

# A Comparison of Efficiency for Homogeneity of Variance Tests under Log-normal Distribution

Kunlanan Sritan and Bumrungrsak Phuenaree\*

Department of Mathematics, Faculty of Science,  
Burapha University, 20131, Chonburi, Thailand.

\*Corresponding author's email: bumrungrsak [AT] buu.ac.th

**ABSTRACT**— *In this paper, we compare five homogeneity of variance tests which are Bartlett's test, Levene's test, Cochran's test, O'Brien's test and Jackknife test. Considering their ability to control probability of type I error and the power of the test, when groups of population are distributed in log-normal distribution. The equal sample sizes are defined as 10, 15, 30 and 50 at the significance is 0.05. The results show that the Levene's test become the best test for the high skewed distribution. For the lowed skew distribution, the Cochran's test is the best test when a variance of group is different to the others. Moreover, Bartlett's test provides the highest power when variances of all populations are different.*

**Keywords**— Homogeneity of variance test, Log-normal distribution, Type I error, Power of the test

## 1. INTRODUCTION

The homogeneity of variance is the important assumption of the analysis of variance (ANOVA). There are many test statistics used to test the variance equality. Some classical tests are very sensitive to departure from normality; Bartlett's test and Cochran's test (Snedecor and Cochran, 1983), whereas some test can be used when data are non-normal; Levene's test (Levene, 1960). The well-known alternative test statistics for checking the equality of variance under different situations are O'Brien's test and Jackknife test (O'Brien, 1981; Miller, 1968). However, in many statistical literatures reported that there is no test which is best test for all distributions. They have some different weak points in difference situation.

Many climate variables, for example water vapor and rainfall can be represented with the skew distribution. One of the most important the probability skew distributions to describes the characteristic of some climate variable is log-normal distribution (Foster, Bevis and Raymond, 2006; Sharma and Singh, 2010). The purpose of this work is to compare the efficiency of Bartlett's test, Levene's test, Cochran's test, O'Brien's test and Jackknife test in term of the robustness and the power of the test when the data are lognormal distribution.

## 2. METHODS

The well-known statistical test of the homogeneity of variances tests; Bartlett's test, Levene's test, Cochran's test, O'Brien's test and Jackknife test are considered in this work. In this section, we consider the homogeneity of variance of  $k$  populations given sample  $X_{ij} : i = 1, 2, \dots, k; j = 1, 2, \dots, n_i$ , where  $\bar{X}_i$  and  $S_i^2$  are mean and variance of the  $i$ th group, respectively.

### 2.1 Bartlett's test

Bartlett's test was proposed by Bartlett (1937) for testing the homogeneity of variances. It can be computed as

$$B = \frac{(n-k) \ln(S_p^2) - \sum_{i=1}^k (n_i - 1) \ln(S_i^2)}{1 + \left( \frac{1}{3(k-1)} \right) \left( \left( \sum_{i=1}^k \frac{1}{(n_i - 1)} \right) - \frac{1}{(n-k)} \right)}$$

$$\text{, where } S_i^2 = \frac{\sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2}{n_i - 1}, S_p^2 = \frac{\sum_{i=1}^k (n_i - 1) S_i^2}{n - k}; i = 1, 2, 3, \dots, k, n = \sum_{i=1}^k n_i.$$

## 2.2 Levene's test

This test was proposed by Levene (1960) for testing the homogeneity of variances. It can be defined as

$$L = \frac{(n - k) \sum_{i=1}^k n_i (\bar{Z}_i - \bar{Z})^2}{(k - 1) \sum_{i=1}^k \sum_{j=1}^{n_i} (Z_{ij} - \bar{Z}_i)^2}$$

$$\text{, where } Z_{ij} = (X_{ij} - \bar{X}_i)^2, \bar{Z}_i = \frac{\sum_{j=1}^{n_i} Z_{ij}}{n_i}, \bar{Z} = \frac{\sum_{i=1}^k n_i \bar{Z}_i}{n}.$$

## 2.3 Cochran's test

This test uses to examine the homogeneity of variances. Cochran's test statistic can be calculated as

$$C = \frac{S_{\max}^2}{\sum_{i=1}^k S_i^2}$$

,where  $S_{\max}^2$  is the largest sample variances and  $\sum_{i=1}^k S_i^2$  is the sum of all variances.

## 2.4 O'Brien's test

O'Brien's test was proposed by O'Brien (1981). This test modifies the concept of analysis of variance F-test based on  $V_{ij}$ .

$$V_{ij} = \frac{(n_i - 1.5)n_i (X_{ij} - \bar{X}_i)^2 - 0.5 S_i^2 (n_i - 1)}{(n_i - 1)(n_i - 2)}$$

$$\text{, where } S_i^2 = \frac{\sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2}{n_i - 1} \quad \text{and} \quad \bar{X}_i = \frac{\sum_{j=1}^{n_i} X_{ij}}{n_i}.$$

The O'Brien test statistic will be the F-value computed on applying the usual analysis of variance procedure on  $V_{ij}$ .

## 2.5 Jackknife test

The Jackknife test is defined as following:

$$J_k = \frac{\sum_{i=1}^k n_i (\bar{u}_i - \bar{u}_{..})^2 / (k - 1)}{\sum_{i=1}^k \sum_{j=1}^{n_i} (u_{ij} - \bar{u}_i)^2 / (n - k)}$$

,where  $\bar{X}_{i(j)} = \frac{\sum_{i \neq j} X_{ij}}{n_i - 1}$  is the mean of  $i$ th treatment group excluded  $j$ th sample.

$S_{i(j)}^2 = \frac{\sum_{i \neq j} (X_{ij} - \bar{X}_{i(j)})^2}{n_i - 2}$  is the variance of  $i$ th treatment group excluded  $j$ th sample.

$$u_{ij} = n_i \log S_i^2 - (n_i - 1)S_{i(j)}^2, \bar{u}_i = \frac{\sum_{j=1}^{n_i} u_{ij}}{n_i}, \bar{u}_{..} = \frac{\sum_{i=1}^k \sum_{j=1}^{n_i} u_{ij}}{n} \text{ for } j = 1, 2, 3, \dots, n_i, i = 1, 2, 3, \dots, k.$$

### 2.6 Log-normal distribution

The probability distribution is

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left[-\frac{(\ln x - \mu)^2}{2\sigma^2}\right]; 0 \leq x < \infty, -\infty < \mu < \infty, \sigma > 0$$

,where  $\mu$  is the location parameter and  $\sigma$  is the shape parameter.

### 2.7 Type I error rate and power of the test

The type I error rate is the probability of false rejection. This value should be close to the significant level ( $\alpha$ ). However, the Bradley limit ( $0.5\alpha, 1.5\alpha$ ) is used in order to investigate that type I error in control.

The power of the test is the probability of rejecting false hypotheses. The test will be called the best method when it shows the highest power. However, we consider only the test that can control type I error.

## 3. SIMULATION STUDY

This study uses Monte Carlo simulation to evaluate the performance of all tests; Bartlett's test (B), Levene's test (L), Cochran's test (C), O'Brien's test (OB) and Jackknife test (J). The lognormal distribution data are generated with the parameters shown in Table 1-2. There are three skewness values; 6.185, 1.750, 0.378 and balance-sample sizes;  $n = 10, 15, 30$  and  $50$ . The number of population groups are set at 3 and 4. The variance ratios are 1:1:1, 1:1:4, 1:2:4, 1:1:1:4 and 1:2:4:8. The significant level is 0.05.

This research uses repeating 5,000 times for each case. The results for type I error rates and the power of the test are shown in Table 3-7.

Table 1: The parameters of lognormal distribution for three groups

| Skewness | Variance ratios | Population 1 |          | Population 2 |          | Population 3 |          |
|----------|-----------------|--------------|----------|--------------|----------|--------------|----------|
|          |                 | $\mu$        | $\sigma$ | $\mu$        | $\sigma$ | $\mu$        | $\sigma$ |
| 6.185    | 1:1:1           | -0.771       | 1.0      | -0.771       | 1.0      | -0.771       | 1.0      |
|          | 1:1:4           | -0.771       | 1.0      | -0.771       | 1.0      | -0.078       | 1.0      |
|          | 1:2:4           | -0.771       | 1.0      | -0.424       | 1.0      | -0.078       | 1.0      |
| 1.750    | 1:1:1           | 0.504        | 0.5      | 0.504        | 0.5      | 0.504        | 0.5      |
|          | 1:1:4           | 0.504        | 0.5      | 0.504        | 0.5      | 1.197        | 0.5      |
|          | 1:2:4           | 0.504        | 0.5      | 0.851        | 0.5      | 1.197        | 0.5      |
| 0.378    | 1:1:1           | 2.068        | 0.125    | 2.068        | 0.125    | 2.068        | 0.125    |
|          | 1:1:4           | 2.068        | 0.125    | 2.068        | 0.125    | 2.761        | 0.125    |
|          | 1:2:4           | 2.068        | 0.125    | 2.414        | 0.125    | 2.761        | 0.125    |

Table 2: The parameters of lognormal distribution for four groups

| Skewness | Variance ratios | Population 1 |          | Population 2 |          | Population 3 |          | Population 4 |          |
|----------|-----------------|--------------|----------|--------------|----------|--------------|----------|--------------|----------|
|          |                 | $\mu$        | $\sigma$ | $\mu$        | $\sigma$ | $\mu$        | $\sigma$ | $\mu$        | $\sigma$ |
| 6.185    | 1:1:1:1         | -0.771       | 1.0      | -0.771       | 2.0      | -0.771       | 1.0      | -0.771       | 1.0      |
|          | 1:1:1:4         | -0.771       | 1.0      | -0.771       | 2.0      | -0.771       | 1.0      | -0.078       | 1.0      |
|          | 1:2:4:8         | -0.771       | 1.0      | -0.424       | 2.0      | -0.078       | 1.0      | 0.269        | 1.0      |
| 1.750    | 1:1:1:1         | 0.504        | 0.5      | 0.504        | 0.5      | 0.504        | 0.5      | 0.504        | 0.5      |
|          | 1:1:1:4         | 0.504        | 0.5      | 0.504        | 0.5      | 0.504        | 0.5      | 1.197        | 0.5      |
|          | 1:2:4:8         | 0.504        | 0.5      | 0.851        | 0.5      | 1.197        | 0.5      | 1.544        | 0.5      |
| 0.378    | 1:1:1:1         | 2.068        | 0.125    | 2.068        | 0.125    | 2.068        | 0.125    | 2.068        | 0.125    |
|          | 1:1:1:4         | 2.068        | 0.125    | 2.068        | 0.125    | 2.068        | 0.125    | 2.761        | 0.125    |
|          | 1:2:4:8         | 2.068        | 0.125    | 2.414        | 0.125    | 2.761        | 0.125    | 3.107        | 0.125    |

#### 4. RESULTS

Table 3: Type I error rates for three groups

| Skewness | Sample sizes | Test           |                |                |                |                |
|----------|--------------|----------------|----------------|----------------|----------------|----------------|
|          |              | B              | L              | C              | OB             | J              |
| 6.185    | 10           | 0.5858         | <b>0.0434*</b> | 0.5010         | <b>0.0490*</b> | 0.1544         |
|          | 15           | 0.6426         | <b>0.0354*</b> | 0.5618         | <b>0.0410*</b> | 0.1508         |
|          | 30           | 0.7292         | <b>0.0396*</b> | 0.6576         | <b>0.0314*</b> | 0.1436         |
|          | 50           | 0.7654         | <b>0.0406*</b> | 0.7064         | <b>0.0310*</b> | 0.1350         |
| 1.750    | 10           | 0.2398         | <b>0.0354*</b> | 0.2356         | <b>0.043*</b>  | 0.0850         |
|          | 15           | 0.2894         | <b>0.0356*</b> | 0.2670         | <b>0.0458*</b> | 0.0954         |
|          | 30           | 0.3476         | <b>0.0412*</b> | 0.3102         | <b>0.0428*</b> | 0.0970         |
|          | 50           | 0.3530         | <b>0.0422*</b> | 0.3168         | <b>0.0432*</b> | 0.0860         |
| 0.378    | 10           | <b>0.0624*</b> | <b>0.0340*</b> | <b>0.0674*</b> | <b>0.0324*</b> | <b>0.0598*</b> |
|          | 15           | <b>0.0618*</b> | <b>0.0256*</b> | <b>0.0640*</b> | <b>0.0370*</b> | <b>0.0534*</b> |
|          | 30           | <b>0.0706*</b> | <b>0.0402*</b> | <b>0.0744*</b> | <b>0.0448*</b> | <b>0.0574*</b> |
|          | 50           | <b>0.0718*</b> | <b>0.0452*</b> | <b>0.0674*</b> | <b>0.0478*</b> | <b>0.0548*</b> |

\* Type I error rates fell in Bradley's limit.

Table 4: Type I error rates for four groups

| Skewness | Sample sizes | Test           |                |                |                |                |
|----------|--------------|----------------|----------------|----------------|----------------|----------------|
|          |              | B              | L              | C              | OB             | J              |
| 6.185    | 10           | 0.6940         | <b>0.0464*</b> | 0.5742         | <b>0.0594*</b> | 0.1734         |
|          | 15           | 0.7788         | <b>0.0302*</b> | 0.6660         | <b>0.0384*</b> | 0.1756         |
|          | 30           | 0.8332         | <b>0.0392*</b> | 0.7416         | <b>0.0320*</b> | 0.1522         |
|          | 50           | 0.8784         | <b>0.0470*</b> | 0.8092         | <b>0.0360*</b> | 0.1498         |
| 1.750    | 10           | 0.2978         | <b>0.0398*</b> | 0.2704         | <b>0.0500*</b> | 0.1078         |
|          | 15           | 0.3480         | <b>0.0318*</b> | 0.3176         | <b>0.0504*</b> | 0.1078         |
|          | 30           | 0.4186         | <b>0.0410*</b> | 0.3604         | <b>0.0432*</b> | 0.0986         |
|          | 50           | 0.4580         | <b>0.0490*</b> | 0.3922         | <b>0.0406*</b> | 0.0920         |
| 0.378    | 10           | <b>0.0594*</b> | <b>0.0294*</b> | <b>0.0662*</b> | <b>0.0386*</b> | <b>0.0480*</b> |
|          | 15           | <b>0.0686*</b> | <b>0.0324*</b> | <b>0.0700*</b> | <b>0.0424*</b> | <b>0.0580*</b> |
|          | 30           | <b>0.0636*</b> | <b>0.0418*</b> | <b>0.0642*</b> | <b>0.0440*</b> | <b>0.0560*</b> |
|          | 50           | <b>0.0690*</b> | <b>0.0426*</b> | <b>0.0702*</b> | <b>0.0470*</b> | <b>0.0518*</b> |

\* Type I error rates fell in Bradley's limit.

Table 5: Power of the test for skewness value 6.185

| Variance ratios | Sample sizes | Test |                |   |        |   |
|-----------------|--------------|------|----------------|---|--------|---|
|                 |              | B    | L              | C | OB     | J |
| 1:1:4           | 10           | -    | <b>0.1278*</b> | - | 0.1080 | - |
|                 | 15           | -    | <b>0.1690*</b> | - | 0.1234 | - |
|                 | 30           | -    | <b>0.3694*</b> | - | 0.1858 | - |
|                 | 50           | -    | <b>0.5888*</b> | - | 0.2762 | - |
| 1:2:4           | 10           | -    | <b>0.0898*</b> | - | 0.0834 | - |
|                 | 15           | -    | <b>0.1114*</b> | - | 0.0840 | - |
|                 | 30           | -    | <b>0.2438*</b> | - | 0.1168 | - |
|                 | 50           | -    | <b>0.4086*</b> | - | 0.1686 | - |
| 1:1:1:4         | 10           | -    | <b>0.1356*</b> | - | 0.1226 | - |
|                 | 15           | -    | <b>0.1794*</b> | - | 0.1362 | - |
|                 | 30           | -    | <b>0.3758*</b> | - | 0.1870 | - |
|                 | 50           | -    | <b>0.5764*</b> | - | 0.2620 | - |
| 1:2:4:8         | 10           | -    | <b>0.1408*</b> | - | 0.1210 | - |
|                 | 15           | -    | <b>0.2256*</b> | - | 0.1360 | - |
|                 | 30           | -    | <b>0.4722*</b> | - | 0.1942 | - |
|                 | 50           | -    | <b>0.7354*</b> | - | 0.2812 | - |

\*The highest power of the test

Table 6: Power of the test for skewness value 1.750

| Variance ratios | Sample sizes | Test |                |   |        |   |
|-----------------|--------------|------|----------------|---|--------|---|
|                 |              | B    | L              | C | OB     | J |
| 1:1:4           | 10           | -    | <b>0.2518*</b> | - | 0.2352 | - |
|                 | 15           | -    | <b>0.3906*</b> | - | 0.3258 | - |
|                 | 30           | -    | <b>0.7678*</b> | - | 0.5950 | - |
|                 | 50           | -    | <b>0.9470*</b> | - | 0.8018 | - |
| 1:2:4           | 10           | -    | <b>0.1574*</b> | - | 0.1446 | - |
|                 | 15           | -    | <b>0.2534*</b> | - | 0.2108 | - |
|                 | 30           | -    | <b>0.5862*</b> | - | 0.4064 | - |
|                 | 50           | -    | <b>0.8424*</b> | - | 0.6092 | - |
| 1:1:1:4         | 10           | -    | <b>0.2566*</b> | - | 0.2458 | - |
|                 | 15           | -    | <b>0.4014*</b> | - | 0.3594 | - |
|                 | 30           | -    | <b>0.7848*</b> | - | 0.6362 | - |
|                 | 50           | -    | <b>0.9534*</b> | - | 0.8400 | - |
| 1:2:4:8         | 10           | -    | <b>0.3108*</b> | - | 0.2630 | - |
|                 | 15           | -    | <b>0.5216*</b> | - | 0.3826 | - |
|                 | 30           | -    | <b>0.9080*</b> | - | 0.6626 | - |
|                 | 50           | -    | <b>0.9950*</b> | - | 0.8702 | - |

\*The highest power of the test

In Table 3-4, the type I error rates of Levene's test and O'Brien's test fall in Bradley limit, i.e., these tests can control type I error for all situations. However, all tests can control type I error when the skewness is small. Table 5-7 show the power of the test that can control type I error. Levene's test has the higher power than O'Brien's test for high skewness values regardless of the variance ratios and sample sizes. For fairly symmetric data, the Cochran's test is the best test when a variance of group is different to the others. Moreover, Bartlett's test had the highest power when variances of all populations are different. However, all tests perform well when the variance ratio is 1:2:4:8 and sample size is 50.

Table 7: Power of the test for skewness value 0.378

| Variance ratios | Sample sizes | Test           |                |                |                |                |
|-----------------|--------------|----------------|----------------|----------------|----------------|----------------|
|                 |              | B              | L              | C              | OB             | J              |
| 1:1:4           | 10           | 0.5624         | 0.3578         | <b>0.6160*</b> | 0.3570         | 0.4048         |
|                 | 15           | 0.7560         | 0.5788         | <b>0.8026*</b> | 0.6176         | 0.6226         |
|                 | 30           | 0.9686         | 0.9294         | <b>0.9770*</b> | 0.9382         | 0.9256         |
|                 | 50           | 0.9982         | 0.9954         | <b>0.9990*</b> | 0.9956         | 0.9952         |
| 1:2:4           | 10           | <b>0.4076*</b> | 0.2232         | 0.3620         | 0.215          | 0.3074         |
|                 | 15           | <b>0.5938*</b> | 0.3872         | 0.5242         | 0.3958         | 0.5056         |
|                 | 30           | <b>0.9042*</b> | 0.8022         | 0.8128         | 0.7964         | 0.8430         |
|                 | 50           | <b>0.9912*</b> | 0.9716         | 0.9622         | 0.9712         | 0.9768         |
| 1:1:1:4         | 10           | 0.5764         | 0.3836         | <b>0.6544*</b> | 0.4234         | 0.3776         |
|                 | 15           | 0.7688         | 0.5908         | <b>0.8222*</b> | 0.6562         | 0.6004         |
|                 | 30           | 0.9734         | 0.9372         | <b>0.9832*</b> | 0.9604         | 0.9300         |
|                 | 50           | 0.9992         | 0.9970         | <b>0.9994*</b> | 0.9978         | 0.9948         |
| 1:2:4:8         | 10           | <b>0.7326*</b> | 0.4562         | 0.5804         | 0.4116         | 0.5670         |
|                 | 15           | <b>0.9338*</b> | 0.7544         | 0.7758         | 0.7098         | 0.8374         |
|                 | 30           | <b>0.9980*</b> | 0.9914         | 0.9782         | 0.9820         | 0.9924         |
|                 | 50           | <b>1.0000*</b> | <b>1.0000*</b> | 0.9998         | <b>1.0000*</b> | <b>1.0000*</b> |

\*The highest power of the test

## 5. CONCLUSION

From the results, it can be seen that the Levene's test becomes the best test for the high skewed distribution. For the lowed skew distribution, the Cochran's test is the best test when a variance of group is different to the others. Moreover, Bartlett's test has the highest power when variances of all populations are different.

## 6. ACKNOWLEDGEMENT

This study was supported by the Faculty of Science, Burapha University, Thailand

## 7. REFERENCES

- [1] Foster, J., Bevis, M. & Raymond, W., "Precipitable water and the lognormal distribution", Journal of Geophysical Research Atmospheres., vol.111, no.15, 2006
- [2] Levene, H. "Robust Tests for Equality of Variances", In Contributions to Probability and Statistics. California: Stanford University Press. Pp.278-292, 1960
- [3] Miller, R.G. Jr. "Jackknifing variances", Annals of Mathematical Statistics, vol.39, pp.567-582, 1968
- [4] O'Brien, R. G. "A simple test for variance effects in experimental designs". Psychological Bulletin, vol.89, no.3, pp.570-574, 1981.
- [5] Sharma, M.A. & Singh, J.B. "Use of Probability Distribution in Rainfall Analysis", New York Science Journal, vol 3, no.9, 2010
- [6] Snedecor, G.W. & Cochran, W.G. Statistical Methods ( 6<sup>th</sup> ed.). Oxford and IBH, New Delhi., 1983