



# Asian Journal of Mathematics & Statistics

ISSN 1994-5418

**science**  
alert

**ANSI***net*  
an open access publisher  
<http://ansinet.com>



## Research Article

# Probability of Type I Error and Power of Some Parametric Test: Comparative Approach

<sup>1</sup>Enegelele Dennis, <sup>2</sup>Biu O. Emmanuel and <sup>3</sup>Otaru O.A. Paul

<sup>1</sup>Department of Mathematical and Computing Sciences, Kola Daisi University, Ibadan, Nigeria

<sup>2</sup>Department of Mathematics and Statistics, University of Port Harcourt, Port Harcourt, Nigeria

<sup>3</sup>Department of Mathematics Sciences, Federal University Lokoja, Nigeria

## Abstract

**Background and Objectives:** Assumptions in statistics are mostly violated when testing hypotheses, hence, the use of inappropriate statistical tests results to invalid research conclusions. Most real life data are void of these assumptions resulting to difficulty in analysis using either parametric or non-parametric tests. The objective of the study is to examine the probability of type I error rate and the power of the parametric tests. **Materials and Methods:** To find out probability of type I error and power of some parametric tests such as; Bartlett's, Cochran's, Hartley's and O'Brien test were taken under three conditions; normal and non-normal distributions, equal and unequal sample variances and equal sample size. **Results:** Results showed that all tests were very robust when normality assumption was achieved. But when normality assumption was violated, Hartley's and Cochran tests could not control the type I error applying chi-square and Gamma distribution. For power, Bartlett's, Hartley's and O'Brien tests were most powerful than Cochran test irrespective of the normality assumption and equality of variance. However, Cochran test is more robust than Hartley's test when the distribution is chi-square while, the Hartley test is more robust when the distribution is gamma. **Conclusion:** It is concluded that care should be taken in the choice of an appropriate statistical test when assumption of normality is violated.

**Key words:** Robustness, power, normality, variances, distribution, assumptions

**Citation:** Enegelele Dennis, Biu O. Emmanuel and Otaru O.A. Paul, 2020. Probability of type I error and power of some parametric test: Comparative approach. Asian J. Math. Stat., 13: 7-13.

**Corresponding Author:** Otaru O.A. Paul, Department of Mathematics Sciences, Federal University Lokoja, Nigeria

**Copyright:** © 2020 Enegelele Dennis *et al.* This is an open access article distributed under the terms of the creative commons attribution License, which permits unrestricted use, distribution and reproduction in any medium, provided the original author and source are credited.

**Competing Interest:** The authors have declared that no competing interest exists.

**Data Availability:** All relevant data are within the paper and its supporting information files.

**INTRODUCTION**

The analysis of variance (ANOVA) is the most powerful method for testing hypotheses when the assumptions of normality, homogeneity of variance and independence of errors are achieved<sup>1,2</sup>. Statistical test results are greatly distorted when any of these assumptions are not met, leading to invalid inference<sup>3</sup>. However, test of sample homogeneity of variance are often use in various application of statistical analysis prior to the use of analysis of variance. There are two classes for testing equality of variance, the parametric and non-parametric test, however, this study only considered parametric test. Classic parametric methods are based on certain assumptions so as to produce exact results; the assumptions underlying them (e.g., normality and homoscedasticity) must be fulfilled<sup>4</sup>. Besides normality assumptions, all parametric tests assume random samples, independence within samples and mutual independence between samples<sup>5</sup>. These assumptions are hardly satisfied when analyzing real-life data and thus violated in time series<sup>6</sup>. There are two basic criteria for testing equality of variance, robustness and power.

Parametric tests are significant test which assume certain distribution of the data, interval level of measurement and homogeneity of variances when two or more samples are compared. Most significant tests are Levene test, Bartlett test, Jackknife test, Sharma test, Cochran test, Hartley test and O'Brien test<sup>7,8</sup>. This study considered Bartlett test, Cochran test, Hartley test and O'Brien test for equality of variance in time series. Assuming the expected mean ( $\mu$ ) = 10 and variances are equal and unequal. A hypothesis to test for equality of variances corresponding to m samples is in the form:

$$H_0 = \sigma_1^2 = \sigma_2^2 = \dots = \sigma_m^2 \tag{1}$$

$$H_1 = \sigma_{i1}^2 \neq \sigma_{i2}^2$$

where, the inequality at least for one pair of subscripts  $i_1, i_2$ .

Bartlett's test statistic is designed to test for equality of variances across groups against alternative that variances are unequal for at least two groups. The sampling distribution of the test is approximated by chi-square distribution with m-1 degrees of freedom and random samples m are drawn from an independent normal populations<sup>9</sup>. The test statistic is:

$$B = 2.3026 * G \tag{2}$$

$$G = \frac{(N-m) \log S_p^2 - \sum_{j=1}^m (n_j - 1) \log S_j^2}{1 + \frac{1}{3(m-1)} \sum_{j=1}^m \left( \frac{1}{n_j - 1} \right) - \frac{1}{N-m}} \tag{3}$$

$$S_p^2 = \frac{\sum_{j=1}^m (n_j - 1) S_j^2}{N - k} \tag{4}$$

where, N is the total sample size, k is the number of groups,  $S_p^2$  is the pooled variance and  $S_j^2$  is the sample variance from the jth sample.

In Bartlett's test, sample size of the groups need not be equal, however, sample size should be larger than 5<sup>10</sup>. When comparing statistic for power and robustness, Bartlett's test is most used in several experimental cases<sup>11</sup>. The disadvantage of the test is the assumption that all population follows a normal distribution. The null hypothesis of equal variances is rejected if B is larger than the critical value.

Cochran test is computationally simpler than Bartlett's test, it is used to test homogeneity of variances. The test is also affected by non-normality<sup>12</sup> and it is a good choice for checking homogenous variance if robustness and power against non-normality is needed.

The test is defined as:

$$C = \frac{\text{Largest } S_i^2}{\sum_{i=1}^k S_i^2} \tag{5}$$

where,  $S_i^2$  is the sample variance of the ith group, k is the number of groups and  $df = k - 1$ ; each of the k groups has n-1 degree of freedom.

The hypothesis is rejected if  $C > C_{k,v}$  for a giving  $\alpha, v = n - 1$ , n is number of observations per group. The null hypothesis is rejected concluding that the variances are heterogeneous at 5% level of significance. Cochran's test performed well in power for equal sample<sup>11</sup>. The test is also useful when the variance increases in succession by a constant ratio.

Hartley's  $F_{max}$  utilizes only maximum and minimum variances of groups under test. The test is used in case of equal sample size. The test is defined as:

$$F = \frac{S_{max}^2}{S_{min}^2} \tag{6}$$

Where:

$$S_{\max}^2 = \max(S_1^2, S_2^2, \dots, S_m^2)$$

$$S_{\min}^2 = \min(S_1^2, S_2^2, \dots, S_m^2)$$

F-ratio of the Hartley test is different from F-ratio that is produced in ANOVA. If the variances are similar to each other, then the F-ratio will be close to 1 otherwise the more the variances differ, the larger the F-ratio will be. If the F-ratio is very close to 1, it is safe to conclude that the data probably show equality of variance. If the F-ratio is quite a bit larger than 1, the table of F-max values is used so, as to determine the likelihood of obtaining the F-ratio by chance<sup>6</sup>. If the  $F_{\max} > F_{\max(k, n-1)}$ , variances are heterogeneous.

This test is used to test homogeneity of variance. The null hypothesis states that observation under consideration comes from a population with the same variance<sup>13</sup>. Original series is transformed such that the means of the transformed series reflect the variance of the original series<sup>14</sup>. The transformation is given as:

$$Y_{ij} = \frac{n_i(n_i - 1.5)(X_{ij} - \bar{X}_i)^2 - 0.5SS_i}{(n_i - 1)(n_i - 2)} \quad (7)$$

where, n is the number of observation in ith group,  $\bar{x}_i$  is the mean of ith group,  $X_{ij}$  is the observation at ith row and jth column,  $SS_i$  is the sum of square of group i:

$$\sum_{i=1}^n (X_{ij} - \bar{X}_i)^2$$

If all group has equal sample size:

$$Y_{ij} = \frac{N(N - 1.5)(X_{ij} - \bar{X}_i)^2 - 0.5SS_i}{(N - 1)(N - 2)} \quad (8)$$

where, N is the number of observation per group. When null hypothesis is true, this test statistic has approximately  $F_{k-1, N-k}$  distribution.

In the literature, statistical test have different methods to test data and they also possesses some disadvantages. Bartlett's and O'Brian test are disposed to violation of the normality assumption. Cochran's and Hartley's tests are relatively a good choice for checking equality of variance if robustness against non-normality is needed<sup>15,16</sup>.

The rationale for this study is to examine the probability of type I error rate and the power of the parametric tests considered in this study under three conditions when, (i) The data distributions are normal and non-normal, (ii) The sample size were equal and (iii) The sample variance were equal and unequal.

## MATERIALS AND METHODS

This study was carried out between March-October, 2019 at the Department of Mathematics and Statistics, University of Port Harcourt, Port Harcourt, Nigeria. The study adopted design of Vorapongsathorn *et al.*<sup>8</sup>.

**Robustness and power:** Robustness and power are two criteria used to detect the test for equality of variance under violation of assumption. Robustness is the ability to control type I error when there are small departures from assumption. Also, it is the capability of the test not to wrongly detect non-homogeneous groups when the data is not normally distributed. Therefore, a statistical test is robust if departure of the empirical type I error  $\phi$  from the normal level of significance ( $\alpha$ ) is not greater than the predetermined value<sup>8,15,17</sup>. The study relied on the Cochran limit to test for robustness, which is:

- At 0.01 significance level,  $\phi$  value is between 0.007-0.015
- At 0.05 significance level,  $\phi$  value is between 0.04-0.06

Thus, a statistical test is called robust when it's empirical alpha values lies within the Cochran limit<sup>18</sup>. Error rate could not be controlled by a test when any of the probability of type I error is below or exceed the Cochran limit. The power of a test is the probability of rejecting a null hypothesis when it should be rejected. It is the probability of not committing a type II error in a simulation experiments. This study computes power by subtracting the empirical probability of a type II error  $\hat{\beta}$  from 1. Type II error is an error made by wrongly accepting or failure to reject a false null hypothesis i.e.:

- Power = 1-the empirical probability of a type II error = (1- $\hat{\beta}$ ):

$$\text{Power} = \frac{\text{Number of } H_0 \text{ failed to reject when } H_0 \text{ is true}}{\text{Number of replications 5,000 times}}$$

The maximum total power of a test can have is 1 and the minimum is zero<sup>18</sup>.

**Computation:** The data used in this study were generated by using R programme under different setting and three distributions namely; normal, gamma and chi-square distribution.

**Normal distribution or Gaussian:** If X is a continuous random variable which follows a normal distribution with mean  $\mu$  and variance  $\sigma^2$ , then its p.d.f is given by:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad -\infty < x < \infty, \sigma^2 > 0 \quad (9)$$

**Gamma distribution:** If X is a random variable that follows a gamma distribution and  $\alpha$  and  $\beta$  are the parameters, then:

$$f(x) = \begin{cases} \frac{x^{\alpha-1} e^{-x/\beta}}{\beta^\alpha \Gamma\alpha} & x > 0 \quad \beta > 0 \quad \alpha > 0 \\ 0 & \text{elsewhere} \end{cases} \quad (10)$$

The mean and variance of the Gamma distribution are  $\alpha\beta$  and  $\alpha\beta^2$ , respectively. The values of  $\alpha = 2.5$  and  $\beta = 0.6$ .

**Chi-square distribution:** Suppose  $X \sim \chi^2_k$ , where k is the degree of freedom, then:

$$f(x) = \begin{cases} \frac{1}{2^{k/2} \Gamma(k/2)} e^{-x/2} x^{(k/2)-1}, & 0 < x < \infty, \\ 0 & \text{elsewhere} \end{cases} \quad (11)$$

The mean and variance of the chi-square distribution are k and 2k, respectively<sup>18</sup>.

**Generation of groups of populations and proportion of sample variances:** The population used in this study was generated in three groups with the same distribution as

Normal, gamma and chi-square distribution. This study generated equal sample size of 45, 60 and 75, respectively. The three populations considered in the study have sample variances in each case that were in the ratios 1:1:1 (under  $H_0$ ); 1:1:2 and 1:2:4 (under  $H_1$ )<sup>1</sup>. The theoretical alpha value for testing the equality of variance was defined as 0.01 and 0.05, respectively.

Normal, gamma and chi-square distribution at different combination of sample size for three populations were simulated and the empirical type 1 error and power of tests were investigated. In other to estimate the empirical type 1 error and power estimates, nominal 1 and 5% level was used with 5,000 simulations. The simulated data was used for computing Bartlett's test, Cochran test, Hartley test and O'Brien test. The critical region of the respective tests statistic was compared with the values obtained. The values that rejected the null hypothesis were recorded and those that failed to reject the null hypothesis were also recorded for type II error. The probability of the type II error was subtracted from 1.0 so, as to get the power of the test. The computation of this process was repeated in all conditions.

## RESULTS

Table 1 showed the empirical alpha values of statistical tests controlled by type 1 error for different distribution considered in the study at  $\alpha = 0.01$  and 0.05, respectively, when equality of variance was achieved. When normality condition was achieved, the various test statistic considered in the study were robust as the empirical alpha values of the statistical tests lies within the Cochran limit (0.007-0.015 at 0.01 significant level and 0.04-0.06 at 0.05 significant level, respectively). When normality condition was violated and the distribution was chi-square distribution, Bartlett test and O'Brien test controlled type 1 error at  $\alpha = 0.01$  and 0.05, respectively (Robust) with an average sample size of 60. When the distribution was gamma, none of the tests could control the type 1 error for robustness at  $\alpha = 0.05$ . However, the

Table 1: Type 1 error for equal variance hypothesis of statistical tests and distributions

Nominal level of significance	Sample sizes $N_i$	Normal distribution				Chi-square distribution				Gamma distribution			
		B	C	H	O	B	C	H	O	B	C	H	O
0.01	45,45,45	0.007*	0.010*	0.007*	0.007*	0.006	0.025	0.004	0.005	0.007*	0.025	0.007*	0.044
	60,60,60	0.008*	0.012*	0.008*	0.009*	0.074*	0.026	0.002	0.068	0.008*	0.026	0.007*	0.056
	75,75,75	0.010*	0.012*	0.009*	0.012*	0.010*	0.026	0.003	0.012*	0.015*	0.029	0.009*	0.067
0.05	45,45,45	0.041*	0.052*	0.044*	0.047*	0.003	0.127	0.003	0.045*	0.031	0.124	0.034	0.125
	60,60,60	0.047*	0.058*	0.049*	0.052*	0.044*	0.128	0.001	0.055*	0.033	0.132	0.037	0.223
	75,75,75	0.053*	0.060*	0.058*	0.059*	0.058*	0.128	0.004	0.059*	0.762	0.145	0.047	0.321

B: Bartlett's test, C: Cochran's test, H: Hartley's test, O: O'Brien's test, \*empirical alpha values of the statistical tests

Table 2: Power of statistical tests for unequal variances (1:1:2)

Nominal level of significance	Sample sizes $N_i$	Normal distribution				Chi-square distribution				Gamma distribution			
		B	C	H	O	B	C	H	O	B	C	H	O
0.01	45,45,45	0.479	0.482	0.497	0.478	0.402	0.390	0.410	0.417	0.497	0.377	0.494	0.356
	60,60,60	0.699	0.582	0.866	0.788	0.500	0.489	0.500	0.588	0.697	0.577	0.594	0.367
	75,75,75	0.720	0.682	0.896	0.800	0.730	0.576	0.620	0.590	0.749	0.580	0.795	0.678
0.05	45,45,45	0.797	0.712	0.983	0.854	0.730	0.749	1.000	0.600	0.806	0.585	0.971	0.845
	60,60,60	0.897	0.814	0.981	0.920	0.700	0.844	1.000	0.650	0.887	0.614	0.969	0.865
	75,75,75	0.998	0.911	0.981	0.992	0.890	0.946	1.000	0.720	0.995	0.898	0.974	0.888

Critical 0.00-1.00, B: Bartlett's test, C: Cochran's test, H: Hartley's test, O: O'Brien test

Table 3: Power of statistical tests for unequal variances (1:2:4)

Nominal level of significance	Sample sizes $N_i$	Normal distribution				Chi-square distribution				Gamma distribution			
		B	C	H	O	B	C	H	O	B	C	H	O
0.01	45,45,45	0.960	0.984	0.998	0.967	1.000	0.990	1.000	0.967	0.999	0.981	0.996	0.958
	60,60,60	1.000	0.984	0.998	0.994	1.000	0.989	1.000	0.985	0.999	0.980	0.997	0.989
	75,75,75	1.000	0.984	0.998	0.996	1.000	0.990	1.000	0.978	1.000	0.982	0.997	0.991
0.05	45,45,45	0.998	0.922	0.989	0.986	1.000	0.949	1.000	0.978	0.996	0.903	0.982	0.976
	60,60,60	0.999	0.921	0.990	0.990	1.000	0.947	1.000	1.000	0.997	0.902	0.983	0.987
	75,75,75	0.999	0.921	0.990	1.000	1.000	0.948	1.000	1.000	0.998	0.911	0.987	0.992

Critical 0.00-1.00, B: Bartlett's test, C: Cochran's test, H: Hartley's test, O: O'Brien test

Table 4: Generation of responses for the three distributions, equal sample size and unequal sample variances

Distribution/ level of significance	Sample size $n_i$	Under $H_0$		Under $H_1$		
		Under $H_0$ , the ratio of variance = 1:1:1	Under $H_0$ , the ratio of variance = 1:1:1	Ratio of variance = 1:1:2	Ratio of variance = 1:2:4	
Normal $\alpha = 0.01$	45, 45, 45 ...	$X_i \sim N(\mu, \sigma^2), \forall i$	$X_i \sim N(\mu, \sigma^2)$	$X_1 \sim N(\mu, \sigma^2)$	$X_1 \sim N(\mu, \sigma^2)$	
	60, 60, 60 ...	$\mu_1 = \mu \forall i$	$X_2 \sim N(\mu, \sigma^2)$	$X_2 \sim N(\mu, \sigma^2)$	$X_2 \sim N(\mu, 2\sigma^2)$	
	75, 75, 75 ...	$\sigma_1^2 = \sigma^2 \forall i$	$X_3 \sim N(\mu, 2\sigma^2)$	$X_3 \sim N(\mu, 2\sigma^2)$	$X_3 \sim N(\mu, 4\sigma^2)$	
$\alpha = 0.05$	$\forall i (i = 1, 2, 3)$	$H_0 = \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma_m^2 = 1:1:1$	$H_1 = \sigma_1^2: \sigma_2^2: \sigma_3^2 = \sigma_1^2: \sigma_2^2: 2\sigma_3^2 = 1:1:2$	$H_1 = \sigma_1^2: \sigma_2^2: \sigma_3^2 = \sigma_1^2: 2\sigma_2^2: 4\sigma_3^2 = 1:2:4$	$H_1 = \sigma_1^2: \sigma_2^2: \sigma_3^2 = \sigma_1^2: 2\sigma_2^2: 4\sigma_3^2 = 1:2:4$	
	Gamma $\alpha = 0.01$	45, 45, 45 ...	$X_i \sim G(\alpha, \beta) \forall i$	$X_1 \sim G(\alpha, \beta)$	$X_1 \sim G(\alpha, \beta)$	$X_1 \sim G(\alpha, \beta)$
		60, 60, 60 ...	$\mu_1 = \alpha\beta \forall i$	$X_2 \sim G(\alpha, \beta)$	$X_2 \sim G(\alpha, \beta)$	$X_2 \sim G(\alpha/2, 2\beta)$
75, 75, 75 ...		$\sigma_1^2 = \alpha\beta^2 \forall i$	$X_3 \sim G(\alpha/2, 2\beta)$	$X_3 \sim G(\alpha/2, 2\beta)$	$X_3 \sim G(\alpha/4, 4\beta)$	
$\alpha = 0.05$	$\forall i (i = 1, 2, 3)$	$H_0 = \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \alpha\beta^2 = 1:1:1$	$H_1 = \sigma_1^2: \sigma_2^2: \sigma_3^2 = \alpha\beta^2: \alpha\beta^2: 2\alpha\beta^2 = 1:1:2$	$H_1 = \sigma_1^2: \sigma_2^2: \sigma_3^2 = \alpha\beta^2: 2\alpha\beta^2: 4\alpha\beta^2 = 1:2:4$	$H_1 = \sigma_1^2: \sigma_2^2: \sigma_3^2 = \alpha\beta^2: 2\alpha\beta^2: 4\alpha\beta^2 = 1:2:4$	
	Chi-square $\alpha = 0.01$	45, 45, 45 ...	$X_i \sim \chi^2(\alpha, \beta)$	$X_1 \sim \chi^2(n), \mu_1 = n, \sigma_1^2 = 2n$	$X_1 \sim \chi^2(n), \mu_1 = n, \sigma_1^2 = 2n$	$X_1 \sim \chi^2(n), \mu_1 = n, \sigma_1^2 = 2n$
		60, 60, 60 ...	$\mu_1 = n$	$X_2 \sim \chi^2(n), \mu_2 = n, \sigma_2^2 = 2n$	$X_2 \sim \chi^2(2n), \mu_2 = 2n, \sigma_2^2 = 4n$	$X_2 \sim \chi^2(2n), \mu_2 = 2n, \sigma_2^2 = 4n$
75, 75, 75 ...		$\sigma_1^2 = 2n \forall i$	$X_3 \sim \chi^2(2n), \mu_3 = 2n, \sigma_3^2 = 4n$	$X_3 \sim \chi^2(4n), \mu_3 = 4n, \sigma_3^2 = 8n$	$X_3 \sim \chi^2(4n), \mu_3 = 4n, \sigma_3^2 = 8n$	
$\alpha = 0.05$	$\forall i (i = 1, 2, 3)$	$H_0 = \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = 2n = 1:1:1$	$H_1 = \sigma_1^2: \sigma_2^2: \sigma_3^2 = 2n:2n:4n = 1:1:2$	$H_1 = \sigma_1^2: \sigma_2^2: \sigma_3^2 = 2n:4n:8n = 1:2:4$	$H_1 = \sigma_1^2: \sigma_2^2: \sigma_3^2 = 2n:4n:8n = 1:2:4$	

sample size did not affect robustness of Bartlett and Hartley's tests as their empirical alpha values of the statistical tests lies within the Cochran limit.

Furthermore, Table 2 showed the power of statistical tests for different distribution considered in the study at  $\alpha = 0.01$  and 0.05, respectively when variances are unequal (1:1:2). When the sample size was 45 for all the distribution, the power of all the tests were below 0.5 at  $\alpha = 0.01$ . This implies that the tests could identify a faulty null hypothesis. However, when the sample size increased up to 70, the power of the tests for the different distribution improved significantly.

Table 3 showed the power of statistical tests for different distribution considered in the study at  $\alpha = 0.01$  and 0.05, respectively when variances are unequal (1:2:4). The results showed that Bartlett's and Hartley's tests had the maximum power of 1.0 for the different distributions irrespective of the sample size.

Finally, Table 4 showed that how the populations used in the study were generated in three different groups with the same distribution as normal, gamma and chi-square. The sample variances in each case for the population considered in the study were in the ratios 1:1:1 (under  $H_0$ ); 1:1:2 and 1:2:4 (under  $H_1$ ).

## DISCUSSION

Bartlett's, Cochran, Hartley's and O'Brien tests were very robust when normality assumption was achieved at  $\alpha = 0.01$  and  $0.05$  irrespective of the sample size as the empirical alpha value lies with the Cochran limit. The result is in accordance with the findings of Hatchavanich<sup>17</sup>. However, when normality assumption was violated and the distribution was chi-square, Bartlett's and O'Brien tests could control the type 1 error rate (robust) when the sample size was 60 and 70, respectively at  $\alpha = 0.01$  and  $0.05$ . The result does not support the findings of Sharma and Kibria<sup>2</sup>. According to their findings, Bartlett's test was non-robust when normality assumption was violated and the sample size was 30. In this study, the empirical alpha value for Bartlett's test when the sample size was 60 and 70 and the distribution was chi-square (normality violated) are 0.074 and 0.010 at  $\alpha = 0.01$ . The empirical alpha value for O'Brien test when the sample size was 70 is 0.012 (Table 1). This result is in accordance with the finding of Lee *et al.*<sup>5</sup> when it was reported that the O'Brien test performed very robust as it could control the type I error rate across all population distributions, except for small sample size. Bartlett's and Hartley's test were the only tests that were robust when the population was gamma distribution at  $\alpha = 0.01$ . When the nominal level of significance was 0.05 and the distribution was gamma, none of the statistical test considered in the study was robust as their respective empirical alpha value lies outside the Cochran limit of 0.007-0.015 for  $\alpha = 0.01$  and 0.004-0.006 for  $\alpha = 0.05$ . The power of each statistical test for all distributions when variance ratio was 1:1:2 was less than 0.5 at  $\alpha = 0.01$ . This indicated that the tests could detect faulty null hypotheses. Bartlett's test maintained good statistical power of 0.998 and 0.995, respectively when normality assumption was achieved and also when the population was gamma distribution with a sample size of 70 at  $\alpha = 0.05$ . This result is in accordance with the finding of Hatchavanich<sup>17</sup>. This study also revealed that Hartley's test had the maximum power of 1.00 irrespective of sample size and variance ratio 1:1:2 when normality assumption was violated, but the distribution was chi-square at  $\alpha = 0.05$  (Table 2).

Finally, the power of each statistical test for all distributions considered in the study when variance ratio was 1:2:4 revealed that, even when normality assumption was violated, all the tests had power greater than 0.8 with an average sample size of 45 at  $\alpha = 0.01$  and  $0.05$ , respectively (Table 3).

In this study, four parametric (Bartlett's, Cochran's, Hartley's and O'Brien) tests for constant variance were considered and compared with respect to power and

robustness. When normality assumption was achieved, all statistical tests were robust at  $\alpha = 0.01$  and  $0.05$ , respectively. This result is in accordance with the findings of Vorapongsathorn *et al.*<sup>8</sup>. However, when normality condition was violated and the population was chi-square, the O'Brien test statistic was more robust at  $\alpha = 0.05$  than the Cochran test statistic considered in the literature when the average sample size is 45. Thus, this study recommends that care should be taken in choosing a statistical test when assumption of normality is violated.

## CONCLUSION

This study showed that when distribution is normal and variance is equal, empirical type I error of Bartlett's, Cochran, Hartley's and O'Brien tests satisfy Cochran limits, thus, Cochran test is more robust than Hartley's test when the distribution is chi-square and Hartley's test is more robust when the distribution is gamma. All tests were very robust when normality assumption was achieved, but when violated, Hartley's and Cochran tests could not control the type I error for chi-square and gamma distribution, respectively. For power, Bartlett's, Hartley's and O'Brien tests were most powerful than Cochran test irrespective of the normality assumption and equality of variance.

## SIGNIFICANCE STATEMENT

This study discovers the most powerful Bartlett's, Hartley's and O'Brien tests that can be beneficial for data which follows chi-square and gamma distribution. This study will help the researcher to uncover the critical areas of assumption violation applying other distributions and statistical tests for robustness and power.

## REFERENCES

1. Ruscio, J. and B. Roche, 2012. Variance heterogeneity in published psychological research. *Methodology*, 8: 1-11.
2. Sharma, D. and B.M.G. Kibria, 2013. On some test statistics for testing homogeneity of variances: A comparative study. *J. Stat. Comput. Simul.*, 83: 1944-1963.
3. Kim, Y.J. and R.A. Cribbie, 2018. ANOVA and the variance homogeneity assumption: Exploring a better gatekeeper. *Br. J. Math. Stat. Psychol.*, 71: 1-12.
4. Fox, J., 2008. *Applied Regression Analysis and Generalized Linear Models*. 2nd Edn., Sage, Thousand Oaks, C.A.
5. Lee, H.B., G.S. Katz and A.F. Restori, 2010. A monte carlo study of seven homogeneity of variance tests. *J. Math. Stat.*, 6: 359-366.

6. Zimmerman, D.W., 2004. A note on preliminary tests of equality of variances. *Br. J. Math. Stat. Psychol.*, 57: 173-181.
7. Filliben, J.J. and A. Heckert, 2000. Levene's test for equality of variances. Online. <http://www.nist.gov/itl/div898/handbook/eda/section3/eda358.htm>
8. Vorapongsathorn, T., S. Taejaroenkul and C. Viwatwongkasem, 2004. A comparison of type I error and power of Bartlett's test, Levene's test and Cochran's test under violation of assumptions. *Songklanakarin J. Sci. Technol.*, 26: 537-547.
9. Montgomery, D.C., 1997. *Design and Analysis of Experiments*. 4th Edn., John Wiley and Sons, New York.
10. Li, X., W. Qiu, J. Morrow, D.L. DeMeo, S.T. Weiss, Y. Fu and X. Wang, 2015. A comparative study of tests for homogeneity of variances with application to DNA methylation data. *PloS One*, Vol. 10, No. 12 10.1371/journal.pone.0145295.
11. Boos, D.D. and C. Brownie, 2004. Comparing variances and other measures of dispersion. *Stat. Sci.*, 19: 571-578.
12. Lemeshko, B.Y. and S.B. Lemeshko, 2008. Power and robustness of criteria used to verify the homogeneity of means. *Meas. Tech.*, 51: 950-959.
13. Lemeshko, B.Y., S.B. Lemeshko and A.A. Gorbunova, 2010. Application and power of criteria for testing the homogeneity of variances. Part I. Parametric criteria. *Meas. Tech.*, 53: 237-246.
14. Mu, Z., 2006. Comparing the statistical tests for homogeneity of variances. M.Sc. Thesis, Department of Mathematics, East Tennessee State University, USA.
15. Wilcox, R.R., 2010. *Fundamentals of Modern Statistical Methods: Substantially Improving Power and Accuracy*. 2nd Edn., Springer-Verlag, New York.
16. Field, A.P., J.N.V. Miles and Z.C. Field, 2012. *Discovering Statistics Using R: And Sex and Drugs and Rock 'n' Roll*. Sage, London.
17. Hatchavanich, D., 2014. A comparison of type I error and power of Bartlett's test, Levene's test and O'Brien's test for homogeneity of variance tests. *S. Asian J. Sci.*, 3: 181-194.
18. Lemeshko, B.Y. and V.M. Ponomarenko, 2006. Distributions of the statistics used for testing hypotheses of equal variances for non-normal law error. *Nauch. Vestnik NGTU*, Vol. 23, No. 2.