ESKİŞEHİR TECHNICAL UNIVERSITY JOURNAL OF SCIENCE AND TECHNOLOGY
A- APPLIED SCIENCES AND ENGINEERING

2019, 20(4), pp. 503 - 514, DOI: 10.18038/estubtda.574511

TESTING THE POPULATION INVERSE-COEFFICIENTS OF VARIATION AND ITS APPLICATION

Nihan POTAS [1*], Hamza GAMGAM [2]

[1] Health Care Management Department, Faculty of Economics and Administrative Science, Ankara Hacı Bayram Veli University, Ankara, Turkey
[2]Department of Statistics, Faculty of Science, Gazi University, Ankara, Turkey

ABSTRACT

In this paper, some tests are introduced and compared for testing the equality of inverse-coefficients of variation. Monte-Carlo simulation method is used for comparisons. In this simulation study, various simulation scenarios were designed with different population numbers ($k = 3$, 6), sample sizes, parameter values and type I error rates ($\alpha = 0.01$, 0.05). The tests were compared in terms of type I error rate and power in these scenarios. When the sample sizes are small, the D and WT tests showed good results in terms of type I error, but the LR and ST tests did not give good results. As the sample sizes increased, the experimental type I error rates of the LR and ST tests converged to nominal type I error and all tests showed good results in general. While the sample sizes were equal, it was found that the LR test was the most powerful test and the ST test sometimes yielded good results. For these sample sizes, the D test yielded the worst results. When the sample sizes are different, the LR and D tests are powerful than the other tests, and the ST test is the worst test in terms of power. As expected, as the sample sizes and nominal type I error rate increased, the powers of the tests also increased. In addition, an application for the tests was made on real data. It was seen that the results of this application and simulation study coincide.

Keywords: Inverse-Coefficient of Variation, Monte-Carlo Simulation, Likelihood Ratio Test, Score Test, Wald Test

1. INTRODUCTION

The mean value per unit standard deviation is called the inverse-coefficient of variation. In other words, the inverse-coefficient of variation indicates how much the mean changes according to the standard deviation, so it is a measure of relative variability. Even if the means and variances of the examined groups are different, they may have the same relative variability. Therefore, the inverse-coefficient of variation is an important parameter in many areas such as engineering, psychiatry, biology, physics, finance and health. Because of this importance, hypothesis testing is needed.

The test statistic of the equality of the inverse-coefficients of variation were first performed by Bowman and Shenton [2]. In normal distributions, the hypothesis of the equality of the inverse-coefficients of variation is proposed by Doornbos and Dijkstra [5] and Bennett [1]. They compared the Likelihood ratio test and Non-Central t test by using the simulation method and made recommendations. However, for the Likelihood ratio test $(k > 2)$, it contains equations that are not solved algebraically. Nairy and Rao [10], using the theorem of Lehmann and Casella [9] proposed the two-step maximum likelihood estimators to solve these equations. Singh [14] conducted a similar study in Bowman and Shenton [2]. Sharma and Krishnan [12], using sample inverse-coefficients of variation, made inferences about population inverse-coefficients of variation [10]. Singh [14] has determined that the test of the equality of the inverse-coefficients of variation in k-normal distribution populations is easy and understandable because it contains less numerical terms than the test of the equality of the coefficients of variation. Chaturvedi and Rani [3] for the inverse-coefficients of variation in normal distribution have developed a sequential procedure to establish a fixed-width confidence interval. Nairy and Rao [10] have proposed three new tests of inverse-coefficients of variation. These suggested tests are equivalent to other tests that test the equality of coefficients of variation. Kalkur and Rao [7] obtained the Bayesian estimator of the inverse-coefficient of variation of the normal distribution using five different objective priors.

*Corresponding Author: nihan_potas@hotmail.com
Received: 10.06.2019 Published: 30.12.2019

In this study, it is emphasized that the test statistics of equality of inverse-coefficients of variation in normal distributions as in other studies are examined. A simulation comparison was made to determine the strengths and weaknesses of these test statistics. It is aimed to reinforce the findings by applying the survival data.

In the second section of the study, the non-central t test, the wald test, the likelihood ratio test and the score test were introduced for the test of the hypothesis of the inverse-coefficient of variation.

In the third section, a simulation study was performed to compare the tests in the study. In the simulation study, there are comparisons of the type I error and power of the tests. These comparisons are made for different sample sizes, type I error and population number and the results are given in the figures and tables. Thus, a wide comparison was made in many aspects of the tests in the study.

In the fourth section, the survival data of 188 patients who underwent bone marrow transplantation in bone marrow transplantation units were used as real data. According to the previous determination of the disease groups, whether there is any difference among the variations of survival times was investigated with inverse-coefficient of variation.

In the last section, the results of the study were evaluated. Suggestions were made based on the results which obtained from the study.

In the next section, some test statistics are used to test the hypothesis of the equality of the inverse-coefficients of variation. These test methods are the non-central t test, the wald test, the likelihood ratio test and the score test. The decision rules for testing the test statistics and hypotheses are given for each test method.

2. SOME TESTS FOR THE EQUALITY OF INVERSE-COEFFICIENTS OF VARIATION

Let, $X_{i1}, X_{i2}, \ldots, X_{in_i}, i = 1,2, \ldots, k,$ be i.i.d normal random variables with μ_i means and σ_i^2 variance. Let show that, $E(X_{ij}) = \mu_i$ and $Var(X_{ij}) = \sigma_i^2,\ i = 1,2, \ldots, k$ and $j = 1,2, \ldots, n_i$. $\gamma_i = \mu_i / \sigma_i$ is the population inverse-coefficient of variation for the i-th population. The inverse-coefficient of variation for the sample is defined as $\hat{\gamma}_i = \overline{X}_i / S_i$ where $\overline{X}_i$ and S_i^2 are the sample mean and variance for the i-th sample. The equality of inverse-coefficients of variation hypothesis,

$$H_0: \gamma_i = \gamma, i = 1,2, \ldots, k, \gamma \text{ is known} \quad (1)$$

against

$$H_1: \gamma_i \neq \gamma_{j^*}, i \neq j^*\ i, j^* = 1,2, \ldots, k \text{ for at least one pair of } (i, j^*)$$

hypothesis, it is desired to be tested. In the following section, some of the test statistics proposed for the test of the hypothesis of the equality of inverse-coefficients of variation are introduced.

2.1. Non-Central t Test

Doornbos and Dijkstra [5] developed the Non-Central t-test by using the distribution of the sample inverse-coefficient of variation. The sample inverse-coefficient of variation is defined as $\hat{\gamma}_i = \overline{X}_i / S_i$, $i = 1,2, \ldots, k$ and $n = \sum_{i=1}^{k} n_i$. $\tilde{\gamma}_n$ is the weighted mean of the inverse-coefficients of variation, is defined as $\tilde{\gamma}_n = \sum_{i=1}^{k} n_i\, \hat{\gamma}_i / n$ and $T = \sum_{i=1}^{k} n_i (\hat{\gamma}_i - \tilde{\gamma}_n)^2$. The expected value of $\hat{\gamma}_i$ is

$$E(\hat{\gamma}_i) = \left(\frac{n_i - 1}{2}\right)^{\frac{1}{2}} \frac{\Gamma\left[\frac{1}{2}(n_i - 2)\right]\gamma}{\Gamma\left[\frac{1}{2}(n_i - 1)\right]} = e_i\gamma, i = 1,2, \dots, k. \tag{2}$$

From Eqs. (2), Doornbos and Dijkstra [5] showed that the variance was

$$Var(\hat{\gamma}_i) = \frac{1}{n_i}\left(1 + \frac{\gamma^2}{2}\right) \tag{3}$$

by using Stirling's formula. The γ parameter in the given hypothesis Eqs. (1), is one of the status of γ^2 and the unbiased estimator for γ^2 is given as

$$\hat{\gamma}^2 = \frac{\sum_{i=1}^{k} n_i \hat{\gamma}_i^2 - \sum_{i=1}^{k} \frac{n_i - 1}{n_i - 3}}{\sum_{i=1}^{k} \frac{n_i(n_i - 1)}{n_i - 3}}. \tag{4}$$

The expected value of T is

$$\begin{aligned} E(T) &= E\left[\sum_{i=1}^{k} n_i \hat{\gamma}_i^2 - N\tilde{\gamma}_n^2\right] \\ &= \sum_{i=1}^{k} \frac{(N - n_i)(n_i - 1)}{N(n_i - 3)} \\ &\quad + \gamma^2 \left\{\sum_{i=1}^{k} \frac{(N - n_i)(n_i - 1)}{N(n_i - 3)} + \frac{1}{N}\left[\sum_{i=1}^{k} n_i^2 e_i^2 - \left(\sum_{i=1}^{k} n_i^2 e_i^2\right)^2\right]\right\}. \end{aligned}$$

When $\hat{\gamma}^2$ is placed instead of γ^2, $E(T)$ becomes $\hat{E}(T)$. However, when the sample size is large enough, the expected value of T is become $\hat{E}(T) \approx (1 + 1/2\,\gamma^2)(k - 1)$. According to this Non-Central t-test statistic is

$$D = (k - 1)\frac{T}{\hat{E}(T)} \tag{5}$$

[5]. The D statistic shows asymptotically χ^2 distribution with $(k - 1)$ degree of freedom.

2.2. Wald Test

$\gamma = [\gamma_1, \gamma_2, \dots, \gamma_k]^T$ vector is the unknown parameters for the inverse-coefficients of variation of k population and the $\hat{\gamma} = [\hat{\gamma}_1, \hat{\gamma}_2, \dots, \hat{\gamma}_k]^T$ vector is also the maximum likelihood estimator of these parameters. For inverse-coefficient of variation, the Wald Test is solved by similar methods. Here the null hypothesis is

$$H_0: h_1 = [\gamma_1 - \gamma_2, \gamma_2 - \gamma_3, \dots, \gamma_{k-1} - \gamma_k]^T = 0. \tag{6}$$

Let $\hat{h}_1 = [\hat{\gamma}_1 - \hat{\gamma}_2, \hat{\gamma}_2 - \hat{\gamma}_3, \dots, \hat{\gamma}_{k-1} - \hat{\gamma}_k]^T$ be the estimator of h_1 and the H matrix is defined as

$$H(\gamma) = \left\|\frac{\delta h(\gamma)}{\delta \gamma}\right\|.$$

The variance of the estimator of γ parameter is

$$\hat{V}_1 = diag\left[\left(1+{}^{\hat{\gamma}_1^2}/_2\right)/n_1, \left(1+{}^{\hat{\gamma}_2^2}/_2\right)/n_2, \dots, \left(1+{}^{\hat{\gamma}_k^2}/_2\right)/n_k\right] \quad (7)$$

[8], and the Wald test statistic is

$$WT = \hat{h}_1^T\left[H\hat{V}_1H^T\right]^{-1}\hat{h}_1 \quad (8)$$

[11]. The Wald Test statistic for the inverse-coefficients of variation has a chi-square distribution with $(k-1)$ degrees of freedom.

2.3. Likelihood Ratio Test

Under the null hypothesis H_0, the likelihood function is

$$L_0 = \prod_{i=1}^{k}\left(\frac{1}{\sigma_i\sqrt{2\pi}}\right)^{n_i}\exp\left[-\frac{1}{2}\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(\frac{x_{ij}-\gamma\sigma_i}{\sigma_i}\right)^2\right]. \quad (9)$$

Nairy and Rao [10] have shown that their derivatives

$$\sum_{i=1}^{k}\left[\frac{2n_i\hat{\gamma}_i}{-\gamma(\hat{\gamma}_i)+\sqrt{(\gamma^2)(\hat{\gamma}_i^2)+4(1+\hat{\gamma}_i^2)}} - n_i\gamma\right] = 0 \quad (10)$$

and

$$\sigma_i = -\frac{1}{2}\left[\gamma\overline{X}_i - \sqrt{(\gamma^2)\left(\overline{X}_i^2\right)+4\left(S_i^2+\overline{X}_i^2\right)}\right] \; i = 1,2,\dots,k \quad (11)$$

by using γ and σ_i parameters by utilizing the Eqs. (9). It is known that these equations do not have a numerical solution for $k > 2$. Therefore, Nairy and Rao [10] used two-step estimators instead of one-step estimators. They made a different recommendation based on two-step estimators for the test statistics.

When the estimator $\hat{\gamma}_i$ of the inverse-coefficient of variation is asymptotically normal distributed, $\tilde{\gamma} = \sum n_i\hat{\gamma}_i/n_i$ is the $\sqrt{n}$-consistent estimator of γ. According to that, in the Eqs. (10) $f(\gamma)$ is placed on the left side. In order to find the two-step estimator, if the first-order derivative of $f(\gamma)$ is taken,

$$\frac{\delta f}{\delta\gamma} = \sum_{i=1}^{k}2n_i\left\{\hat{\gamma}_i\gamma - \frac{\gamma^2\hat{\gamma}_i^2+2(1+\hat{\gamma}_i^3)}{\sqrt{(\gamma^2)(\hat{\gamma}_i^2)+4(1+\hat{\gamma}_i^2)}}\right\}$$

equality occurs. Here, instead of the γ parameter, the estimator $\tilde{\gamma}$ is placed. Then, the two-step estimator is $\hat{\hat{\gamma}} = \tilde{\gamma} - f(\tilde{\gamma})/f'(\tilde{\gamma})$. In Eqs. (11), γ is replaced by $\hat{\hat{\gamma}}$, $\hat{\hat{\sigma}}_i$ is obtained. For the inverse-coefficients of variation, the Likelihood Ratio test statistic is

$$-2\ln\lambda = \sum_{i=1}^{k}n_i\ln\left(\frac{\hat{\hat{\sigma}}_i^2}{S_i^2}\right) \quad (12)$$

[9]. The test statistic is also known as the LR test. This test statistic has asymptotically a χ^2 distribution with $(k-1)$ degree of freedom [4, 13].

2.4. Score Test

Under H_0 hypothesis, the likelihood function is

$$\ln L_0 = -\frac{1}{2}\sum_{i=1}^{k} \ln \sigma_i^2 - \frac{1}{2}\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(\frac{x_{ij}-\gamma\sigma_i}{\sigma_i}\right)^2.$$

Using the Fisher information matrix indicated by $\mathrm{I}(\gamma)$ and making the necessary simplifications, Nairy and Rao [10] proposed the Score Test statistic for the equality of the inverse-coefficients of variation. The Score Test statistic was proposed as

$$ST = \frac{\left(2+\hat{\hat{\gamma}}^2\right)}{2}\sum_{i=1}^{k}\frac{a_i^2}{n_i} \tag{13}$$

[10]. Here, $a_i = \sum_{j=1}^{n_i}\left(X_{ij}-\hat{\hat{\gamma}}\hat{\hat{\sigma}}_i\right)/\hat{\hat{\sigma}}_i$ where $\hat{\hat{\gamma}}$ and $\hat{\hat{\sigma}}_i$ are two-step estimators given in the Likelihood Ratio Test in Eqs. (10) and Eqs. (11). Score Test statistic shows χ^2 distribution with $(k-1)$ degree of freedom.

3. SIMULATION STUDY

This section covers comparisons of the tests in terms of type I error rate and power. In this simulation study, data generated from normal distributions were used, and 0.01 and 0.05 values were used for nominal type I error rate. This simulation study consists of two parts. In the first part, four tests in the study were compared in terms of type I error rate. Under H_0, sample sizes for the comparisons for type I error rate were taken as 10, 20, 30, 40 and 50, and population numbers were taken as 3 and 6. The results of power comparisons for the tests were given in the second part. In power comparisons, data generated from normal distributions were used under H_1. Population number was taken as 3, and different sample sizes were used. In the simulation study, the number of iterations for each simulation scenario is 10000 and R 3.5.1 program code was prepared for each test method. The results were presented in Figures and Tables. The algorithm steps for the simulation study are as follows:

Step 1. Select the sample sizes (n_i), type I error rates (α), the number of iterations (M) and population number (k).

Step 2. Generate a random sample sequence of the probability density function of the normal distribution with the selected scenarios under the hypothesis using R 3.5.1 program.

Step 3. Calculate the inverse-coefficient of variations test statistics values defined in section 2 by using the sequence of n_i random samples $(x_{i1}, x_{i2}, \dots, x_{in_i}, i = 1,2,\dots,k)$.

Step 4. Calculate the critical value for testing the null hypothesis and compare each simulated test statistics values with the critical value.

Step 5. Repeat (Step 3 and Step 4) M times, then compute rejection rates (the rejection number divides into M) according to (Step 4) comparisons.

3.1. Comparisons for type I error rate

In the comparisons for Type I error rate, the data generated from normal distributions with an inverse-coefficient of variation of 2.5 were used. Then, the tests given in the second section were used to test the hypothesis of the equality of the inverse-coefficients of variation. In this way, the rejection rates of the null hypothesis were calculated. These rates which are known as experimental type I error rates were calculated by dividing the rejection number of the hypothesis of equality of inverse-coefficients of variation to 10000. Nominal type I error rates for the tests were taken as 0.01 and 0.05. When the sample sizes are 10, 20, 30, 40, 50, and number of populations are 3 and 6, the experimental type I error rates for the tests were calculated, and given in Figure 1.

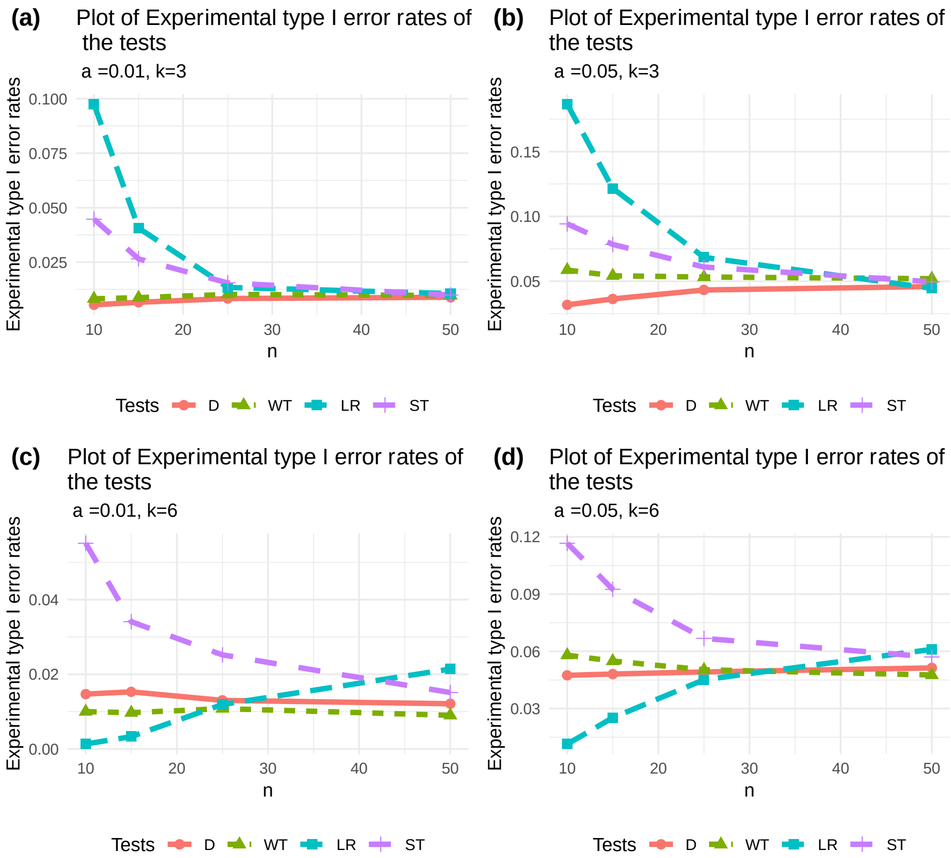

Figure 1. The experimental type I error rates for $\gamma = 2.5$ in case of different k and n values (a-d).

The results given in Figure 1 can be interpreted as follows. When the nominal I type error rate is 0.01 and population number is 3, the D and WT tests for small sample sizes gave the experimental type I error rates very close to the nominal type I error rate, but the LR and ST tests yielded poor results.

On the other hand, when the sample sizes increase, the experimental type I error rates for all tests are close to the nominal type I error rate. When the nominal I type error rate is 0.01 and population number is 6, the WT test for all sample sizes is the best test for the type I error rate, and the D test also gave good results. In this scenario, when the sample sizes increased, the experimental type I error rates for the ST test are close to nominal type I error rate, but these ratios for the LR test were quite different from 0.05.

When the nominal I type error rate is 0.05 and population number is 3, the WT test for small sample sizes gave very good results in terms of type I error rate, and these rates for the other tests were not close to 0.05. As the sample size increased, it was found that the experimental type I error rates of the D, LR and ST tests approached to nominal I type error rate, and the error rates of these tests for $n = 50$

were quite close to nominal type I error rate. Similar to the results in Figure 1 (c), when nominal type I error rate is 0.05 and population number is 6, the D and WT tests for small sample sizes showed good results in terms of type I error rate, but the LR and ST tests yielded poor results. On the other hand, although these ratios for the D and ST tests are very close to 0.05 for large sample sizes, these ratios for the LR and ST tests are about 0.06.

Consequently, although the experimental type I error rates of the LR and ST tests for small sample sizes are not close to the nominal type I error, the experimental type I error rates of all tests for large sample sizes are generally close to nominal type I error rate.

3.2. Power Comparisons

In order to compare the tests in terms of power, simulation scenarios with various sample sizes from three normal distributions were used. The cases where the sample sizes are equal and not equal were considered. When the sample sizes are not equal, the cases where the differences between the sample sizes are both small and large, and only one of the sample sizes is different is taken into account. In this simulation study, 0.05 and 0.01 values were used for nominal type I error rate. The following three scenarios are used for normal distributions having various means and variances.

$$C1 = N(50{,}400), N(50{,}225), N(50{,}225)$$
$$C2 = N(50{,}400), N(50{,}225), N(50{,}100)$$
$$C3 = N(50{,}400), N(50{,}100), N(50{,}100).$$

As can be easily seen in the above scenarios, only one of the inverse-coefficients of variations is different in C1 and C3 scenarios, but all three of these coefficients are different in the C2 scenario. Under H_1, the hypothesis of the equality of the inverse-coefficients of variation was tested for 10000 times with the data generated in each of these scenarios. These tests were performed according to the significance level of both 0.01 and 0.05. Then, the experimental power values for each of these tests were calculated by dividing the rejection number of the hypothesis of the equality of the inverse-coefficients of variation to 10000. As is known, the degree of difference of the distribution parameters, sample sizes and the type I error rate level are factors affecting the power of the test. For this reason, the above simulation scenarios have been created. The experimental power values of the tests were given in Tables 1-4.

Table 1. While the sample sizes are equal, and $\alpha = 0.01$, 0.05, the experimental power values of the tests for the C1, C2 and C3 scenarios

		C1	C2	C3	C1	C2	C3
α	**Test**	n_i =(10,10,10)			n_i =(15,15,15)		
0.01	**D**	.0112	.0634	.0465	.0228	.1945	.1980
	WT	.0329	.1315	.2920	.0503	.2682	.4967
	$-2\ln\lambda$	.0338	.4577	.2883	.0960	.8017	.5897
	ST	.0894	.1911	.3256	.0965	.2678	.4732
0.05	**D**	.0550	.2196	.2233	.0995	.4539	.5168
	WT	.1288	.3597	.5291	.1660	.5340	.7183
	$-2\ln\lambda$	.1787	.7971	.6318	.2803	.9529	.8274
	ST	.1682	.3285	.4984	.1905	.4564	.6888
α	**Test**	n_i =(25,25,25)			n_i =(50,50,50)		
0.01	**D**	.0510	.5265	.5854	.1738	.9379	.9724
	WT	.0967	.5610	.7912	.2493	.9342	.9895
	$-2\ln\lambda$	.2338	.9876	.8936	.5354	.9999	.9974
	ST	.1290	.4606	.7638	.2766	.8956	.9926
0.05	**D**	.1794	.7764	.8508	.4014	.9865	.9954
	WT	.2503	.8003	.9240	.4709	.9879	.9976
	$-2\ln\lambda$	.4527	.9968	.9674	.7083	.9999	.9997
	ST	.2697	.7274	.9271	.4869	.9886	.9995

As seen above, the experimental power values in Table 1 are for equal sample sizes. In the C2 scenario where three of the inverse-coefficients of variation are different, the $-2\ln\lambda$ test is the most powerful test for both levels of significance. Similar results were obtained in the C1 and C3 scenarios where only one of the inverse-coefficients of variation was different. Except in the case of equal sample sizes 10 and 15, and $\alpha = 0.01$, in all other cases it is seen that the $-2\ln\lambda$ test is the most powerful test for both levels of significance. While each of the sample sizes was 10 and $\alpha = 0.01$, the ST test gave the largest power values in the C1 and C3 scenarios. On the other hand, while $\alpha = 0.05$, the ST test had the greatest power in the C1 scenario, while the $-2\ln\lambda$ test gave the greatest power value in the C3 scenario. As expected, the powers of the tests increased as the sample sizes increased. Similarly, when the nominal type I error rate increases, it seems that the power of the tests increases as expected.

As seen in Table 1, whereas three of the inverse-coefficients of variation were different, the D test for small sample sizes and the ST test for large sample sizes gave the worst results in terms of power. On the other hand, when only one of the inverse-coefficients of variation is different, it was observed that the D test gave the worst results in terms of power in all the scenarios considered.

Table 2. While the sample sizes are all different and $\alpha = 0.01, 0.05$, experimental power values of the tests for the C1, C2 and C3 scenarios

		C1	C2	C3	C1	C2	C3
α	**Test**	n_i =(9,10,11)			n_i =(14,15,16)		
0.01	**D**	.0160	.0956	.0735	.0294	.2363	.2381
	WT	.0350	.1490	.2872	.0526	.2820	.4928
	$-2\ln\lambda$	.0396	.4884	.2900	.1044	.8182	.5920
	ST	.0667	.1782	.3037	.0843	.2602	.4615
0.05	**D**	.0809	.2812	.3016	.1199	.4927	.5651
	WT	.1326	.3781	.5195	.1679	.5450	.7140
	$-2\ln\lambda$	.1994	.8078	.6208	.2937	.9452	.8224
	ST	.1397	.3083	.4558	.1675	.4460	.6630
α	**Test**	n_i =(24,25,26)			n_i =(49,50,51)		
0.01	**D**	.0591	.5617	.6192	.1874	.9429	.9704
	WT	.0987	.5695	.7859	.2478	.9402	.9874
	$-2\ln\lambda$	.2450	.9858	.8999	.5395	.9999	.9968
	ST	.1164	.4580	.7483	.2626	.9004	.9911
0.05	**D**	.1990	.7991	.8587	.4055	.9863	.9952
	WT	.2507	.8068	.9127	.4563	.9863	.9980
	$-2\ln\lambda$	.4455	.9980	.9581	.7013	.9998	.9989
	ST	.2497	.7206	.9061	.4643	.9853	.9993

Table 2 includes the experimental power values of the tests while the sample sizes are different. Similar to the results in Table 1, in the C2 scenario where all three inverse-coefficients of variation are different, the $-2\ln\lambda$ test is clearly the most powerful test. Except the cases where the sample sizes are 9,10,11 and $\alpha = 0.01$, the $-2\ln\lambda$ test again gave the highest power values. However, it was seen that the ST test gave the greatest power values when the sample sizes are 9,10,11 and $\alpha = 0.01$. As expected, when the nominal type I error rate increases, it seems that the power of the tests increases. When the results in Table 2 are examined, it is seen that the power of the tests increases as the sample sizes increase.

On the other hand, when all of the inverse-coefficients of variation are different, the ST and D tests generally gave lower power values. However, when only one of these coefficients is different, the D test gave the lowest power values.

Table 3. While the sample sizes are all different and $\alpha = 0.01, 0.05$, experimental power values of the tests for the C1, C2 and C3 scenarios

		C1	C2	C3	C1	C2	C3
α	**Test**	n_i =(3,13,14)			n_i =(4,20,21)		
0.01	**D**	.1690	.2830	.2647	.1976	.4416	.4284
	WT	.0296	.1382	.1791	.0384	.2267	.2559
	$-2\ln\lambda$	.1389	.5497	.1986	.1579	.6933	.2643
	ST	.1146	.1355	.1348	.0848	.1571	.1400
0.05	**D**	.3096	.4911	.4626	.3515	.6428	.6010
	WT	.1080	.3343	.3113	.1200	.4629	.4019
	$-2\ln\lambda$	.3237	.7090	.4037	.3169	.8109	.4681
	ST	.1575	.2447	.2338	.1384	.3234	.2605
α	**Test**	n_i =(6,34,35)			n_i =(9,70,71)		
0.01	**D**	.2627	.6660	.6403	.4064	.9195	.8517
	WT	.0564	.4244	.3868	.0756	.7907	.5701
	$-2\ln\lambda$	.2215	.8733	.4086	.3372	.9747	.5819
	ST	.0606	.2393	.1659	.0566	.6370	.2089
0.05	**D**	.4142	.8240	.7500	.5479	.9673	.8942
	WT	.1450	.6807	.5384	.1917	.9193	.7022
	$-2\ln\lambda$	.3600	.9274	.5657	.4377	.9840	.6706
	ST	.1155	.5272	.3069	.1274	.8858	.3876

As seen above, the differences between the sample sizes in Table 3. are generally larger than those of Table 2. In the C2 scenario where all of inverse-coefficients of variation are different, it is clear that the $-2\ln\lambda$ test is the most powerful test. In the C1 and C3 scenarios where one of the inverse- coefficients of variation is different, the D test has the greatest power values. As expected, when both the nominal I-type error rate and the sample sizes increase, it is seen that the power values of the tests increase.

On the other hand, when all of the inverse-coefficients of variation were different, the ST test gave the lowest power values. In the C1 and C3 scenarios where one of the inverse-coefficients of variation were different, the WT test for small sample sizes gave lower power values than the other tests. However, similar to the results from previous tables, it was found that the ST test for large sample sizes gave lower power values than the other tests.

Table 4. While one of the sample sizes is different and $\alpha = 0.01, 0.05$, experimental power values of the tests for the C1, C2 and C3 scenarios

k =3		C1	C2	C3	C1	C2	C3
α	**Test**	n_i =(4,4,22)			n_i =(10,10,25)		
0.01	**D**	.4044	.5823	.5894	.1797	.5451	.6041
	WT	.0318	.2284	.2137	.0505	.3822	.4442
	$-2\ln\lambda$	.3151	.5395	.4803	.2263	.8136	.6502
	ST	.1231	.1116	.1566	.0309	.1754	.2585
0.05	**D**	.5573	.7424	.7406	.3440	.7545	.7780
	WT	.1165	.4250	.3864	.1570	.6238	.6500
	$-2\ln\lambda$	.4259	.6641	.5958	.3912	.9107	.7915
	ST	.1704	.2176	.2675	.0842	.3933	.4429
α	**Test**	n_i =(14,14,47)			n_i =(15,15,120)		
0.01	**D**	.3498	.8539	.8617	.6517	.9830	.9714
	WT	.0850	.6501	.6493	.1140	.8356	.7666
	$-2\ln\lambda$	.3732	.9382	.8101	.4897	.9270	.7918
	ST	.0257	.3331	.3584	.0571	.4061	.3280
0.05	**D**	.5104	.9337	.9313	.7624	.9932	.9818
	WT	.2101	.8339	.8095	.2524	.9259	.8686
	$-2\ln\lambda$	.5017	.9671	.8736	.5465	.9465	.8263
	ST	.0817	.6127	.5951	.1329	.6379	.5488

Table 4 contains the experimental power values of the tests when only one of the sample sizes is different and quite larger than the others. In the C2 scenario in which all of inverse-coefficients of variation were different, the D and $-2\ln\lambda$ tests gave higher power values than the other two tests. In the C1 and C2 scenarios where only one of the inverse-coefficients of variation is different, the DT test gave the greatest power values for most cases. However, it was observed that the $-2\ln\lambda$ test for the sample sizes of 10, 10 and 25 generally yielded better power results than the other tests. Therefore, while the sample sizes and type I error rate increase, it is seen that the powers of the tests generally increase as expected.

On the other hand, while all of inverse-coefficients of variation were different, the ST test yielded the worst results in terms of power. Except the C1 scenario, where the sample sizes were 4, 4, and 22, the ST test yielded the worst results in all scenarios considered. In the case of C1 where the sample sizes are 4, 4 and 22, the WT test yielded the worst power values.

4. REAL DATA APPLICATION

In real data application, the data collected by Gezgen et al. [6] from 188 patients who underwent bone marrow transplantation in bone marrow transplantation units between 2010 -2016 were used. Data were obtained from bone marrow transplantation units by using a retrospective study. According to the experts in the bone marrow transplantation unit and the previous scoring studies, they grouped the diseases into risk factors.

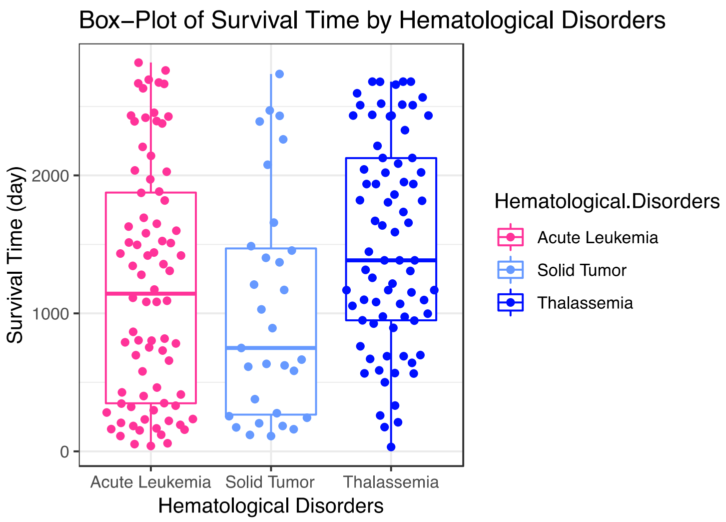

Figure 2. Box-plot of survival time according to hematological disorders

It was aimed to investigate whether there is a statistical difference among the variations of survival times of patients according to the grouped diseases. Box-plot in Figure 2 was drawn to observe the structure of the distributions. According to this box-plot, it was observed that the distributions of Acute Leukemia and Solid Tumor were right-skewed, and Thalassemia distribution is left-skewed. The results of descriptive statistics and analysis of disease groups are given in Table 5.

Table 5. Results of descriptive and test statistics of hematological diseases

Hematological Disorders	n	$\overline{X}_i$	S_i	$\hat{\gamma}_i$	D	WT	LR	ST
Thalassemia	77	1512.377	749.8169	2.0169	4.1155	7.8791[a]	9.7793[a,b]	7.2223[a]
Acute Leukemia	80	1208.925	862.3277	1.4019				
Solid Tumor	31	1032.839	820.5978	1.2586				

[a,b] Denotes statistical significance level at 0.05 and 0.01

As can be seen from the results of the analysis in Table 5, the equality of the inverse-coefficients of variation with the WT and ST tests could not be rejected at the significance level of 0.01. But, it is seen that the hypothesis of the equality of inverse-variation coefficients will be rejected with these tests according to the p-values in Table 5 at significance level of 0.05. On the other hand, the equality of inverse-variation coefficients by the D test cannot be rejected at both significance levels of 0.01 and 0.05. Besides, it is clear that the hypothesis of the equality of inverse-variation coefficients with LR test will be rejected at both significance levels of 0.01 and 0.05. As it will be remembered, according to the simulation results in the previous section, the most powerful test was found to be the LR ($-2\ln\lambda$) test in general. It is understood that the result of this application supports the simulation findings.

5. CONCLUSIONS

This section is on the interpretation of the results of the simulation study for the comparison of the tests in terms of type I error rate and power**.** When the tests were compared with respect to type I error rate, it was observed that the D and WT tests for small sample sizes gave the experimental type I error rates quite close to the nominal type I error rate in all scenarios considered. In these sample sizes, it was seen that the LR ($-2\ln\lambda$) test in case where number of populations is 3, and ST test in case where number of populations were 6 yielded the worst experimental type error I rates. In scenarios with a population number of 3, it is understood that the experimental type I error rates of the LR and ST tests converge to the nominal type I error, and that the experimental type I error rates of all tests for $n = 50$ are almost equal to the nominal type I error. In scenarios where the number of populations is 6, it is seen that the experimental type I error rates of the D and WT tests for large sample sizes are quite close to the nominal type I error. On the other hand, the experimental type I error rates of LR and ST tests for these sample sizes have improved slightly, but these ratios for the ST test are slightly different from the nominal type I error.

As a result, although the experimental type I error rates of the LR and ST tests for small sample sizes are not close to the nominal type I error, the experimental type I error rates of all tests for large sample sizes are generally close to nominal type I error rate.

In scenarios where the sample sizes are equal and the inverse-coefficients of variation are different, the $-2\ln\lambda$ test is the most powerful test. In this scenario, the D test for small sample sizes yielded poor results in terms of power, whereas the ST test for large sample sizes yielded poor results. On the other hand, when only one of the inverse-coefficients of variation is different, the most powerful test is the -2lnL test, while the worst test is the D test.

In scenarios where the sample sizes are different but close to each other (Table 2.) and three of the inverse-coefficients of variation are different, the $-2\ln\lambda$ test is generally powerful than the other tests. For this scenario, the ST and D tests gave very low experimental power results. When only one of the inverse-coefficients of variation was different, it was observed that the $-2\ln\lambda$ test was the most powerful test and the D test had the lowest power.

In the scenario where the sample sizes are different and one is quite small (Table 3.) compared to the others, when all of the inverse-coefficients of variation are different, it was found that the most powerful test was the $-2\ln\lambda$ test and the ST test had the lowest power results. When only one of the inverse-coefficients of variation was different, it was observed that the D test was the most powerful test and the ST test had the lowest power.

In the scenario where the sample sizes are different and one is quite larger (Table 4.) compared to the others, when all of the inverse-coefficients of variation are different, the D and $-2\ln\lambda$ tests were found to be powerful than the other tests and the ST test yielded the worst results. On the other hand, when only one of the inverse-coefficients of variation was different from the others, the D test was generally better than the other tests in terms of power, and the ST test was the test with the lowest power.

Survival data of 188 patients who underwent bone marrow transplantation from bone marrow transplantation units were collected by a retrospective study. In order to determine whether there is any difference among the variations of survival times in the disease groups, the inverse coefficient of variation was used. Then, the hypothesis of the equality of the inverse-coefficients of variation was tested. According to the results of the analysis, the LR statistic was significant at both significance levels, but the D test was not significant at the significance levels of 0.01 and 0.05 (Table 5.). Thus, it is understood that these results are consistent with the simulation results.

REFERENCES

[1] Bennett BM. On an Approximate Test for Homogeneity of Coefficients of Variation. Contributions to Applied Statistics 1976; 169-171.

[2] Bowman KO, Shenton LR. Moment series for the coefficients of variation in Weibull sampling. American Statistical Associations 1981;148-153.

[3] Chaturvedi A, Rani U. Fixed width confidence interval estimation of the inverse coefficient of variation in a normal population. Microelectronics and Reliability 1996; 36: 1305-1308.

[4] Cox DR, Hinkley DV. Theoretical Statistics, Chapman and Hall. Londan, 111-119, 1974.

[5] Doornbos, R., Dijkstra, J.B. A Multi-Sample Test for The Equality of Coefficients of Variation in Normal Population. Commun.Statist-Simula.Computa. 1983; 12: 147-158.

[6] Gezgen Kesen, G., Potas, N., Visal Okur, F., Ok Bozkaya, I. Aksu, T, Kuşkonmaz, B.B., Uçkan Çetinkaya, D. Özbek, N. Y. Çocuklarda Hematopoetik Kök Hücre Naklisonrasi Prognozu Belirlemede Modifiye Ebmt Risk skorlamasinin Kullanımı. 10. Ulusal Kemik İliği Transplantasyonu ve Hücresel Tedaviler Kongresi; 1-3 Mart 2018; Antalya. Türkiye: pp. 99.

[7] Kalkur AT, Rao A. Bayes Estimator for Coefficient of Variation and Inverse Coefficient of Variation for the Normal Distribution. International Journal of Statistics and Systems 2017; 12:721-732.

[8] Kendall M, Stuart A. The Advanced Theory of Statistics. Charles Griffin and Co., Londan, 205-210, 1977.

[9] Lehmann, E.L., Casella, G. Theory of Point Estimation 2ed. Springer-Verlag, New York, 454-457,1998.

[10] Nairy KS, Rao KA. Test of Coefficients of Variation of Normal Population. Commun.Statist-Simula.Computa. 2003; 32: 641-661.

[11] Rao KA, Vidya R. On Performance of a Test for Coefficients of Variation. Calcutta Statistical Association Bulletin 1992; 42: 87-95.

[12] Sharma KK, Krishnan H. Asymptotic sampling distribution of inverse coefficient of variation and its Applications. IEEE- Transactions on Reliability 1994; 43:630-633.

[13] Silvey SD. Statistical Inference. Chapmann & Hall, London 130-135, 1975.

[14] Singh M. Behaviour of Sample Coefficient of Variation Drawn From Several Distributions. Sankhya 1993; 55:65-76.