. Since the variance of $\hat{\alpha}$ is $\frac{2\alpha_0^2}{n}$, it is approximated by substituting $\hat{\alpha}$ in its variance. Under H'_0, the statistic in (2) is given as

$$Z_{score} \cong \sqrt{\frac{2\hat{\alpha}^2}{n}} \left[-\sum_{i=1}^{n} \ln(X_i) + \frac{n}{2\hat{\alpha}} + n\ln\left(n/\sum_{i=1}^{n} X_i^{-1}\right) \right].$$

From the probability statement, $1-\gamma = P(-Z_{1-\gamma/2} \le Z_{score} \le Z_{1-\gamma/2})$, it can be simply written as $1-\gamma = P(l_s \le \theta \le u_s)$. Therefore, the $(1-\gamma)100\%$ confidence interval for θ based on the Score method is given by

$$CI_S = [l_s, u_s] = \left[\frac{1}{\sqrt{\frac{n}{2\left(z_1 - Z_{\gamma/2}\sqrt{\frac{n}{2\hat{\alpha}^2}}\right)} - 2}}, \frac{1}{\sqrt{\frac{n}{2\left(z_1 + Z_{\gamma/2}\sqrt{\frac{n}{2\hat{\alpha}^2}}\right)} - 2}} \right],$$

where $z_1 = \sum_{i=1}^{n} \ln(X_i) - n\ln\left(n/\sum_{i=1}^{n} X_i^{-1}\right)$ and $Z_{\gamma/2}$ is the $\gamma/2$-upper quantile of the standard normal distribution. Therefore, we will reject the null hypothesis, $H_0 : \theta = \theta_0$, if

$$\theta_0 < \frac{1}{\sqrt{\frac{n}{2\left(z_1 - Z_{\gamma/2}\sqrt{\frac{n}{2\hat{\alpha}^2}}\right)} - 2}} \text{ or } \theta_0 > \frac{1}{\sqrt{\frac{n}{2\left(z_1 + Z_{\gamma/2}\sqrt{\frac{n}{2\hat{\alpha}^2}}\right)} - 2}}.$$

3.2. Wald method

The Wald statistic is an asymptotic statistic derived from the property of the MLE (Gaffke et al., 2002). The general form of the Wald statistic under the null hypothesis $H_0' : \alpha = \alpha_0$ is defined as

$$W_2 = (\hat{\alpha} - \alpha_0)^T \left[I^{\alpha\alpha}(\hat{\alpha}, \hat{\beta}) \right]^{-1} (\hat{\alpha} - \alpha_0),$$

where $I^{\alpha\alpha}(\hat{\alpha}, \hat{\beta})$ is the estimated variance of $\hat{\alpha}$ obtained from the first row and the first column of $I^{-1}(\hat{\alpha}, \hat{\beta})$. Using the information of partial derivatives from the previous subsection, the inverse matrix is given by

$$I^{-1}(\hat{\alpha}, \hat{\beta}) = \begin{bmatrix} \frac{2\hat{\alpha}^2}{n} & -\frac{2\hat{\alpha}^2}{\sum_{i=1}^{n} X_i^{-1}} \\ -\frac{2\hat{\alpha}^2}{\sum_{i=1}^{n} X_i^{-1}} & \frac{n\hat{\alpha}(2\hat{\alpha}-1)}{\left(\sum_{i=1}^{n} X_i^{-1}\right)^2} \end{bmatrix},$$

where $I^{\alpha\alpha}(\hat{\alpha}, \hat{\beta}) = \frac{2\hat{\alpha}^2}{n}$. Therefore, under H_0', we obtain the Wald statistic

$$Z_{wald} \cong \sqrt{\frac{n}{2\hat{\alpha}^2}} (\hat{\alpha} - \alpha), \tag{3}$$

which has the limiting distribution of a standard normal distribution. Thus, the $(1-\gamma)100\%$ confidence interval for θ based on the Wald method is given by

$$CI_W = [l_w, u_w] = \left[\frac{1}{\sqrt{\hat{\alpha} - 2 + Z_{\gamma/2}\sqrt{\frac{2\hat{\alpha}^2}{n}}}}, \frac{1}{\sqrt{\hat{\alpha} - 2 - Z_{\gamma/2}\sqrt{\frac{2\hat{\alpha}^2}{n}}}} \right],$$

where $Z_{\gamma/2}$ is the $\gamma/2$-upper quantile of the standard normal distribution. Therefore, we will reject the null hypothesis, $H_0 : \theta = \theta_0$, if

$$\theta_0 < \frac{1}{\sqrt{\hat{\alpha} - 2 + Z_{\gamma/2}\sqrt{\frac{2\hat{\alpha}^2}{n}}}} \quad \text{or} \quad \theta_0 > \frac{1}{\sqrt{\hat{\alpha} - 2 - Z_{\gamma/2}\sqrt{\frac{2\hat{\alpha}^2}{n}}}}.$$

4. Simulation Study and Results

In this study, two statistical methods for testing the population CV in an IG distribution are considered. Since a theoretical comparison is not possible, a Monte Carlo simulation was conducted using the R version 4.1.3 statistical software (Ihaka and Gentleman, 1996) to compare the performance of the test statistics. The methods were compared in terms of their attainment of empirical type I error rates and the powers of their performance. We count the number of times for each test that the null hypothesis was rejected when H_0 was true, to obtain the empirical type I error rates. In addition, the number of times for each test, that the null hypothesis was rejected when H_0 was not true, was counted to obtain the power of the test. The simulation results are presented for the significance level γ= 0.05, since a) γ= 0.05 is widely used to compare the power of the test and b) similar conclusions were obtained for other values of γ.

To observe the behaviour of small, moderate and large sample sizes, we used n = 25, 50, 75, 100 and 200. Each Monte Carlo experiment consisted of 10,000 replications. The data were generated from an IG distribution with β =1 and α was adjusted to obtain the required coefficient of variation θ. We set θ = 0.10, 0.15, 0.20 and 0.30.

As can be seen in the simulation results shown in Tables 1-4, the empirical type I error rates of the Wald method were close to the nominal significance level of 0.05 for all sample sizes while those of the Score method were close to the nominal significance level of 0.05 for larger sample sizes. Note that the Score method had a high empirical type I error rate when sample sizes were small. The Score method performed well in terms of the power of the test for $\theta < \theta_0$. On the other hand, the Wald method performed better for $\theta > \theta_0$. We observed a general pattern; when the sample size increases, the power of the test also increases and the empirical type I error rate approaches 0.05. Also, the power increases as the value of the CV departs from the hypothesized value of the CV. It was observed that for large sample sizes, the performance of the test statistics did not differ greatly in the sense of power and the attainment of the nominal significance level of the test. However, a significant difference was observed for small sample sizes.

Table 1. Empirical type I error rates (bold numeric) and powers of tests for IG(102,1), θ =0.10.

n	Method	θ_0								
		0.06	0.07	0.08	0.09	0.10	0.11	0.12	0.13	0.14
25	Score	0.6696	0.2744	0.0610	0.0413	**0.1155**	0.2711	0.4674	0.6806	0.8383
	Wald	0.9417	0.7357	0.3994	0.1433	**0.0431**	0.0483	0.1055	0.2268	0.3772
50	Score	0.9853	0.7963	0.3257	0.0583	**0.0820**	0.2996	0.6087	0.8623	0.9722
	Wald	0.9977	0.9455	0.6460	0.2192	**0.0432**	0.0962	0.2933	0.5913	0.8365
75	Score	0.9996	0.9561	0.5740	0.1078	**0.0680**	0.3319	0.7339	0.9436	0.9948
	Wald	1.0000	0.9883	0.7927	0.2888	**0.0420**	0.1454	0.4792	0.8088	0.9681
100	Score	1.0000	0.9911	0.7512	0.1785	**0.0688**	0.3806	0.8169	0.9816	0.9995
	Wald	1.0000	0.9982	0.8867	0.3592	**0.0469**	0.1976	0.6333	0.9316	0.9967
200	Score	1.0000	1.0000	0.9799	0.4410	**0.0595**	0.5595	0.9690	0.9999	1.0000
	Wald	1.0000	1.0000	0.9919	0.5863	**0.0496**	0.4089	0.9279	0.9992	1.0000

Table 2. Empirical type I error rates (bold numeric) and powers of tests for IG(46.44, 1), θ =0.15.

n	Method	θ_0								
		0.11	0.12	0.13	0.14	0.15	0.16	0.17	0.18	0.19
25	Score	0.1695	0.0592	0.0323	0.0542	**0.1131**	0.1966	0.3222	0.4602	0.5987
	Wald	0.6244	0.3933	0.1982	0.0933	**0.0398**	0.0350	0.0559	0.1014	0.1703
50	Score	0.6296	0.3168	0.0985	0.0401	**0.0794**	0.1973	0.3845	0.5925	0.7685
	Wald	0.8673	0.6352	0.3253	0.1241	**0.0476**	0.0580	0.1421	0.2815	0.4585
75	Score	0.8726	0.5564	0.2101	0.0549	**0.0702**	0.2085	0.4603	0.7160	0.8807
	Wald	0.9595	0.7910	0.4546	0.1564	**0.0446**	0.0814	0.2292	0.4591	0.6954
100	Score	0.9612	0.7378	0.3319	0.0750	**0.0651**	0.2276	0.5131	0.7890	0.9465
	Wald	0.9884	0.8882	0.5468	0.1856	**0.0465**	0.1005	0.3030	0.5954	0.8423
200	Score	0.9996	0.9757	0.6977	0.1775	**0.0556**	0.3102	0.7294	0.9599	0.9979
	Wald	0.9998	0.9909	0.8137	0.2988	**0.0501**	0.1901	0.5836	0.9112	0.9923

Table 3. Empirical type I error rates (bold numeric) and powers of tests for IG(27, 1), θ =0.20.

n	Method	θ_0								
		0.16	0.17	0.18	0.19	0.20	0.21	0.22	0.23	0.24
25	Score	0.0531	0.0332	0.0383	0.0683	**0.1146**	0.1693	0.2406	0.3386	0.4419
	Wald	0.3825	0.2363	0.1356	0.0727	**0.0421**	0.0319	0.0382	0.0617	0.0920
50	Score	0.2990	0.1337	0.0525	0.0418	**0.0782**	0.1537	0.2766	0.4017	0.5590
	Wald	0.6224	0.4029	0.2099	0.0950	**0.0446**	0.0501	0.0924	0.1523	0.2578
75	Score	0.5438	0.2697	0.0988	0.0428	**0.0677**	0.1550	0.3047	0.4879	0.6761
	Wald	0.7710	0.5278	0.2700	0.1097	**0.0497**	0.0571	0.1323	0.2445	0.4195
100	Score	0.7210	0.4130	0.1611	0.0571	**0.0616**	0.1630	0.3386	0.5600	0.7552
	Wald	0.8735	0.6327	0.3462	0.1282	**0.0485**	0.0642	0.1731	0.3494	0.5535
200	Score	0.9734	0.7889	0.4010	0.1028	**0.0545**	0.2039	0.4865	0.7950	0.9462
	Wald	0.9889	0.8803	0.5573	0.1895	**0.0495**	0.1178	0.3402	0.6589	0.8870

Table 4. Empirical type I error rates (bold numeric) and powers of tests for IG(13.11, 1), θ =0.30.

n	Method	θ_0								
		0.26	0.27	0.28	0.29	0.30	0.31	0.32	0.33	0.34
25	Score	0.0324	0.0431	0.0594	0.0784	**0.1094**	0.1444	0.1834	0.2214	0.2787
	Wald	0.1879	0.1306	0.0915	0.0557	**0.0429**	0.0328	0.0308	0.0345	0.0457
50	Score	0.0858	0.0535	0.0427	0.0534	**0.0736**	0.1136	0.1653	0.2352	0.3164
	Wald	0.2997	0.2058	0.1178	0.0717	**0.0481**	0.0410	0.0513	0.0759	0.1034
75	Score	0.1832	0.0934	0.0534	0.0426	**0.0623**	0.1103	0.1775	0.2544	0.3573
	Wald	0.4103	0.2550	0.1497	0.0818	**0.0511**	0.0474	0.0676	0.1048	0.1649
100	Score	0.2806	0.1494	0.0678	0.0442	**0.0603**	0.1024	0.1809	0.2853	0.4150
	Wald	0.5001	0.3153	0.1690	0.0878	**0.0497**	0.0474	0.0816	0.1320	0.2227
200	Score	0.6430	0.3718	0.1668	0.0652	**0.0517**	0.1115	0.2307	0.4030	0.5970
	Wald	0.7682	0.5246	0.2832	0.1242	**0.0513**	0.0618	0.1361	0.2686	0.4456

5. An Empirical Application

To illustrate the applicability of the two statistical methods for testing the CV introduced in the previous section, we used annual rainfall data in millimetres obtained from the Hydrology Irrigation Center for the Upper Northeastern Region, the Royal Irrigation Department, Thailand (http://hydro-3.rid.go.th). The annual rainfall amounts were measured at the Irrigation Station, Mueang District, Chaiyaphum, Thailand from 1998 to 2021. The descriptive statistics are as follows: sample size = 23, mean = 1088.44 mm, standard deviation (SD) = 245.79 mm, CV = 0.226, coefficient of skewness = 0.886, and kurtosis = 0.946. The distribution of the annual rainfall amount is right-skewed and it has heavy-tailed data distribution. The histogram, density plot, Box and Whisker plot, and inverse gamma quantile-quantile (Q-Q) plot shown in Figure 1 confirm that the fitted distribution for the annual rainfall amounts is not symmetric.

Table 5 reports the Akaike information criterion (AIC) (Akaike, 1974) results to check the fitting of the distribution for the annual rainfall amounts in Chaiyaphum. The AIC is defined as $AIC = -2\ln L + 2k$, where L is the likelihood function and k is the number of parameters. The results show that the annual rainfall amounts in Chaiyaphum follow an IG distribution because the AIC value for this distribution was the smallest. The annual rainfall amounts in Chaiyaphum had an IG distribution with shape parameter $\hat{\alpha}$= 26.8951 and scale parameter $\hat{\beta}$= 23588.810, while the MLE for the CV is $\hat{\theta}$= 0.2148.

Our interest was in testing the population CV of the annual rainfall amounts in Chaiyaphum. Suppose the researcher wanted to test the claim that a population CV equals 0.25. The null and alternative hypotheses are respectively given as follows:

$$H_0 : \theta = 0.25 \text{ versus } H_1 : \theta \neq 0.25.$$

The lower and upper critical values of both test statistics were shown in Table 6. The null hypothesis H_0 was not rejected since $0.1390 \leq \theta_0 \leq 0.2843$ and $0.1721 \leq \theta_0 \leq 0.3634$ using test statistics based on the Score and Wald methods, respectively. We conclude that the population CV of the annual rainfall amounts in Chaiyaphum does not differ from 0.25 at the 0.05 significance level.

Table 5. Results of AIC for the annual rainfall amounts in Chaiyaphum, Thailand.

Normal	Cauchy	Exponential	Weibull	Gamma	Inverse Gamma
321.4549	328.0113	369.6550	323.7112	319.2704	**318.2490**

Table 6. Critical values of test statistics based on the Score and Wald methods for the significance level of 0.05

Method	Critical values	
	Lower	Upper
Score	0.1390	0.2843
Wald	0.1721	0.3634

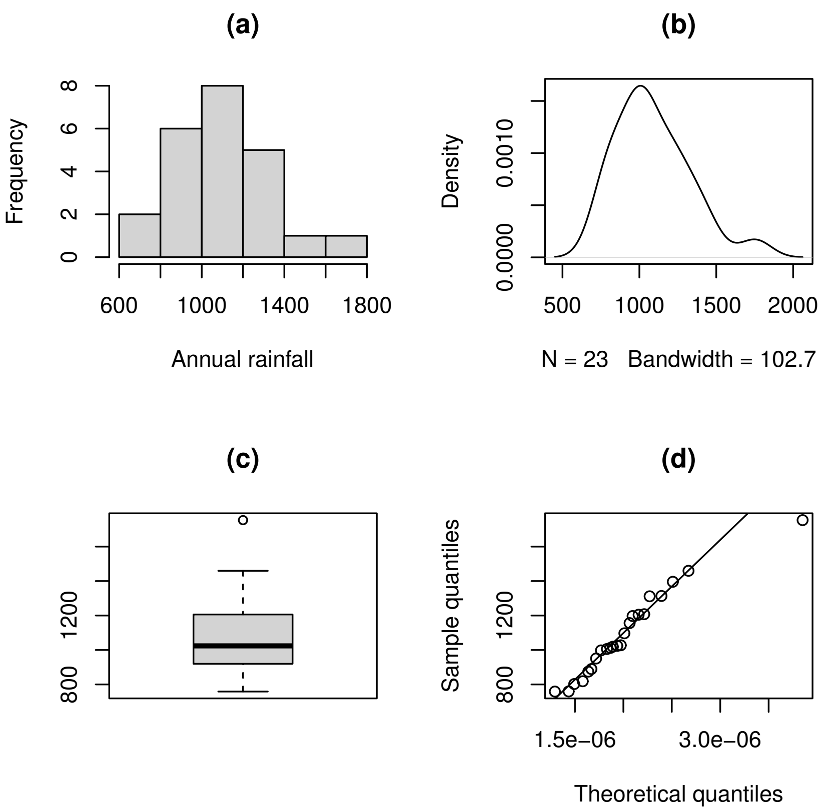

Figure 2: (a) histogram (b) density plot (c) Box and Whisker plot (d) inverse gamma Q-Q plot of the annual rainfall amounts in Chaiyaphum, Thailand

6. Conclusions and Discussion

The aim of this study is to identify potential methods that can be recommended to practitioners for testing the population CV in an IG distribution. A general pattern was observed (as expected); as the sample size increased, the power of the test also increased and the empirical type I error rates approached 0.05. Moreover, the power increased as the value of CV departed from the hypothesized value of the CV. It can be observed that for large sample sizes, the performance of both methods did not differ greatly in terms of the power and attaining the nominal size of the test. However, a significant difference was observed for small sample sizes.

In this study, two statistical methods for testing the population CV in an IG distribution were derived. Based on the simulation results, it is evident that the Wald method performed better than the Score method in terms of the empirical type I error rate. The Score method performed well in the sense of the power of the test when the population CV was smaller than the hypothesized value of the CV. On the other hand, the Wald method performed better when the population CV was greater than the hypothesized value of the CV. In summary, we would recommend the Wald method for testing since its empirical type I error rate is close to the nominal significance level. Furthermore, Kaewprasert et al. (2020) concluded that the best method for estimating confidence interval for the CV of the IG distribution was the Wald method. The conclusions of this study were consistent with the study of Kaewprasert et al. (2020). In addition, the researchers can apply the proposed methods for testing the population CV in an IG distribution with other data sets fitted well to an IG distribution. For example, the IG distribution has been used for the hitting time distribution of a Wiener process. Future research could focus on the one-tailed hypothesis testing.

Acknowledgements

The author would like to thank editor and the reviewers for the valuable comments and suggestions to improve this paper.

References

Abid, S. H., Al-Hassany, S. A., (2016). On the inverted gamma distributions. *International Journal of Systems Science and Applied Mathematics*, Vol. 1, pp. 16–22.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, Vol. 19, pp. 716–723.

Albatineh, A. N., Boubakari, I. and Kibra, B. M. G., (2017). New confidence interval estimator of the signal-to-noise ratio based on asymptotic sampling distribution. *Communication in Statistics-Theory and Methods*, Vol. 46, pp. 574–590.

Archana, V., Rao, K. A., (2011). Improved estimators of coefficient of variation in a finite population, *Statistics in Transition new series*, Vol. 12, pp. 357–380.

Calif, R., Soubdhan, T., (2016). On the use of the coefficient of variation to measure spatial and temporal correlation of global solar radiation. *Renewable Energy*, Vol. 88, pp. 192–199.

Eartheclipse, (2022). *What is a Drought?*. [online]. Retrieved from https://eartheclipse.com/natural-disaster/causes-and-effects-of-drought.html, 2 April 2022.

Faber, D. S., Korn, H., (1991). Applicability of the coefficient of variation method for analyzing synaptic plasticity. *Biophysical Journal*, Vol. 60, pp. 1288–1294.

Gaffke, N., Heiligers, B. and Offinger, H., (2002). On the asymptotic null-distribution of the Wald statistic at singular parameter points. *Statistics & Decisions*, Vol. 20, pp. 379–398.

Gelman, A., (2006). Prior distributions for variance parameters in hierarchical models, *Bayesian Analysis*, Vol. 1, pp. 515–533.

Glen, A. G., Leemis, L. M., eds. (2017). Computational Probability Applications. *International Series in Operations Research & Management Science*, Vol. 247, Cham: Springer.

Google Maps, (2023). Chaiyaphum, Available from: https://www.google.co.th/maps/@16.0082727, 101.33255,9z?hl=en [Accessed 8 February 2023].

Ihaka, R., Gentleman, R., (1996). R: a language for data analysis and graphics. *Journal of Computational and Graphical Statistics*, Vol. 5, 299–314.

Kaewprasert, T., Niwitpong, S. A. and Niwitpong, S., (2020). Confidence interval for coefficient of variation of inverse gamma distributions, In: HUYNH, V. N., ENTANI, T., Jeenanunta, C., Inuiguchi, M., Yenradee, P. (Eds.). *Lecture Notes in Artificial Intelligence: Vol.12482. Integrated Uncertainty in Knowledge Modelling and Decision Making*, pp. 407–418, Cham: Springer.

Kaewprasert, T., Niwitpong, S. A. and Niwitpong, S., (2023). Confidence intervals for the ratio of the coefficients of variation of inverse-gamma distributions. *Applied Science and Engineering Progress*, Vol. 16, doi.org/10.14416/ j.asep.2021.12.002.

Kang, C. W., Lee, M. S., Seong, Y. J. and Hawkins, D. M., (2007). A control chart for the coefficient of variation. *Journal of Quality Technology*, Vol. 39, pp. 151–158.

Kay, S. M., (1993). Fundamentals of statistical signal processing: Estimation theory. Englewood Cliffs: Prentice-Hall.

Khadka, D., Babel, M. S., Shrestha, S., Virdis and S. G. P., Collins, M., (2021). Multivariate and multi-temporal analysis of meteorological drought in the northeast of Thailand. *Weather and Climate Extremes*, Vol. 34, https://doi. org/10.1016/j.wace.2021.100399.

Llera, S., Beckmann, C. F., (2016). Estimating an inverse gamma distribution. *Technical report*, Radbound University Nijmegen, Donders Institute for Brain Cognition and Behavior. ar.Xiv: 1605.01019v2.

Milevsky, M. A., Posner, S. E., (1998). Asian options, the sum of log-normals and the reciprocal gamma distribution. *The Journal of Financial and Quantitative Analysis*, Vol. 33, pp. 203–218.

Nationthailand, (2022). *Nine hospitals in provinces suffer impact from drought*, [online]. Retrieved from https://www.nationthailand.com/in-focus/30380811, 5 April 2022.

Pang, W. K., Yu, B. W. T., Troutt, M. D. and Hou, S. H., (2008). A simulation-based approach to the study of coefficient of variation of dividend yields. *European Journal of Operational Research*, Vol. 189, pp. 559–569.

Pattayamail, (2022). *Drought situation in Chaiyaphum reaches crisis point*, [online]. Retrieved from https://www.pattayamail.com/thailandnews/drought-situation-in-chaiyaphum-reaches-crisis-point-260702, 5 April 2022.

Prabnakorn, S., Maskey, S., Suryadi, F. X. and de Fraiture, C., (2018). Rice yield in response to climate trends and drought index in the Mun River Basin, Thailand. *Science of the Total Environment*, Vol. 621, pp. 108–119.

Rao, C. R., (1948). Large sample tests of statistical hypotheses concerning several parameters with applications to problems of estimation. *Mathematical Proceedings of the Cambridge Philosophical Society*, Vol. 44, pp. 50–57.

Rao, C. R., (2005). Score Test: Historical Review and Recent Developments. In: Advances in Ranking and Selection, Multiple Comparisons, and Reliability. Boston: Birkhäuser, pp. 3–20.

RamÍRez-Espinosa, P. Lopez-Martinez, F. J., (2019). On the utility of the inverse gamma distribution in modeling composite fading channels. In: 2019 IEEE Global Communications Conference (GLOBECOM). Waikoloa, Hawaii, USA, 9–13 December 2019, New York: Curran Associates.

Rasheed, H. A., Sultan, A. J., (2015). Bayesian estimation of the scale parameter for inverse gamma distribution under Linex loss function. *International Journal of Advanced Research*, Vol. 3, pp. 369–375.

Reed, G. F., Lynn, F. and Meade, B. D. (2002). Use of coefficient of variation in assessing variability of quantitative assays. *Clinical and Diagnostic Laboratory Immunology*, Vol. 9, pp. 1235–1239.

Rivera, P. A., CalderÍN-Ojeda, E., Gallardo D. I. and Gómez H. W., (2021). A compound class of the inverse gamma and power series distributions. *Symmetry*, Vol. 13, https://doi.org/10.3390/sym13081328.

Singh, R., Mishra, M., (2019). Estimating population coefficient of variation using a single auxiliary variable in simple random sampling. *Statistics in Transition new series*, Vol. 20, pp. 89–111.

Srichaiwong, P., Ardwichai, S., Tungchuvong, L. and Kenpahoom, S., (2020). The live weir innovation at Chi river watershed, Chaiyaphum province, Thailand. *Bioscience Biotechnology Research Communications*, Vol. 13, pp. 103–107.

Thangjai, W., Niwitpong, S., (2020). Comparing particulate matter dispersion in Thailand using the Bayesian confidence intervals for ratio of coefficients of variation, *Statistics in Transition new series*, Vol. 21, pp. 41–60.

Wichitarapongsakun, P., Sarin, C., Klomjek, P. and Chuenchooklin, S. (2016). Rainfall prediction and meteorological drought analysis in the Sakae Krang River basin of Thailand. *Agriculture and Natural Resources*, Vol. 50, pp. 490–498.

Yoo, S. K., Cotton, S. L., Zhang, L. and Sofotasios, P. C., (2019). The inverse gamma distribution: a new shadowing model, In: 2019 8th Asia-Pacific Conference on Antennas and Propagation (APCAP), Incheon, Korea (South), 4–7 August 2019, pp. 475–476.

Zhang, L., Zhang, Y. Y., (2022). The Bayesian posterior and marginal densities of the hierarchical gamma–gamma, gamma–inverse gamma, inverse gamma–gamma, and inverse gamma–inverse gamma models with conjugate priors. *Mathematics*, Vol. 10, https://doi.org/10.3390/math10214005.