Süleyman Demirel Üniversitesi
Fen Bilimleri Enstitüsü Dergisi
Cilt 22, Sayı 3, 1125-1131, 2018

Süleyman Demirel University
Journal of Natural and Applied Sciences
Volume 22, Issue 3, 1125-1131, 2018

DOI: 10.19113/sdufenbed.469282

A Monte Carlo Simulation Study Robustness of MANOVA Test Statistics in Bernoulli Distribution

Mustafa ŞAHİN[1], Şeyma KOÇ[*1]

[1]Kahramanmaraş Sütçü İmam University, Department of Agriculture, 4600, Kahramanmaraş

(Alınış / Received: 26.04.2018, Kabul / Accepted: 28.09.2018, Online Yayınlanma / Published Online: 10.10.2018)

Keywords
Manova test statistics,
Simulation study,
Monte Carlo

Abstract: The aim of this study is to compare the robustness of Manova test statistics against Type I error rate using the Monte Carlo simulation technique. In the method, numbers are generated according to constant and increasing variance for *g=3,4,5* group *p=3,5,7* dependent variables *n=10,30,60* sample size using the R. Numbers have been produced using these 54 combinations. Pillai Trace test statistic has been the least deviating from the nominal α =0.05 value. Wilk Lambda and Hotelling-Lawley Trace test results were close to each other. The researchers can decide according to the comparison results of the analysis's suggested decision stage.

MANOVA Test İstatistiklerinin Monte-Carlo Simülasyonu ile Bernoulli Dağılımında Karşılaştırılması

Anahtar Kelimeler
Manova test istatistiği,
Simülasyon çalışması,
Monte Carlo

Özet: Bu çalışmanın amacı, Manova test istatistiklerinin sağlamlığını Monte Carlo simülasyonunu kullanılarak I.tip hata bakımından kıyaslamaktır. Yöntemde, sayılar g = 3,4,5 grup için p = 3,5,7 bağımlı değişkene ait n = 10,30,60 örneklem büyüklüğü kullanılarak sabit ve artan varyansta R programlama dili kullanılarak üretilmiştir. 54 kombinasyonda hesaplanan I.Tip hatalardan, nominal α =0.05 değerinden en az uzaklaşan test istatistiği Pillai İz test istatistiği olmuştur. Wilk Lambda ve Hotelling-Lawley İz test istatistikleri ise birbirlerine yakın sonuç vermişlerdir. Araştırıcılar analizlerinin karar aşamasında önerilen kıyaslama sonuçlarına göre karar verebilirler.

1. Introduction

The one-way multivariate analysis of variance (one-way MANOVA) is used to determine whether there are any differences between independent groups on more than one continuous dependent variable. The most important assumptions are multivariate normality and homogeneity of variance-covariance matrices. The most well known and widely used MANOVA test statistics are Wilk's Λ, Pillai, Lawley-Hotelling, and Roy's test.

Wilk's Λ: Wilks' lambda [1] is a test statistic used in multivariate analysis of variance (MANOVA) to test whether there are differences between the means of identified groups of subjects on a combination of dependent variables. Wilks' lambda is the oldest multivariate test statistic, and is the most widely used.

Let,
T: Total sums of squares and cross-products matrix,
B: Between-group sums of squares and cross-products matrix,
W: Within-group sums of squares and cross-products matrix,
p: Number of dependent variables in each groups,
g: The number of groups g ≥ 2.
$\bar{x}$: Overall sample mean vectors,
n_i : sample size for the i-th group,
S_i : sample covariance matrix for the i − th sample
Thus B and W matrix can be expressed by
$B = \sum_{i=1}^{g} n_i(\bar{x}_i - \bar{x})(\bar{x}_i - \bar{x})'$

$$W = \sum_{i=1}^{g}(n_i - 1)\, S_i \quad (1)$$

The Wilks' Lambda statistic is the ratio of the within generalized dispersion to the total generalized dispersion

$$\Lambda = \frac{|W|}{|B+W|} = \frac{|W|}{|T|} \quad (2)$$

takes values between zero and one. The Wilks' Lambda can be obtained as a product of eigenvalues which can be obtained by the eigenvalues of the matrix of BW^{-1} by following method

$\Lambda = \prod_{i=1}^{s} \frac{1}{1+\lambda_i}$ where s = min(p, g−1) and the rank of the B matrix and the expression λ_i are eigenvalues of the BW^{-1} matrix. According to Johnson and Wichern [2] the Wilks' Lambda performs, in a multivariate setting, with a combination of dependent variables -

the same role as the F-test performs in a one-way analysis of variance. Bartlett, M.S. [3] using a chi-square test instead of an F-distribution test. Bartlett's test is a modi cation of the corresponding likelihood ratio test designed to make the approximation of thechi-square distribution better at all stages as formuled,

$$V = -[N - 1 - (p + g)/2] ln\Lambda \qquad (3)$$

Hotelling-Lawley Trace (T): The Hotelling ve Lawley Trace , statistic which defined as follows [4, 5],

$$T = trace(BW^{-1}) = \sum_{i=1}^{s} \lambda_i \qquad (5)$$

The F distribution can be used to test the T statistic [5]. T is the trace of the BW^{-1} matrix [6],

Pillai's Trace (V): Pillai trace statistic can be interpreted as the proportion of variance in the dependent variables which is accounted for by variation in the independent variables [7]. The V statistics where s, m, n parameters are as follows;

$$s = min\,(g-1, p), \begin{Bmatrix} m = \frac{|p-(g-1)|-1}{2} \\ n = \frac{N-p-g-1}{2} \times \frac{2n+s+1}{2m+s+1} \times \frac{V}{s-V} \end{Bmatrix} (6)$$

closed F distribution with s(2m+s+1*)* and (2n +s+1) degrees of freedom [8], [9].

Roy's Largest Root (R): If the big eigenvalue of the matrix of BW^{-1} is denoted by λ_{max} Roy's R statistic is given by

$$R = \sum_{i=1}^{s} \frac{\lambda_{max}}{1+\lambda_{max}} \qquad (7)$$

This value is compared to the Heck graph value with parameter s, m, n. If the R statistic is greater than the Heck graph value, it is said to be the difference between the mean vectors [10]. When s = 1, R shows exact F distribution [11].

2. Material and Method

This investigation deals mainly to assess the robustness of MANOVA. To do is the Multivariate Normality assumption was violated see if that will affect Type I error rate. In order to evaluate the robustness of MANOVA the virtual experiment was designed in the following way. In the method, numbers are generated according to constant and increasing variance for *g=3,4,5* group *p=3,5,7* dependent variables *n=10,30,60* sample size using the R. For the significance test 2160000 numbers have been produced using these combinations. That simulation was based on 10,000 replications

The Monte Carlo study manipulated in equal variance ($\sigma_1^2 = \sigma_2^2 = \ldots = \sigma_g^2$) and unequal variance. When establishing the unequal variance, the variance of a dependent variable was first set, then the other dependent variables were multiplied by 3, that mean variance ratio is (1:3). All of the statistical methods were conducted using R (MVNormTest written by Slawomir on 04/12/2012: Normality test for multivariate variables package). In order to test the hypothesis used to compare the mean of more than two groups the Wilks' Lambda(W), Pillai's Trace(V), Hotelling-Lawley Trace(T), Roy's Largest Root test(R) statistics values and their Type I error rate were calculated. If p-value was less than 0.05, the nominal alpha level, the null hypothesis was rejected. The data are produced in the Bernoulli distribution. Scenarios were prepared in 54 different combinations for each test statistic. These operations were repeated 10,000 times and the number of null hypothesis rejections was determined for each test statistic. Experimental Type I error rates were calculated for each test statistic with dividing the rejection number by the repeat number.

3. Result

Monte Carlo test result for R,V,T,W test statistics is given respectively Table 1, 2 and 3.

When group number is g=3, for all values of p, observations are interpreted according to sample size of four test statistics with Figure 2, 3 and 4.

For the Roy test statistic, deviations from Type I error were found to decrease with both constant and increasing variance sample size (n value) and variable number (p=3, p=5, p=7). For Roy test statistic in g=3, the highest deviation was seen in all scenarios when p=3 n=10, constant variance with 0.0592 value.

For Pillai when p=3 per group, both constant and increasing variance, deviations from nominal signifcance level, α =0.05, decrease as the number of sample size (n value) increase. For Pillai test statistic in g=3, the highest deviation was seen in all scenarios when p=5, n=30, in constant variance with 0.0575 value.

For Hotelling-Lawley when p=3 per group, most deviations is seen when the sample size n=60 both constant and increasing variance. When p = 5, the greatest deviation is seen when n = 30, both constant and increasing variance again. As the number of variables p = 7 the highest deviation is seen; when n = 30 for the constant variance and when n = 10 for the increasing variance. For Hotelling-Lawley test statistic in g=3, the highest deviation was seen in all scenarios when p=5, n=30, in constant variance with 0.0584 value.

Table 1. For g=3, p=3, 5, 7; sample size n=10,30,60 experimental type ı error rate with 10000 raplicate

Number of Group	The number of variables	Status of Variance	Sample Size	Roy (R)	Pillai's trace (Y)	Hotelling-lawley (T)	Wilks lambda (W)
3	3	constant	10	0,0592	0,0544	0,0522	0,0572
			30	0,0521	0,0542	0,0519	0,0502
			60	0,0537	0,0538	0,0542	0,0512
		Increase	10	0,0563	0,0549	0,0551	0,0608
			30	0,0541	0,0549	0,0548	0,0533
			60	0,055	0,0494	0,0568	0,0503
	5	constant	10	0,0549	0,0521	0,058	0,0517
			30	0,0526	0,0575	0,0584	0,0558
			60	0,0517	0,0543	0,0484	0,0511
		Increase	10	0,0563	0,0522	0,0533	0,0548
			30	0,0527	0,0549	0,0498	0,0524
			60	0,0545	0,0503	0,0525	0,0549
	7	constant	10	0,0515	0,0509	0,052	0,0529
			30	0,0556	0,055	0,0545	0,0518
			60	0,054	0,0534	0,053	0,0491
		Increase	10	0,058	0,0568	0,0561	0,0547
			30	0,053	0,0573	0,0535	0,0542
			60	0,0517	0,0511	0,0528	0,0494

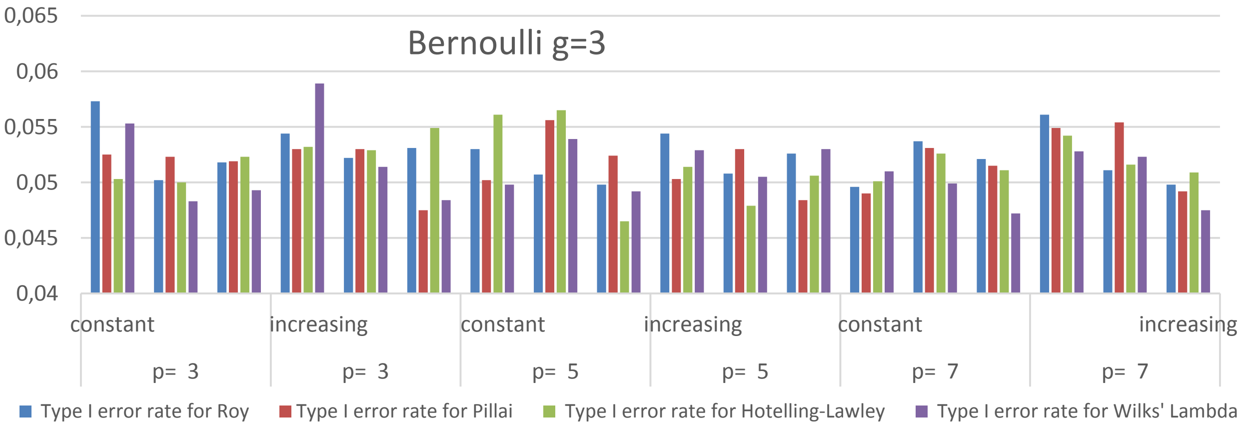

Figure 1. When the group number is g= 3, the states of four test statistic on p=3,5,7

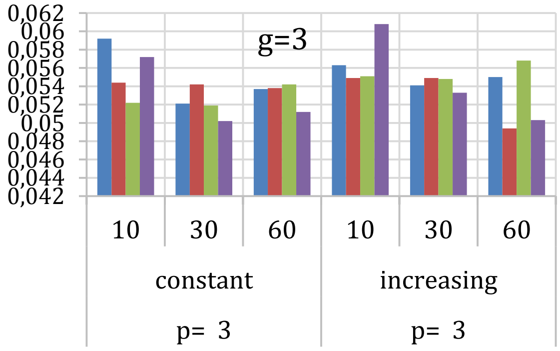

Figure 2. Display for Type I error rates for p = 3

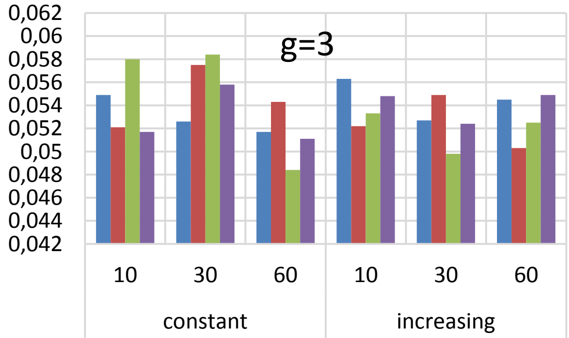

Figure 3. Display for Type I error rates for p = 5

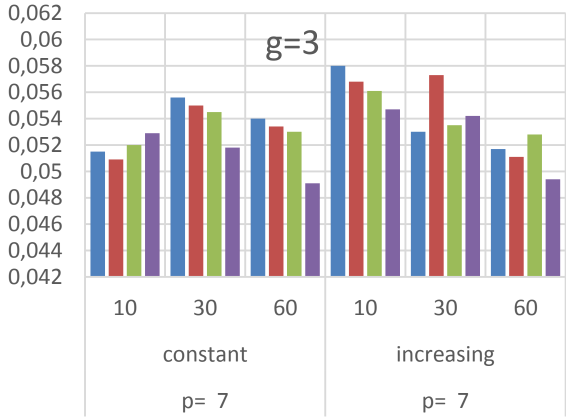

Figure 4. Display for Type I error rates for p = 7

For Wilks Lambda when p=3 per group, both constant and increasing variance, the highest deviation is seen when n=10. As p=5 the highest deviation is seen; when n = 30 for the constant variance and when n = 10 for the increasing variance. . As p=7 both constant and increasing variance, the highest deviation is seen when n=10. For Wilks Lambda test statistic in g=3, the highest deviation was seen in all scenarios when p=3, n=10, in constant variance with 0.0608 value.

Table 2. For g=4, p=3, 5, 7; sample size n=10,30,60 experimental Type I error rate with 10000 raplicate

Number of Group	The number of variables	Status of Variance	Sample Size	Roy (R)	Pillai's trace (Y)	Hotelling-lawley (T)	Wilks lambda (W)
3	3	constant	10	0,0535	0,0499	0,0545	0,0549
			30	0,0599	0,0548	0,0505	0,0552
			60	0,0517	0,0497	0,0508	0,0537
		Increase	10	0,0526	0,052	0,0556	0,0506
			30	0,0532	0,0519	0,0518	0,0528
			60	0,0531	0,0528	0,0502	0,0513
	5	constant	10	0,0534	0,0506	0,0565	0,0551
			30	0,0543	0,0498	0,0533	0,0533
			60	0,0544	0,0502	0,0518	0,0503
		Increase	10	0,0543	0,0541	0,0539	0,0567
			30	0,0518	0,0537	0,0521	0,054
			60	0,0536	0,0508	0,0523	0,0542
	7	constant	10	0,0561	0,0553	0,0517	0,0508
			30	0,0559	0,0538	0,055	0,0526
			60	0,0505	0,0537	0,0513	0,0508
		Increase	10	0,0511	0,0551	0,0567	0,0545
			30	0,0581	0,0526	0,0577	0,0544
			60	0,0557	0,052	0,053	0,0539

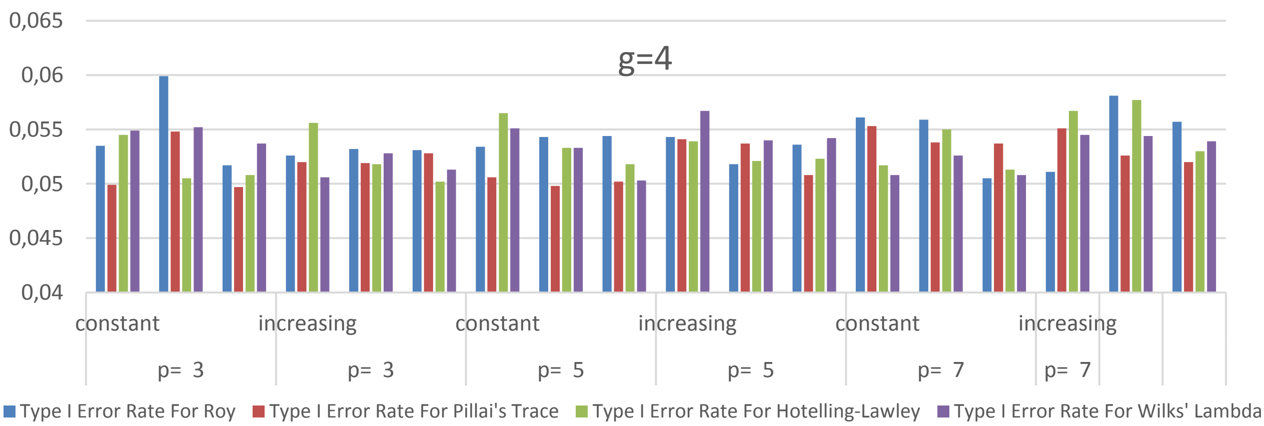

Figure 5. When the group number is g= 4, the states of four test statistic on p=3,5,7

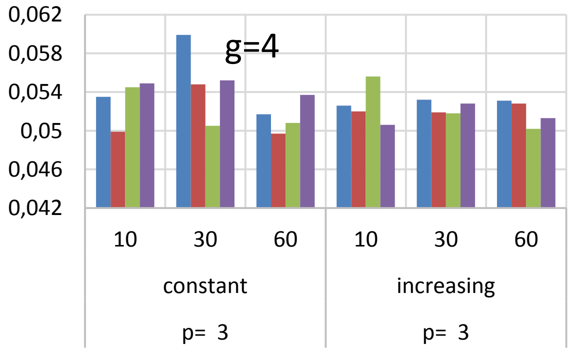

Figure 6. Display for Type I error rates for p = 3

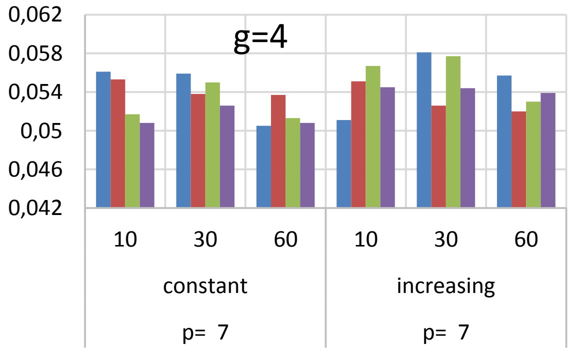

Figure 8. Display for Type I error rates for p =7

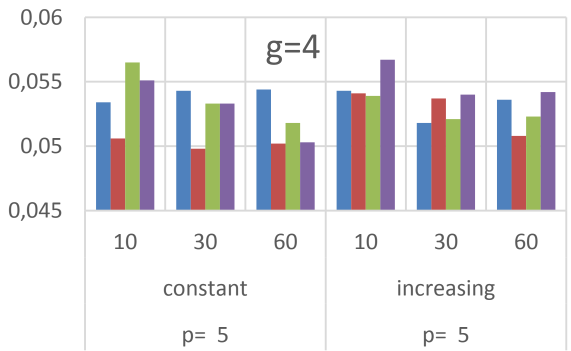

Figure 7. Display for Type I error rates for p = 5

When group number is g=4, for all values of p, observations are interpreted according to sample size of four test statistics with Figure 6, 7 and 8.

For Roy as p=3 per group, both constant and increasing variance, deviations from nominal signifcance level, α =0.05, at when n=30. As p = 5, the greatest deviation is seen when n = 60 for the constant variance, and when n = 10 for the increasing variance. As the number of variables p = 7, the highest deviation is seen when n=10 for constant variance and when n=30 for the increasing variance.

For Roy test statistic in g=4, the highest deviation was seen in all scenarios when p=3, n=30, in constant variance with 0.0599 value.

For Pillai'sTrace when p=3 per group, most deviations is seen when the sample size n=30 for the constant variance and when n = 60 for the increasing variance. When p = 5, 7 the greatest deviation is seen when n = 30, both constant and increasing variance. For Pillai test statistic in g=4, the highest deviation was seen in all scenarios when p = 7 n = 10 with 0.0553 value. For Pillai test statistic in g=4, the highest deviation was seen in all scenarios when p=7, n=10, in constant variance with 0.0553 value.

For Hotelling-Lawley when p=3 and p=5 per group, both constant and increasing variance, the highest deviation is seen when n=10. As p=7 both constant and increasing variance, the highest deviation is seen when n=30. For Hotelling-Lawley test statistic in g=4, the highest deviation was seen in all scenarios when p=7, n=30, in increasing variance with 0.0577 value.

For Wilks'Lambda when p=3 per group, both constant and increasing variance, the highest deviation is seen when n=30. The number of variables p=5 per group, both constant and increasing variance, the highest deviation is seen when n=10. As the number of variables p = 7 the highest deviation is seen; when n = 30 for the constant variance and when n=10 for the increasing variance. For Wilks Lambda test statistic in g=3, the highest deviation was seen in all scenarios when p=5, n=10, in increasing variance with 0.0567 value.

Table 3. For g=5, p=3, 5, 7; sample size n=10,30,60 experimental Type I error rate with 10000 raplicate

Number of Group	The number of variables	Status of Variance	Sample Size	Roy (R)	Pillai's trace (Y)	Hotelling-lawley (T)	Wilks lambda (W)
			10	0,0559	0,0521	0,0517	0,0536
		constant	30	0,0505	0,0524	0,0535	0,051
	3		60	0,0537	0,0523	0,0539	0,0531
			10	0,0535	0,053	0,0495	0,0566
		Increase	30	0,0522	0,0542	0,0538	0,054
			60	0,0549	0,051	0,0544	0,0524
			10	0,0502	0,0508	0,0519	0,0537
		constant	30	0,0518	0,0569	0,0538	0,0555
3	5		60	0,0551	0,0534	0,0531	0,0513
			10	0,0553	0,0517	0,0549	0,0531
		Increase	30	0,0493	0,0462	0,0543	0,0585
			60	0,0543	0,0518	0,0473	0,0496
			10	0,0531	0,0507	0,0541	0,0537
		constant	30	0,0508	0,051	0,055	0,0542
	7		60	0,0487	0,05	0,0508	0,0539
			10	0,052	0,0521	0,0519	0,0565
		Increase	30	0,0524	0,0559	0,0583	0,053
			60	0,0531	0,0518	0,0534	0,0556

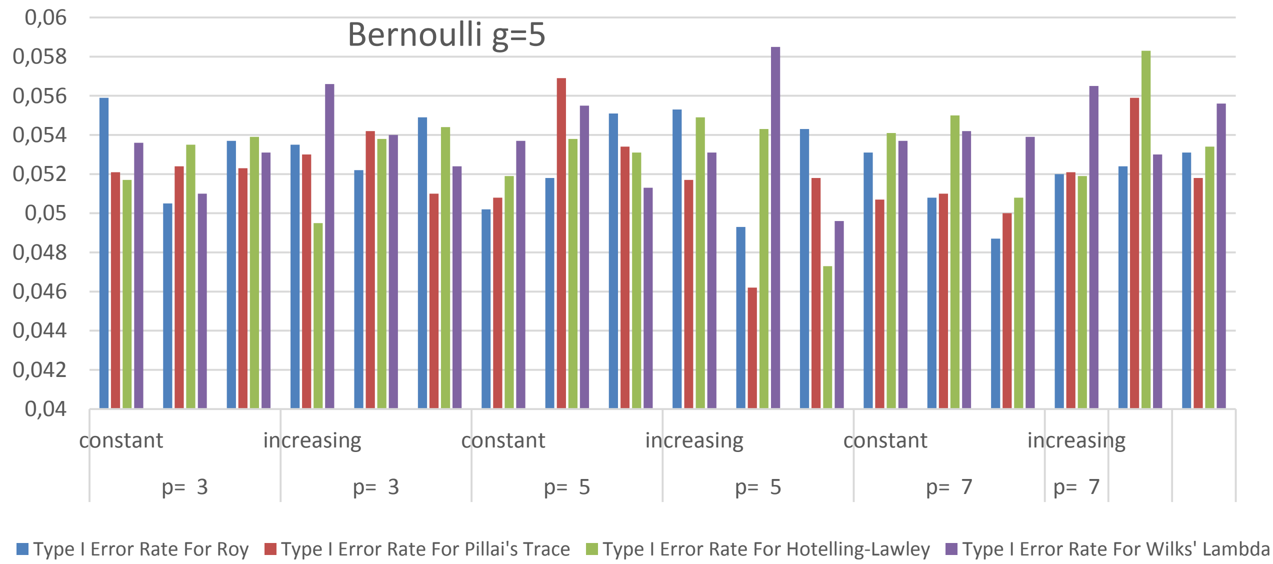

Figure 9. When the group number is g= 5, the states of four test statistic on p=3,5,7

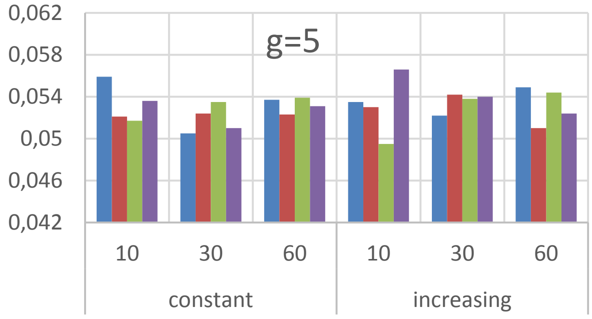

Figure 10. Display for Type I error rates for p = 3

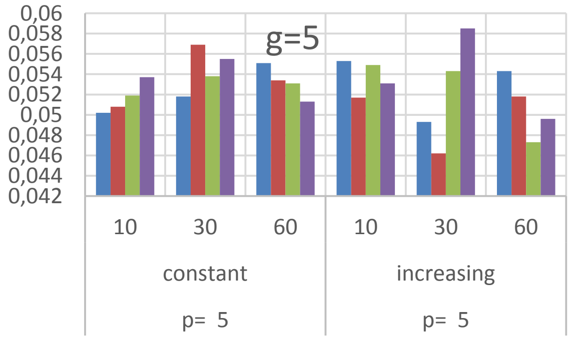

Figure 11. Display for Type I error rates for p = 3

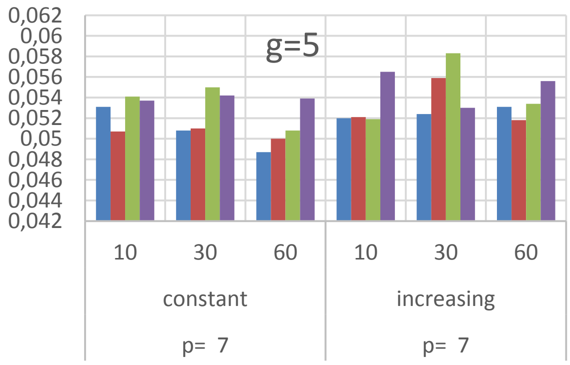

Figure 12. Display for Type I error rates for p =7

When group number is g=5, for all values of p, observations are interpreted according to sample size of four test statistics with Figure 9, Figure 10, Figure 11.

For Roy as p=3 the greatest deviation is seen when n = 30 for the constant variance, and when n = 60 for the increasing variance. As p = 5, the greatest deviation is seen when n = 60 for the constant variance, and when n = 10 for the increasing variance. As the number of variables p = 7, the greatest deviation is seen when n = 10 for the constant variance, and when n = 60 for the increasing variance. For Roy test statistic in g=5, the highest deviation was seen in all scenarios when p=3, n=10, in constant variance with 0.0599 value.

For Pillai's Trace when p=3,5,7 per group, both constant and increasing variance, the highest deviation is seen when n=30. For Pillai test statistic in g=5, the highest deviation was seen in all scenarios when p=5, n=30, in constant variance with 0.0569 value.

For Hotelling-Lawley when p=3 per group, both constant and increasing variance, the highest deviation is seen when n=60. As p=5 the highest deviation is seen; when n = 30 for the constant variance and when n=10 for the increasing variance. As the number of variables p = 7 the highest deviation is seen; when n=30 for both constant and increasing variance. For Hotelling-Lawley test statistic in g=5, the highest deviation was seen in all scenarios when p=7, n=30, in increasing variance with 0.0583 value.

For Wilks'Lambda when p=3 the highest deviation is seen; when n = 10 for both constant and increasing variance. When p = 5, the greatest deviation is seen when n =30 for the constant variance and when n=10 for the increasing variance. When p=7 the highest deviation is seen; when n = 30 for both constant and increasing variance. For Wilks Lambda test statistic in g=5, the highest deviation was seen in all scenarios when p=7, n=30, in increasing variance with 0.0583 value.

4. Conclusion

In this study, 54 design points were created for 10, 30 and 60 observations with 3, 4, 5 variable numbers 3, 5, 7 constant and increasing variance groups for each test statistic. The results of the Monte Carlo simulation run 10,000 times with each design tested are as follows. In cases where the deviation from theType I error rate deviates from the value of 0.05, it is mostly observed in the R test statistic followed by W and T statistics. W and T statistics were given close results in terms of the maximum bias. In the V statistic, the maximum deviation scenarios are less common than the other test statistics. This study suggests that the Pillai Trace statistic works well in the Bernoulli Distribution. Other studies are that found the Pillai Trace test statistic to be reliable in the form of Olson [11], Ito [12], Korin [13], Hopkins and Clay [14] and [8],[9],[10],[11]. details can be generalized as follows. The deviation of the constant variance when the number of groups is g=3 p=3 n=10, the R test statistic is quite high (0.0608). Although this report does not cover why or how this value happened, it is worth noting and perhaps further investigating. This is only a small part of possible research done to evaluate the robustness of MANOVA. The simulations seem to suggest that the larger the sample size, be it group or overall sample size, affects the results. It seems that larger sample sizes seem to help the robustness of the test, such that when n = 60, the deviation of any test statistic is not high. Perhaps because of some multivariate version of the central limit theorem . While the group number is g=5 p=7, the V test statistic is fairly close to 0.05 for all observation values in case of constant variance, and when n=60 this value is exactly 0.05 Small dimension case (g=3,p=3), the larger the sample, more robust V and T becomes against violations in the distribution condition for MANOVA. Large dimension case, growth of g and p size, the increase of variance, negatively affects the V,T,W statistic. Even if equality of variance and small dimension is provided, R is not stable for all groups.

In general, when all the test statistics are examined, the Type I error rates of the Pillai test statistic are the least deviant statistic at nominal α = 0.05 value, as in many studies. However, the theoretical distribution of this statistic is not known precisely. Researchers can produce critical values at different degrees of freedom and Type I error rates with Monte Carlo simulation study and they can submit their recommendation. These are some suggestions to future studies on this topic.

References

[1] Wilks, S.S., 1932. Certain generalizations made in the analysis of variance, Biometrica 24:471-494.

[2] Johnson, R. A.. Wichern D. W., 1982. Applied Multivariate Statistical Analysis. Prentice-Hall, Inc. USA,594s.

[3] Bartlett, M.S., 1954. A Note on the Multiplying Factors for Various chi-square pproximations. Journal of the Royal Statistical Society Series B (Methodological):pp 296-298.

[4] Seber, G. A. F., 1984. Multivariate Observations. John Wiley & sons, Inc., USA,686.

[5] Lawley, D. N., 1939. A generalization of Fisher's z test. Biometrika 30: 467-469.

[6] Hotelling, H., 1931. The generalization of student's ratio. Annals of Mathematical Statistics 2: 360-378.

[7] Pillai, K.C.S., 1955. Some New Test Criteria in Multivariate Analysis. The Annals of Mathematical Statistics 26:117-121.

[8] Davis, A.W.,1980. On The Effects Of Nonnormality On The Likelihood Ratio Criterion Wilks's Moderate Multyvariate. The Journal Of the American Statistical Association 67:419-427.

[9] Davis, A. W., 1982. On The Effects Of The Moderate Multivariate Nonnormality On Roy's Largest Root Tests. The Journal Of the American Statistical Association 77:986-990.

[10] Holloway, L.N., Dunn O.J., 1967. The robestness of Hotelling's T^2. The Journal Of the American Statistical Association 62:124-136.

[11] Olson, C.L., 1974. Copperative Robustness Of Multivariate Analysis Of Six Tests In Variance. The Journal Of the American Association 69 (348): 894-907.

[12] Ito, K., 1969. On The Effect Of Homoscedasticity And Nonnormality Upon Some Multivariate Procedures. In Multivariate Analysis 2:87-120.

[13] Korin, B.P.,1972. Some comment on the Homoscedasticity Criterion M and the multivariate analysis of varia as test T^2, W. and R. Biometrica 59:215-216.

[14] Hopkins, J.W., Clay P.P.F., 1963. Some Bivariate Distribution Of Emprical T^2 And Homoscedasticity Criterion M Under Unecual Variance And Leptokurtosis. The Journal Of the American Statistical Association 58:1048-1053.