Development of the Robust Test for Testing the Homogeneity of Variances and Its Applications

UNCHALEE TONGGUMNEAD[1,*], NIKORN SAENGNGAM[2]
[1]Department of Mathematics and Computer Science,
Rajamangala University of Technology Thanyaburi,
39 Moo 1, Klong 6, Khlong Luang, Pathum Thani 12110,
THAILAND

[2]Department of Technical Education
Rajamangala University of Technology Thanyaburi
39 Moo 1, Klong 6, Khlong Luang, Pathum Thani 12110
THAILAND

*Corresponding Author

Abstract: - The objective of this study is to develop a robust Levene's test for testing the homogeneity of the variances of k datasets ($k = 3$) by reformulating the test in the form of a two-stage regression framework in the absolute different scenario and the squared different scenario. The resultant test statistics comprise $L_{AB}^{OLS}, L_{AB}^{LAD}, L_{AB}^{S}, L_{SQ}^{OLS}, L_{SQ}^{LAD}$, and L_{SQ}^{S}. Simulations of the test statistics draw on a Monte Carlo technique and are repeated 1,000 times constituting three patterns of data distribution: a normal distribution, a logistic distribution, and a lognormal distribution. The differences between the ratios of variances are determined using a non-centrality parameter value. The research results show that the Levene's test statistic performs better in the absolute different scenario than in the squared different scenario. Additionally, the test statistic L_{AB}^{S}, one of the test statistics in the absolute different scenario used to carry out the parameter estimation of the regression model in Stage 1 using the S-estimation method and of the regression model in Stage 2 using the OLS method, is the most efficient in all situations. Simulations of the six test statistics and their applications to actual data lead to comparable results. Based on the findings, it can be concluded that L_{AB}^{S} is a highly efficient test statistic that is robust to logistically, and lognormally distributed data.

Key-Words:- Robust Levene's test, homogeneity of variances, ordinary least squares, least absolute deviation, S-estimation method, heavy-tailed distribution.

Received: August 17, 2024. Revised: December 11, 2024. Accepted: January 4, 2025. Published: March 14, 2025.

1 Introduction

Inferential statistics play an essential role in research in many fields. In most clinical trials, the main interest is to test whether there are differences in the mean outcomes among the treatment groups. A typical test statistic is a t-test for a two-group comparison. In the case of more than two groups, an ANOVA F-test is used to test the equality for all groups, [1]. In economics and finance, ANOVA is a fundamental statistical technique used to compare means between different groups and test the equality hypothesis. Within, this turns into a potent method for evaluating policy efficacy, examining market segmentation, and investigating the economic effects of diverse elements across multiple populations or historical periods, [2]. In the field of educational research, the independent sample t-test is a crucial statistical instrument that provides a methodical and rigorous way to assess the effects of interventions, teaching strategies, and educational policy, [3]. In addition, [4] review examines the quality of reporting for two statistical tests, t-test, and ANOVA, for papers published in a selection of physiology journals in June 2017. Of the 328 original research articles examined, 277 (84.5%) included an ANOVA or t-test or both, and in 95% of the papers that used ANOVA, most papers also omitted the information and assumptions needed to verify ANOVA results. One of the fundamental assumptions for the analysis of variance using the F-test statistic is the homogeneity of the variances of k datasets. Violating this assumption will deteriorate the reliability of hypothesis testing regarding the consequence of violating such an assumption. [5] postulate three possibilities. First, it may stem from the mild effect on the statistical significance level of the F-test statistic of that data, characterized by a large sample size with datasets of equal size and low dataset variances. Alternatively, the violation of the assumption may be attributable to the moderate effect on the statistical significance level of the F-test statistic of the data that feature a large sample size with datasets of unequal size and low dataset variances, thereby resulting in the probability of Type I error lower than the significance level. Finally, it may reflect the strong effect on the statistical significance level of the F-test statistic of the data that typify a small sample size with datasets of unequal size and high dataset variances, thereby contributing to the probability of Type I error higher than the significance level and lower power of the test. No matter which possibility, the violation of the assumption concerning the homogeneity of the variances of k datasets should be strictly avoided. Similarly, [6] states that violating this assumption has a severe consequence on the power of the F-test statistic, especially in the case of datasets of unequal size.

For testing the homogeneity of variances, several methods are available, such as the Box-Anderson test [7], Levene test [8], the Brown-Forsythe test [9], the jackknife [10], Bartlett's test [11], bootstrapping [12]. [13] introduce a test using the generalized p-value approach, and compare it with the Bartlett test for homogeneity of variances. [14] have presented a test statistic based on the computational approach test (CAT), a parametric bootstrap case based on simulation and numerical computations; the CAT method uses the maximum likelihood estimates (MLEs) and does not require knowledge of any sampling distribution. [15] introduce the Standardized Likelihood Ratio Test (SLRT) for Homogeneity of Variance under Normality. [16] have presented a robust test for checking the homogeneity of variance for comparing two-sample tests. A modified structural zero removal method is applied to the Brown–Forsythe transformation. The study results found that robust test statistics are powerful to small or unequal sample sizes across many distributions. [17] propose new test statistics for the homogeneity of several variances against tree-ordered alternatives based on the inferential model (IM) and compare the performance of the developed test statistic with Spurrier's test, test based on isotonic estimators, and test based on sample quasi-range. The results found that the proposed test statistic is the only test used for unequal sample sizes. [18] propose new test statistics for comparing several variances with a control using the marginal inferential model (MIM). The key idea of the MIM is to reduce the dimension of the auxiliary variable, and the MIM test statistic effectively controls the type I error rate and power of the test compared with that of Spurrier's optimal test. [19] purpose A new exact p-value approach for testing variance homogeneity by developing a practically valuable procedure to calculate the null distribution, i.e., the p-value of the restrictive maximum likelihood-ratio (RELR) statistic. [20] suggested an adjusted Bartlett's test (ABT) based on the equal mean principle. [21] re-examined the computational approach test (CAT), initially introduced by [22]. [1] have studied the statistical tests for homogeneity of variance for clinical trials. The study's results found that, for two-sample problems, the Jackknife method tends to outperform others regardless of the variance ratio or the sample size. For more than two groups, Barlett's and Cochran's tests are better when data are nearly normally distributed; otherwise, Levene's test is a better choice for non-normally distributed data. Among these, Levene's test is regarded as one of the most efficient and widely used methods for testing the homogeneity of the variances of k datasets.

Therefore, the present research aims to develop a robust Levene's test that satisfies the requirements concerning normal data distribution and applies it to testing the homogeneity of the variances of k datasets. However, due to the prevalence of actual data involving extreme events (positive or negative ones) that cannot be dealt with merely with normal distribution principles, such as economic, financial, and astrological data, this study reformulates the test using a two-stage regression framework. The research procedures comprise developing and analyzing the robust Levene's test for testing the homogeneity of the variances of k datasets, simulations of the test statistic, and applications of the test statistic to actual datasets.

2 Materials and Methods

2.1 Development of the Test Statistics by Reformulating Levene's Test using a Two-Stage Regression Framework

The purpose of this study is to develop a robust Levene's test for testing the homogeneity of the variances of k datasets by the test hypothesis as follows:

$H_0 : \sigma_1^2 = \sigma_2^2 = ... = \sigma_k^2$ Versus $H_1 : \sigma_i^2 \neq \sigma_j^2$
for some i, j , $i \neq j$ $i, j \in \{1, 2, ..., k\}$,

The test statistics have been developed by using a two-stage regression framework. In regression analysis, when considered in terms of regression framework the linear equations can be expressed using metric notation as:

$$\underline{y} = \underline{X}\underline{\beta} + \underline{\varepsilon} \tag{1}$$

where $\underline{y}$ is $n \times 1$ random vector of response, $\underline{\varepsilon}$ is vector of random error, $\underline{\beta}$ is $(k+1) \times 1$ vector of unknown parameters and $\underline{X}$ is $n \times (k+1)$ metric of scalars. The model in equation (1) is called a full rank model. Namely, the metric $\underline{X}$ is full rank. It can be said that the Least Square Estimator of $\underline{\beta}$ is denoted by $\hat{\underline{\beta}} = (\underline{X}'\underline{X})^{-1}\underline{X}'\underline{y}$. In addition, in applied statistics, "analysis of variance" is often introduced by first considering the one-way classification model with fixed effect. The model in general is given by:

$$y_i = \mu + \tau_j + \varepsilon_i, \; i = 1, 2, ..., n, \; j = 1, 2, ..., k \tag{2}$$

where k is the number of treatments, n_j denotes the number of response available at the j^{th} level, and $n = \sum_{j=1}^{k} n_j$. In matrix notation, the model can be expressed in the form:

$$\underline{y} = \underline{X}\underline{\beta} + \underline{\varepsilon} \tag{3}$$

where $\underline{y}$ is vector of responses of dimension $n \times 1$.

$\underline{\beta}$ is vector of parameter $\beta' = [\mu \;\; \tau_1 \;\; \tau_2, ..., \tau_k]$.
$\underline{X}$ is design metric of dimension $n \times (k+1)$.
$\underline{\varepsilon}$ is $n \times 1$ vector of random error.

when the design metric and vector of the parameters of the new model are as follows:

$$\underline{y} = \left[\begin{array}{c} y_1 \\ y_2 \\ \vdots \\ y_{n_1} \\ \hline y_{n_1+1} \\ y_{n_1+2} \\ \vdots \\ y_{n_1+n_2} \\ \hline \vdots \\ \hline y_{n_1+,...,n_{k-1}+1} \\ y_{n_1+,...,n_{k-1}+2} \\ \vdots \\ y_{n_1+,...,n_{k-1}+n_k} \end{array}\right]_{\sum_{j=1}^{k} n_j \times 1} \qquad \underline{X} = \left[\begin{array}{ccccc} 1 & 1 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ \vdots & 1 & \vdots & & \vdots \\ 1 & 1 & 0 & \cdots & 0 \\ \hline 1 & 0 & 1 & \cdots & 0 \\ 1 & 0 & 1 & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ 1 & 0 & 1 & \cdots & 0 \\ \hline \vdots & & \vdots & & \vdots \\ \hline 1 & 0 & 0 & \cdots & 1 \\ 1 & 0 & 0 & \cdots & 1 \\ \vdots & & \vdots & & \vdots \\ 1 & 0 & 0 & \cdots & 1 \end{array}\right]_{\sum_{j=1}^{k} n_j \times k}$$

$$\underline{\beta} = \begin{bmatrix} \mu \\ \tau_1 \\ \tau_2 \\ \vdots \\ \tau_k \end{bmatrix}_{k \times 1} \qquad \underline{y} = \left[\begin{array}{c} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_{n_1} \\ \hline \varepsilon_{n_1+1} \\ \varepsilon_{n_1+2} \\ \vdots \\ \varepsilon_{n_1+n_2} \\ \hline \vdots \\ \hline \varepsilon_{n_1+,...,n_{k-1}+1} \\ \varepsilon_{n_1+,...,n_{k-1}+2} \\ \vdots \\ \varepsilon_{n_1+,...,n_{k-1}+n_k} \end{array}\right]_{\sum_{j=1}^{k} n_j \times 1}$$

The model in equations (2) and (3) is called a less than full rank model. In general, less than full rank model reason will make $\hat{\underline{\beta}} = (\underline{X}'\underline{X})^{-1}\underline{X}'\underline{y}$ have infinitely many solutions. One often used for the approach of the less than full rank model is reparameterization.

The model can be expressed in the form:

$$y_i = \mu_j + \varepsilon_i , \quad i = 1,2,...,n,\ j = 1,2,...,k \tag{4}$$

when the design metric and vector of the parameters of the new model are as follows:

$$\underline{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{n_1} \\ \hline y_{n_1+1} \\ y_{n_1+2} \\ \vdots \\ y_{n_1+n_2} \\ \hline \vdots \\ \hline y_{n_1+,...,n_{k-1}+1} \\ y_{n_1+,...,n_{k-1}+2} \\ \vdots \\ y_{n_1+,...,n_{k-1}+n_k} \end{bmatrix}_{\sum_{j=1}^{k} n_j \times 1} \qquad \underline{X} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 1 & 0 & \cdots & 0 \\ \hline 0 & 1 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 1 & \cdots & 0 \\ \hline \vdots & \vdots & & \vdots \\ \hline 0 & 0 & \cdots & 1 \\ 0 & 0 & \cdots & 1 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}_{\sum_{j=1}^{k} n_j \times k}$$

$$\underline{\beta} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_k \end{bmatrix}_{k\times 1} \qquad \underline{y} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_{n_1} \\ \hline \varepsilon_{n_1+1} \\ \varepsilon_{n_1+2} \\ \vdots \\ \varepsilon_{n_1+n_2} \\ \hline \vdots \\ \hline \varepsilon_{n_1+,...,n_{k-1}+1} \\ \varepsilon_{n_1+,...,n_{k-1}+2} \\ \vdots \\ \varepsilon_{n_1+,...,n_{k-1}+n_k} \end{bmatrix}_{\sum_{j=1}^{k} n_j \times 1}$$

From the new design metric and vector of the parameters,Thus $\underline{X}$ is $n \times k$ of rank k; it is now full rank. Therefore, the parameters that are estimated from $\hat{\underline{\beta}} = (\underline{X}'\underline{X})^{-1}\underline{X}'\underline{y}$ are unique [23]. In this study a robust Levene's test for testing the homogeneity of the variances of k datasets was developed from the concept of [24], this test using the principles of two-stage regression framework. The procedures are as follows:

Case 1. Absolute different Levene' s test
($z_i = |\hat{\varepsilon}_i| = |y_i - \hat{y}_i|$)

<u>Stage 1</u>. The basic principles are to estimate the parameter $\hat{\underline{y}} = \underline{X}\hat{\underline{\beta}}$ using the Ordinary Least Squares (OLS) method and calculate the error from $z_i = |\hat{\varepsilon}_i| = |y_i - \hat{y}_i|,\ i = 1,2,...,n,\ j = 1,2,...,k$, then working covariance metric is $\Sigma_{stage1} = \sigma^2 I$, where I is identity metric $\hat{\underline{y}} = (\underline{X}'\underline{X})^{-1}(\underline{X}'\underline{y}) = \underline{H}\underline{y}$, $\underline{H}$ is hat matric, $\hat{\varepsilon} \sim N(0, \Sigma(I-H))$, and $\hat{\varepsilon}_i \sim N(0, \sigma_i^2(1-h_{ii}))$, where $z_i = |\hat{\varepsilon}_i| = |y_i - \hat{y}_i|$ has a folder-normal
distribution pattern with the mean being a linear function of, where:

$$E(z_i) = \sigma_i \sqrt{\frac{2}{\pi}(1-h_{ii})},\ i = 1,2,...,n,\ j = 1,2,...,k . \tag{5}$$

In the case of the Absolute difference in Levene's test, z_i is the absolute error or the absolute value of the actual value that deviates from the predicted value, namely, $z_i = |\hat{\varepsilon}_i| = |y_i - \hat{y}_i|$. In addition to estimating the predicted value of $\hat{y}_i$ using OLS method, $\hat{y}_i$ is also estimated using LAD and S-estimation methods.

<u>Stage 2</u>. From Equation (5), the relationship between z_i and σ_i can be arranged in the form of

$\underline{z} = \underline{\delta}\underline{X} + \underline{e}$, or:

$$z_i = \alpha + \delta_1 X_{1i} + \delta_2 X_{2i} + ... + \delta_{(k-1)} X_{(k-1)i} + e_i ,$$
$$i = 1,2,...,n,\ j = 1,2,...,k . \tag{6}$$

where the test hypothesis $H_0 : \sigma_1^2 = \sigma_2^2 = ... = \sigma_k^2$ is reformulated as $H_0 : \delta_1 = \delta_2 = ... = \delta_k$ and the parameter estimation is conducted using the OLS method, thus $e \sim N(0 \Sigma_{stage2}$, $\Sigma_{stage2} = \sigma_z^2 I$, Finally, the test statistics has the following for:

$$L_{AB} = \frac{\sum_{i=1}^{n} (\hat{z}_i - \bar{\hat{z}}_i)^2 / k-1}{\sum_{i=1}^{n} (z_i - \hat{z}_i)^2 / n-k} \quad . \qquad (7)$$

where $z_i = |\hat{\varepsilon}_i| = |y_i - \hat{y}_i|$, $\hat{y}_i$ is the predicted value in stage 1 that performs with OLS, LAD and S-estimation method from stage 1.

$\hat{z}_i$ is the predicted value from the regression of model (6) that estimates the parameter with the OLS method from stage 2.

$\bar{\hat{z}}_i$ is the predicted value from regression of $\underline{z}$ on $\underline{1}$ from equation (6), where $\underline{1}$ is the first column of the transformed design matrix $\underline{X}$.

The reformulation of the absolute different Levene's test in the first stage involves parameter estimation using the ordinary least squares (OLS) method, the Least Absolute Deviation (LAD) method, and the S-estimation method, while the second employs only the OLS method. From equation (7), the test statistic L_{AB} follows an approximate $F_{(k-1,n-k)}$ distribution under the null hypothesis of $H_0 : \sigma_1^2 = \sigma_2^2 = ... = \sigma_k^2$, and a $\chi_{k-1}^2 / k-1$ distribution asymptotically as $n \to \infty$ [24]. In addition, the test statistics in this study are developed in the terms of squared difference Levene's test. The procedures are as follows:

Case 2. Squared difference Levene's test

$z_i = (\hat{\varepsilon}_i)^2 = (y_i - \hat{y}_i)^2$

Stage 1. The basic principles are to estimate the parameter $\underline{\hat{y}} = \underline{X}\underline{\hat{\beta}}$ using the Ordinary Least Squares (OLS) method and calculate the error from $z_i = (\hat{\varepsilon}_i)^2 = (y_i - \hat{y}_i)^2$, $\quad i = 1, 2, ..., n, \ j = 1, 2, ..., k$, where $\underline{\hat{\varepsilon}} = \underline{y} - \underline{\hat{y}} \sim N(0, \Sigma(I-H))$, will say that the n independent standard normal random variable of $\frac{(\hat{\varepsilon})^2}{\Sigma(I-H)}$ is χ_n^2 distribution, and $(\underline{\hat{\varepsilon}})^2 \sim \Sigma(I-H)\chi_n^2$, we get the $E(z_i)$ is a linear function of σ_i, where $\text{var}(\varepsilon_i) = E(\varepsilon_i)^2 - [E(\varepsilon_i)]^2$, and $\lfloor E(\varepsilon_i) \rfloor = 0$, we can rewrite $\text{var}(\varepsilon_i)$ as $\text{var}(\varepsilon_i) = E(\varepsilon_i)^2$. The simplified formula is then:

$$E(\varepsilon_i)^2 = E(z_i) = \sigma_i^2 (1 - h_{ii}),$$

$$i = 1, 2, ..., n, \ j = 1, 2, ..., k \qquad (8)$$

In the case of the Squared difference Levene's test, z_i is the squared error or the squared difference between the actual value that deviates from the predicted value, namely, $z_i = (\hat{\varepsilon}_i)^2 = (y_i - \hat{y}_i)^2$. In addition to estimating the predicted value of $\hat{y}_i$ using the OLS method, $\hat{y}_i$ is also estimated using LAD and S-estimation methods.

Stage 2. From Equation (8), the relationship between z_i and σ_i can be arranged in the form of

$\underline{z} = \underline{\delta}\underline{X} + \underline{e}$, or:

$$z_i = \alpha + \delta_1 X_{1i} + \delta_2 X_{2i} + ... + \delta_{(k-1)} X_{(k-1)i} + e_i,$$
$$i = 1, 2, ..., n, \ j = 1, 2, ..., k, \qquad (9)$$

and the parameter estimation is conducted using the OLS method. Where $\underline{e} \sim N(0, \Sigma_{stage2})$, $\Sigma_{stage2} = \sigma_z^2 I$. Finally, the test statistics has the following for

$$L_{SQ} = \frac{\sum_{i=1}^{n} (\hat{z}_i - \bar{\hat{z}}_i)^2 / k-1}{\sum_{i=1}^{n} (z_i - \hat{z}_i)^2 / n-k} \quad . \qquad (10)$$

where $z_i = (\hat{\varepsilon}_i)^2 = (y_i - \hat{y}_i)^2$, $\hat{y}_i$ is the parameter estimate in stage 1 that perform with OLS, LAD and S-estimation method.

$\hat{z}_i$ is the predicted value from the regression of model (9) that estimates the parameter with the OLS method.

$\bar{\hat{z}}_i$ is the predicted value from regression of $\underline{z}$ on $\underline{1}$ from equation (9), where $\underline{1}$ is the first column of the transformed design mat $\underline{X}$.

The reformulation of the square different Levene's test in the first stage involves parameter estimation using the ordinary least squares (OLS) method, the Least Absolute Deviation (LAD) method, and the S-estimation method, while the second employs only the OLS method. From equation (10), the test statistic L_{SQ} follows approximately $F_{(k-1,n-k)}$ distribution under the null hypothesis of $H_0 : \sigma_1^2 = \sigma_2^2 = ... = \sigma_k^2$, and a $\chi_{k-1}^2 / k-1$ distribution asymptotically as $n \to \infty$ [24].

2.2 Parameter Estimate in a Two-Stage Regression Framework

2.2.1 Ordinary Least Squares (OLS)

Let $\underset{\sim}{y}$ denote the vector of responses , $\underset{\sim}{\varepsilon}$ denote a random vector of residual with mean 0 and variance $\sigma^2 I$, and $\underset{\sim}{\beta}$ is a vector of unknown parameters. The least square estimator of $\underset{\sim}{\beta}$ is $\hat{\underset{\sim}{\beta}}$ that minimize the sum of squares of the residuals $\underset{\sim}{\varepsilon}'\underset{\sim}{\varepsilon} = \sum_{i=1}^{n} \varepsilon_i^2 = (\underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{\beta})'(\underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{\beta})$. The estimator of $\underset{\sim}{\beta}$ is given by $\hat{\underset{\sim}{\beta}} = (\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{X}'\underset{\sim}{y}$, [23].

2.2.2 Least Absolute Deviation (LAD)

LAD is a statistical optimality criterion and the statistical optimization technique that similar to the least squares technique. It is the robust method that minimizes the sum of the absolute value of the residual $\sum_{i=1}^{n}|\varepsilon_i| = \sum_{i=1}^{n}|y_i - X_{ij}\beta_j|$ [25]. The problem can be solved using any linear programming technique, We wish to

$$\min imize \sum_{i=1}^{n}|\underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{\beta}| = \min imize \sum_{i=1}^{n} a_i ,$$

with respect to $\underset{\sim}{\beta}$ and a_i , subject to

$$a_i \ge y_i - X_i\beta_j$$

$$a_i \ge -\lfloor y_i - X_i\beta_j \rfloor ,$$

for $i = 1, 2, ..., n, \ j = 1, 2, ..., k$.

The method of LAD finds applications in many areas, due to its robustness against the outliers compared to the least squares method. At the same time, the LAD method may be limited in the case of unstable solutions or possibly multiple solutions.

2.2.3 S-estimation

S-estimators was proposed by [26]. It is a robust estimation method for regression models that minimize the dispersion for the residuals with considering the minimum robust scale estimator that is determined by the ρ function, and the objective function is:

$$\min \sum_{i=1}^{n} \rho\left(\frac{Y_i - \sum_{i=1}^{n} X_i\beta_j}{\hat{\sigma}_s}\right) = \min \sum_{i=1}^{n} \rho\left(\frac{e_i}{\hat{\sigma}_s}\right),$$

for $i = 1, 2, ..., n, \ j = 1, 2, ..., k$. (11)

where $e_1, e_2, ..., e_n$ is the i^{th} residual , $\hat{\sigma}_s$ is a minimum robust scale estimator, [27], [28]. The procedures of S–estimation is as follows.

1. Estimate regression coefficients on the data with Ordinary Least Square (OLS).
2. Check the assumptions of the classical regression model , and detect outlier in the data set.
3. Calculate $\hat{\beta}_0$ with Ordinary Least Square (OLS) .
4. Calculate the residual with $e_i = y_i - \hat{y}_i$.
5. Calculate $\hat{\sigma}_i$ from

$$\hat{\sigma}_i = \begin{cases} \dfrac{\text{median}\left|e_i \text{-median } e_i\right|}{0.6745} & , \text{ iteration} = 1 \\ \sqrt{\dfrac{i}{nk}\sum_{i=1}^{n} w_i e_i^2} & , \text{ iteration} > 1 \end{cases} .$$

6. Calculate value $u_i = \dfrac{e_i}{\hat{\sigma}_i}$.
7. Calculate weighted value (w_i) from

$$w_i = \begin{cases} \begin{cases} \left[1 - \left(\dfrac{u_i}{1.547}\right)^2\right]^2 & , \ |u_i| \le 1.547 \\ 0 & , \ |u_i| > 1.547 \end{cases} & , \text{ iteration} = 1 \\ \dfrac{\rho(u)}{u^2} & , \text{ iteration} > 1 \end{cases} .$$

8. Calculate $\hat{\beta}_S$ with Weighted Least Square (WLS) method with wighted w_i.
9. Repeat from steps 4 -7 to obtain a convergent value of $\hat{\beta}_S$.

3 Simulation Study

The purpose of this study is to develop a robust Levene's test for testing the homogeneity of the variances of k datasets (k=3) by reformulating the test using a two-stage regression framework. The procedures of simulation study are as follows:

1. Data distribution patterns :Simulations of the six test statistics are performed to address the following three data distribution patterns: a normal distribution, a logistic distribution, and a lognormal distribution. In the case of equlity variance, the values of the location parameter

μ and of the scale parameter σ^2 of three populations are set at 0 and 10, respectively.

2. Determination of the number of populations for hypothesis testing according Table 1. The number of populations for hypothesis testing is determined at three, and the simulations are done for cases of both equal and unequal populations with the total sample sizes equaling 45, 90, and 180 and the average sample sizes equaling 15, 30, and 60, [29].

Table 1. Determination of the number of populations for hypothesis testing

size	sample size	
	equal	unequal
small	(15,15,15)	(10,15,20)
medium	(30,30,30)	(25,30,35)
large	(60,60,60)	(50,60,70)

3. Determination of the differences between the ratios of variances: The differences between the ratios of variances are determined using a non-centrality parameter value (ϕ), [30] .

$$\phi = \frac{\left(\sum_{j=1}^{k}(\sigma_j^2 - \bar{\sigma}^2)^2 / k\right)^{1/2}}{\sigma_1^2} \quad (12)$$

σ_j^2 is the population variance with the j[th] group, $j = 1, 2, \ldots, k$.

σ_1^2 is the population variance with the lowest.

$\bar{\sigma}^2$ is the mean of population variance with k groups.

k is the number of population groups, in this study , $k = 3$.

Table 2. Determination of the ratio of variance by non-centrality parameter (ϕ)

levels		ratio of variance	ϕ
slightly	$(0 < \phi < 1.5)$	1 : 2 :3	0.816
moderately	$(1.5 \le \phi < 3.0)$	1 : 3 : 5	1.633
highly	$(\phi \ge 3.0)$	1 : 5 : 10	3.682

From Table 2, in the case of each population, there are different variances, given the level of difference into three levels: slightly, moderately, and highly, respectively:

- In the case of a slightly different variance, the variance ratio is 1: 2 :3, generate the variance of population group 1, group 2, and group 3 is equal to 10, 20, and 30, respectively. When substituting the variance of each population group according to Equation 12, we get the value ϕ =0.816, which is in the range (0< ϕ<1.5). Figure 1 shows data simulation in the case of slightly different variances.

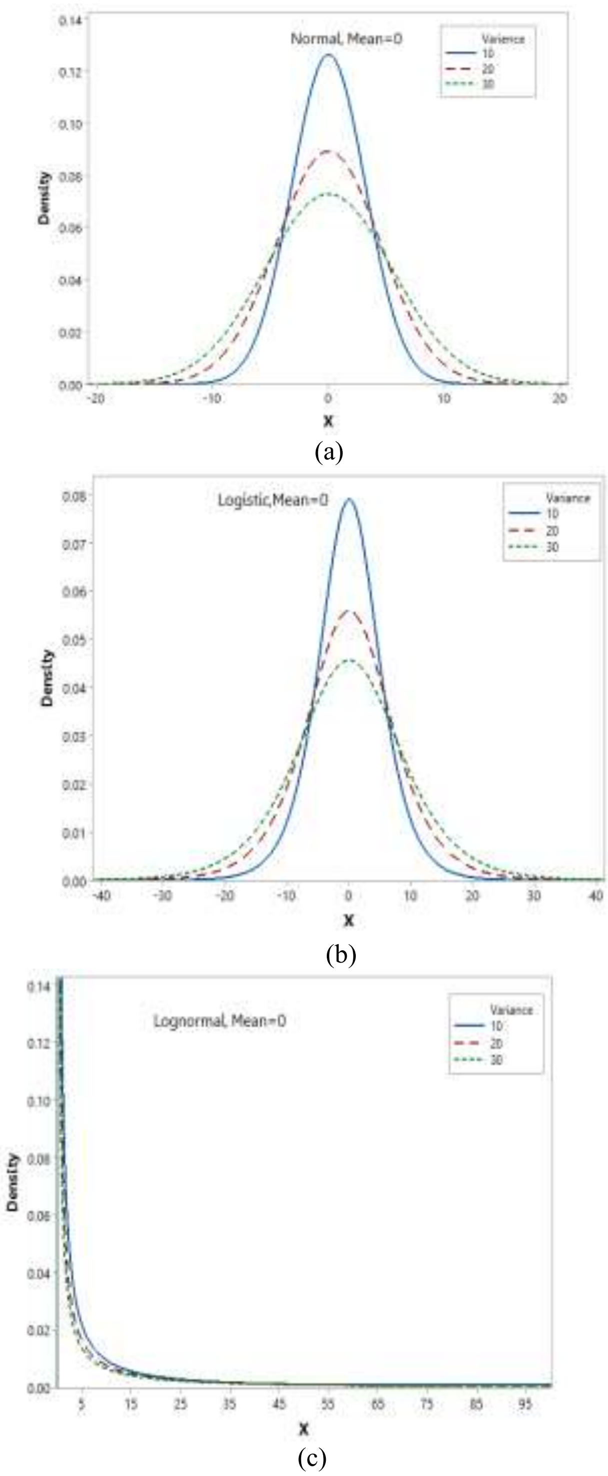
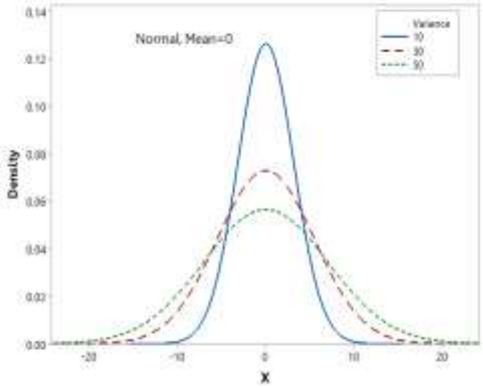
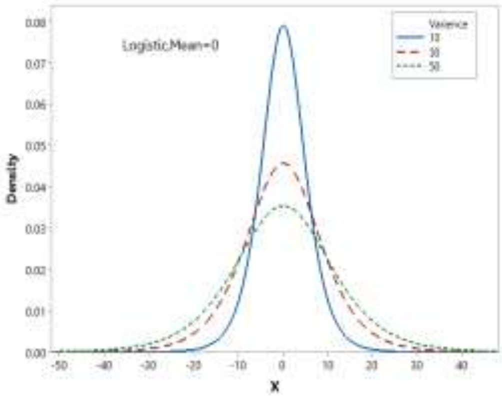


Fig. 1: Illustrate generating three groups of data with a ratio of differences variances of 1:2:3, and the data distribution is in three formats: a) normal distribution, b) Logistic distribution, and c) Lognormal distribution

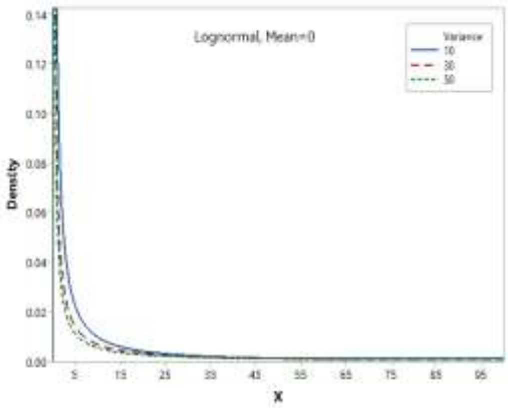
- In the case of a moderately different variance, the variance ratio is 1: 3 : 5, generating the variance of population group 1, group 2, and group 3 is equal to 10, 30, and 50, respectively. When substituting the variance of each population group according to Equation 12, we get the value ϕ =1.633, which is in the range ($1.50\leq \phi<3.0$). Figure 2 shows data simulation in the case of moderately different variances.



(a)



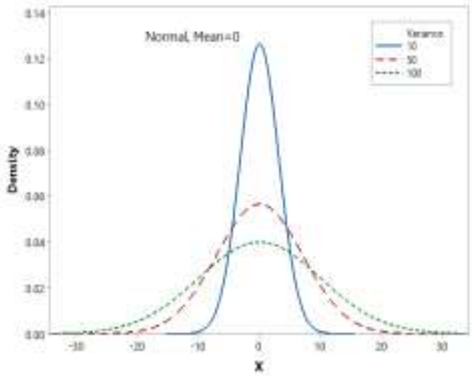
(b)



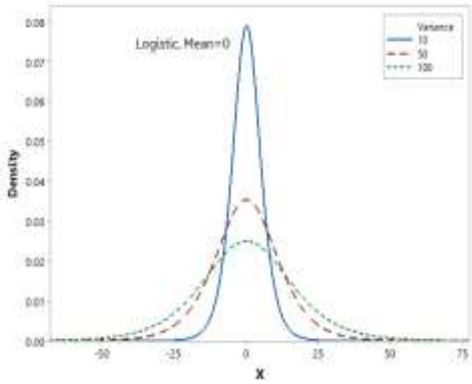
(c)

Fig. 2: Illustrate generating three groups of data with a ratio of differences variances of 1:3:5, and the data distribution is in three formats: a) normal distribution, b) Logistic distribution, and c) Lognormal distribution

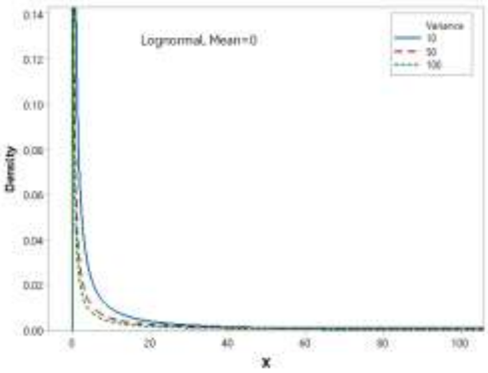
- In the case of a highly different variance, the variance ratio is 1: 5 : 10, generating the variance of population group 1, group 2, and group 3 is equal to 10, 50, and 100, respectively. When substituting the variance of each population group according to Equation 12, we get the value ϕ =3.682, which is in the range ($\phi\geq3.0$). Figure 3 shows data simulation in the case of highly different variances.



(a)



(b)



(c)

Fig. 3: Illustrate generating three groups of data with a ratio of differences variances of 1:5:10, and the data distribution is in three formats: a) normal distribution, b) Logistic distribution, and c) Lognormal distribution

4. Calculation of the Levene's test statistic values in the absolute different scenario and the square

different scenario: The Levene's test statistic values in the absolute different scenario and the squared different scenario are calculated from the parameter estimation in Stage 1 using the OLS method, the LAD method, and the S-estimation method and the parameter estimation in Stage 2 using only the OLS method. As a consequence, the test statistics comprise $L_{AB}^{OLS}, L_{AB}^{LAD}, L_{AB}^{S}, L_{SQ}^{OLS}, L_{SQ}^{LAD}$ and L_{SQ}^{S}, where the symbols AB and SQ represent Levene's test statistic in the absolutely different scenario and the squared different scenario, respectively, and the symbols OLS, LAD, and S represent the parameter estimation methods in Stage 1. Then the calculated test statistic values are compared against the statistical levels of significance, pre-determined at $\alpha = 0.01$ and $\alpha = 0.05$. Also, the probabilities of Type I error, i.e. rejecting the null hypothesis (H_0) when it is true, and the power of the test, i.e. rejecting H_0 when it is false, are calculated from 1,000 replications.

5. Comparison the performance of the test statistics for control the type I error using Bradley's Criteria [31]. The control of Type I error based on Bradley's liberal criterion of robustness , where:
   α represents the occurrence of Type I error.
   $\hat{\alpha}$ represents the estimated value of the occurrence of Type I error.

For Bradley's liberal, a test can be considered robust of the rate of type I error, $\hat{\alpha}$ is within the interval 0.5α and 1.5α. The finding indicates that the control ranges of Type I error when $\alpha = 0.01$ and $\alpha = 0.05$ are $[0.005, 0.015]$ and ,$[0.025, 0.075]$, respectively.

4 Result

The results relating to the ability to control Type I error, i.e. rejecting the null hypothesis (H_0) when it is true, show that in case of normal and logistic distributions, all the six test statistics, $L_{AB}^{OLS}, L_{AB}^{LAD}, L_{AB}^{S}, L_{SQ}^{OLS}, L_{SQ}^{LAD}$ and L_{SQ}^{S} are able to control Type I error in all the situations at the significance levels of both $\alpha = 0.01$ and $\alpha = 0.05$. Conversely, in case of a lognormal distribution, only the test statistic L_{AB}^{S} is efficient in controlling Type I error provided the sample size is large, i.e. (60,60,60) and (50,60,70). The information is shown in Appendix in Table 3, Table 4 and Figure 4 and Figure 5.

The findings relating to the power of the test, i.e. rejecting H_0 when it is false, demonstrate that in case the differences between the ratios of variances are low (1:2:3) at the significance level of $\alpha = 0.01$, the Levene's test fares better in the absolute different scenario than in the squared difference of scenario. Additionally, among all the test statistics, L_{AB}^{S}, one of those in the absolutely different scenario used for the parameter estimation of the regression model in Stage 1 using the S-estimation method, is the most efficient. In addition, another key factor determining the efficiency of the test statistics is the sample size, with large and equal sample sizes strengthening the power of the test and vice versa. Also, the power of the test statistics increases with a normal distribution, followed in order by a logistic distribution and a lognormal distribution. As for the lognormal distribution, Levene's test in the absolute different scenario significantly outperforms its counterpart in the squared difference scenario with the test statistic L_{AB}^{S} being noticeably more efficient than the test statistics L_{AB}^{OLS} and L_{AB}^{LAD}. The results in case the differences between the ratios of variances are low (1:2:3) at the significance level of $\alpha = 0.05$ illustrate a similar trend except for the comparable power of the test statistics $L_{AB}^{OLS}, L_{AB}^{LAD}, L_{AB}^{S}$ regardless of whether the data are normally or lognormally distributed. The information are shown in Appendix in Table 5 and Table 6.

The findings relating to the power of the test, i.e. rejecting H_0 when it is false, demonstrate that in case the differences between the ratios of variances are moderate (1:3:5), the result found that the power of the test is higher than the low ratios of variance. At the significance level of $\alpha = 0.01$, the Levene's test fares better in the absolute different scenario than in the squared difference scenario. Additionally, among all the test statistics, L_{AB}^{S}, one of those in the absolutely different scenario used for the parameter estimation of the regression model in Stage 1 using the S-estimation method, is the most efficient. In addition, another key factor determining the efficiency of the test statistics is the sample size, with large and equal sample sizes strengthening the power of the test and vice versa. Also, the power of the test statistics increases with a normal distribution, followed in order by a logistic distribution and a lognormal distribution. As for the lognormal distribution, Levene's test in the absolute different scenario significantly outperforms its counterpart in the squared different scenario with the test statistic L_{AB}^{S} being noticeably more efficient than the test statistics L_{AB}^{OLS} and L_{AB}^{LAD}. The results in case the differences between the ratios of variances are moderate (1:3:5) at the significance level of $\alpha = 0.05$ illustrate a similar trend except for the comparable power of the test statistics $L_{AB}^{OLS}, L_{AB}^{LAD}, L_{AB}^{S}$ regardless of whether the data are normally or lognormally distributed. The information is shown in Appendix Table 7 and Table 8.

The findings relating to the power of the test, i.e. rejecting H_0 when it is false, demonstrate that in case the differences between the ratios of variances are high (1:5:10), the result found that the power of the test is higher than the low ratios of variance. At the significance level of $\alpha = 0.01$, the Levene's test fares better in the absolute different scenario than in squared different scenario. Additionally, among all the test statistics, L_{AB}^{S}, one of those in the absolute different scenario used for the parameter estimation of the regression model in Stage 1 using the S-estimation method, is the most efficient. In addition, another key factor determining the efficiency of the test statistics is the sample size, with large and equal sample sizes strengthening the power of the test and vice versa. Also, the power of the test statistics increases with a normal distribution, followed in order by a logistic distribution and a lognormal distribution. As for the lognormal distribution, Levene's test in the absolute different scenario significantly outperforms its counterpart in the squared difference with the test statistic L_{AB}^{S} being noticeably more efficient than the test statistics L_{AB}^{OLS} and L_{AB}^{LAD}. The results in case the differences between the ratios of variances are high (1:5:10) at the significance level of $\alpha = 0.05$ illustrate a similar trend except for the comparable power of the test statistics $L_{AB}^{OLS}, L_{AB}^{LAD}, L_{AB}^{S}$ regardless of whether the data are normally or lognormally distributed. The information is shown in Appendix in Table 9 and Table 10 and Figure 6 and Figure 7.

5 Application of the Test Statistics to Actual Data

The application of the test statistics $L_{AB}^{OLS}, L_{AB}^{LAD}, L_{AB}^{S}, L_{SQ}^{OLS}, L_{SQ}^{LAD}$ and L_{SQ}^{S} is carried out to test the homogeneity of the variances of two actual datasets each comprising three subsets of data as follows:

5.1 The average household expenditure in 17 Northern Thai provinces covering the years 2009, 2010, and 2011 [32].

5.2 The average marriage registration in 75 provinces across Thailand covering the years 2009, 2010, and 2011, [33].

The means and standard deviations of the two data sets are presented in Table 11 (Appendix). The data distribution of the two datasets, derived from the Anderson-Darling test, is displayed in Figure 8 and Figure 9.

In terms of the distribution of the data, the first dataset is found to demonstrate both a normal distribution and a logistic distribution at the significance level of $\alpha = 0.05$, while the second features a lognormal distribution at the significance level of $\alpha = 0.0$. In terms of the homogeneity of variances determined from the test statistics $L_{AB}^{OLS}, L_{AB}^{LAD}, L_{AB}^{S}, L_{SQ}^{OLS}, L_{SQ}^{LAD}$ and L_{SQ}^{S}, the findings reveal that the three subsets of data in both the datasets do not differ significantly at the significance level of $\alpha = 0.05$ with the Levene's test yielding a higher p in the absolutely different scenario than in the squared different scenario. Among all the test statistics, L_{AB}^{S} leads to the highest p for both datasets. Additionally, for the first dataset, which features normal and logistic distributions, all the test statistics produce comparable p values. Conversely, for the second, which features a lognormal distribution, the test statistics in the absolute different scenario, $L_{AB}^{OLS}, L_{AB}^{LAD}, L_{AB}^{S}$, bring about a relatively much higher p than those in the squared different scenario, $L_{SQ}^{OLS}, L_{SQ}^{LAD}, L_{SQ}^{S}$. Table 12 (Appendix) shows the results of the data analysis from the actual data. The result found that the null hypothesis $H_0 : \sigma_1^2 = \sigma_2^2 = \sigma_3^2$ is accepted for both data sets. In the first data set, where each group had a normal distribution, the six test statistics gave the test value, and the p values were similar. For the second set of data, where each data group has a Lognormal distribution, the values of the test statistics in the Absolute different Levene's test group are higher than the Square different Levene's test group, with the test statistic L_{AB}^{S} giving the highest p-value.

6 Discussion and Conclusions

The objective of this study is to develop a robust Levene's test for testing the homogeneity of the variances of k datasets ($k=3$) by reformulating the test in the form of a two-stage regression framework. The first stage involves parameter estimation using the Ordinary Least Square (OLS) method, the Least Absolute Deviation (LAD) method, and the S-estimation method, while the second employs only the OLS method. In this study, the results demonstrate the ability to test the homogeneity of the variances of k datasets in the case of normal, logistic, and lognormal distributions and present six test statistics, including $L_{AB}^{OLS}, L_{AB}^{LAD}, L_{AB}^{S}, L_{SQ}^{OLS}, L_{SQ}^{LAD}$, and L_{SQ}^{S}. The results of the study found that the efficiency of the test statistics in the absolute different scenario, $L_{AB}^{OLS}, L_{AB}^{LAD}, L_{AB}^{S}$, is higher than that of the test statistics in the squared different scenario, $L_{SQ}^{OLS}, L_{SQ}^{LAD}, L_{SQ}^{S}$. In addition, among those in the former scenario, the test statistic L_{AB}^{S} is the most efficient in all situations. Additionally, in the case of normal and logistic distributions, the efficiency of the test statistics $L_{AB}^{OLS}, L_{AB}^{LAD}, L_{AB}^{S}$ does not differ significantly in terms of both the ability to control Type I error and the power of the test. Conversely, in case of a lognormal distribution, the test statistic L_{AB}^{S} is clearly more efficient than the test statistics $L_{AB}^{OLS}, L_{AB}^{LAD}$ in both aspects. However, with large and equal sample sizes, the test statistics $L_{AB}^{OLS}, L_{AB}^{LAD}$ fare equally at the significance level of $\alpha=0.05$ regardless of whether the data are normally, logistically, or lognormally distributed. Based on the present findings, the test statistic L_{AB}^{S} is shown to be the most robust to all distribution patterns, especially logistic and lognormal distributions, with the simulation results being consistent with those obtained from the applications to actual data. Therefore, the purpose test statistics are another option of a test statistic that effectively checks the necessary initial assumptions of the test statistic about equality of variances. The proposed test statistics are robust to data with heavier tails, such as logistic distributions, and data with positive skewness, such as lognormal distributions. For further research, an analysis should be extended to the homogeneity of the variances of dependent k datasets to broaden the knowledge in such areas as the paired sample t-test when $k=2$ and the repeated measures ANOVA when $k>2$.

Acknowledgement:
The authors gratefully acknowledge Rajamangala University of Technology Thanyaburi that provides support in research. We are also thankful to those who could not be mentioned here for their kindness and encouragement. And finally, the authors would like to thank the anonymous reviewers for their comments and suggestions.

Declaration of Generative AI and AI-assisted Technologies in the Writing Process
During the preparation of this work the authors used Grammarly in order to check grammar. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

References:
[1] Zhou, Y., Zhu, Y., and Wong, W. K., Statistical tests for homogeneity of variance for clinical trials and recommendations. *Contemporary Clinical Trials Communications*, Vol.33, 101119, 2003, pp. 1-11. https://doi:10.1016/j.conctc.2023.101119.
[2] Nand kishor Kumar, F-Test and Analysis of Variance (ANOVA) in Economics. *Mikailalsys Journal of Mathematics and Statistics*, Vol.2, No.3, 2024, pp. 102-113. https://doi:10.58578/mjms.v2i3.3449.
[3] AKPAN, E. E., and Clark, L. J., Independent T-Test Statistics: It's Relevance in Educational Research. *International Journal of Eminent Scholars*, Vol.10, No.1, 2023, pp.79-88.
[4] Weissgerber, T. L., Garcia-Valencia, O.,Garovic, V. D., Milic, N. M., and Winham, S. J., Why we need to report more than'Data were Analyzed by t-tests or ANOVA'. *Elife*, Vol.7, 2018, e36163. https://doi.org/10.7554/eLife.36163.
[5] Cochran, W.G., and Cox, G.M., *Experimental Design*, New York: John Hiley and Sons, 1976.
[6] Scheffé, H., *The Analysis of Variance*, John Wiley and Sons. Inc., New York, 1959.
[7] Box, G. E., and Andersen, S. L., Permutation theory in the derivation of robust criteria and the study of departures from assumption, *Journal of the Royal Statistical Society: Series B (Methodological)*, Vol.17, No.1, 1955, pp.1-26. https://doi.org/10.1111/j.2517-6161.1955.tb00176.x.
[8] Levene, H., Robust tests for equality of variances, *Contributions to probability and*

statistics. Essays in honor of Harold Hotelling, 1961, pp.279-292, [Online]. https://cir.nii.ac.jp/crid/1573668924369527296 (Accessed Date: October 1, 2024).

[9] Brown, M. B., and Forsythe, A. B., Robust tests for the equality of variances, *Journal of the American Statistical Association*, Vol.69, No. 346, 1974, pp.364-367. https://doi.org/10.2307/2285659.

[10] Bissell, A. F., and Ferguson, R. A.,The Jackknife—Toy, Tool or Two-Edged Weapon?. *Journal of the Royal Statistical Society: Series D (The Statistician),* Vol. 24, No.2, 1975, pp. 79-100. https://doi.org/10.2307/2987663.

[11] Snedecor, G. W., amd Cochran, W. G., *Statistical Methods, eight edition,* Iowa state University press, Ames, Iowa, 1989.

[12] Hall, P., and Wilson, S. R., Two guidelines for bootstrap hypothesis testing. *Biometrics*, Vol. 47, No. 2, 1991, pp. 757-762. https://doi.org/10.2307/2532163.

[13] Liu, X. and Xu, X., A new generalized p-value approach for testing the homogeneity of variances, *Statistics and probability letters*, Vol.80, No. 19, 2010, pp.1486-1491. https://doi.org/10.1016/j.spl.2010.05.017.

[14] Gökpınar, F. and Gökpınar, E., A Computational approach for testing of coefficients of variation in k normal population, *Hacettepe Journal of Mathematics and Statistics*, Vol.44, No.5, 2015, pp.1197-1213. https://doi:10.15672/HJMS.2014317482.

[15] Gökpınar, E., Standardized likelihood ratio test for homogeneity of variance based on likelihood ratio under normality. *Gazi University Journal of Science*, Vol.30, No. 3, 2017, pp. 223-235.

[16] Erps, R. C., and Noguchi, K., A robust test for checking the homogeneity of variability measures and its application to the analysis of implicit attitudes. *Journal of Educational and Behavioral Statistics*, Vol.45, No.4, 2020, pp. 403-425. https://doi:10.3102/1076998619883874.

[17] Kong, J., Jin, H., Lu, H., Lin, J., and Jin, K., An inferential model-based method for testinghomogeneity of several variances against tree-ordered alternatives. *International Journal of Approximate Reasoning*, Vol.152, 2023, pp.344-354. https://doi.org/10.1016/j.ijar.2022.11.006.

[18] Kong, J., and Lu, H., (2024). Comparing the variances of several treatments with that of a control treatment: Theory and applications. *Plos one*, Vol.19, No.1, 2024, e0296376. https://doi.org/10.1371/journal.pone.0296376.

[19] Wang, J., Li, X., and Liang, H., A new exact p-value approach for testing variance homogeneity. *Statistical Theory and Related Fields*, Vol.6, No.1, 2022, pp. 81-86. https://doi:10.1080/24754269.2021.1907519.

[20] Ma, X. B., Lin, F. C., and Zhao, Y., An adjustment to the Bartlett's test for small sample size. *Communications in Statistics-Simulation and Computation*, Vol.44, No.1, 2015, pp. 257-269. https://doi.org/10.1080/03610918.2013.773347.

[21] Gokpinar, E., and Gokpinar, F., Testing equality of variances for several normal populations. *Communications in Statistics-Simulation and Computation*, Vol.46, No.1, 2017, pp. 38-52. https://doi:10.1080/03610918.2014.955110.

[22] Pal, N., Lim, W. K., and Ling, C.H., A computational approach to statistical inferences. *Journal of Applied Probability and Statistics*, Vol.2, No.1, 2007, pp.13–35.

[23] Myers, R. H., and Milton, J. S., *A first course in the theory of linear statistical models*. Kent Publishing Company, 1991.

[24] Soave, D., and Sun, L., A generalized Levene's scale test for variance heterogeneity in the presence of sample correlation and group uncertainty. *Biometrics*, Vol.73, No.3, 2017, pp. 960-971. https://doi.org/10.1111/biom.12651.

[25] Dodge, Y. (2008). *The concise encyclopedia of statistics*. Springer Science and Business Media, 2008.

[26] Rousseeuw, P., and Yohai, V.(1984). *Robust regression by means of S-estimators. In Robust and nonlinear time series analysis*. Springer, New York, NY, 1984.

[27] Susanti, Y., and Pratiwi, H.M., estimation, S estimation, and MM estimation in robust regression. *International Journal of Pure and Applied Mathematics*, Vol.91, No.3, 2014, pp. 349-360. http://dx.doi.org/10.12732/ijpam.v91i3.7.

[28] Salibian-Barrera, M., and Yohai, V. J., A fast algorithm for S-regression estimates. *Journal of computational and Graphical Statistics*, Vol.15, No.2, 2006, pp. 414-427. https://doi.org/10.1198/106186006X113629.

[29] Reiczigel, J., Confidence intervals for the binomial parameter: some new considerations.

Statistics in medicine, Vol.22, No.4, 2003, pp.611-621. https://doi.org/10.1002/sim.1320.

[30] Games, P. A., Winkler, H. B., and Probert, D. A., Robust tests for homogeneity of variance. *Educational and Psychological Measurement*, Vol.32, No.4, 1972, pp. 887-909. https://doi.org/10.1177/001316447203200404

[31] Bradley, J. V., Robustness?. *British Journal of Mathematical and Statistical Psychology*, Vol.31, No.2, 1978, pp. 144-152. https://doi.org/10.1111/j.2044-8317.1978.tb00581.x.

[32] National Statistical Officer Thailand, The average household expenditure by region and province of Thailand, 2009-2018, [Online].. https://www.nso.go.th/nsoweb/nso/survey_detail/qC [accessed January 2024].

[33] National Statistical Officer Thailand, Couple with Marriage Certificate by Region and Province: 2004-2013, [Online]. https://www.nso.go.th/public/e-book/Statistical-Yearbook/SYB-2023/96/ (Accessed Date: October 1, 2024).

Contribution of Individual Authors to the Creation of a Scientific Article (Ghostwriting Policy)

- Unchalee Tonggumnead has formulated or evolved overarching research goals and developed the statistics for testing the equality of variance, carrying out the simulation, writing - the original draft, interacting with editors, and editing before publication.
- Nikorn Saengngam applied the proposed method to the actual data, compared the performance of the developed method with other methods, participated in research design and literature review, prepared and reviewed manuscripts, and revised and edited them before publication.

Sources of Funding for Research Presented in a Scientific Article or Scientific Article Itself

No funding was received for conducting this study.

Conflict of Interest

The authors declare that there is no conflict of interests.

APPENDIX

Table 3. Probabilities of rejection when H_0 is true (Type I error) of the test statistics $L_{AB}^{OLS}, L_{AB}^{LAD}, L_{AB}^{S}, L_{SQ}^{OLS}, L_{SQ}^{LAD}$ and L_{SQ}^{S} in the case of testing the equality of variance of three groups ($\sigma_1^2 = \sigma_2^2 = \sigma_3^2$) based on 1,000 simulations, $\alpha = 0.01$

distribution	sample size	Test statistics					
		L_{AB}^{OLS}	L_{AB}^{LAD}	L_{AB}^{S}	L_{SQ}^{OLS}	L_{SQ}^{LAD}	L_{SQ}^{S}
Normal	(15,15,15)	0.011	0.013	0.012	0.008	0.008	0.009
	(30,30,30)	0.012	0.012	0.014	0.009	0.008	0.008
	(60,60,60)	0.010	0.010	0.009	0.009	0.009	0.010
	(10,15,20)	0.014	0.013	0.014	0.008	0.007	0.009
	(25,30,35)	0.013	0.013	0.014	0.011	0.012	0.011
	(50,60,70)	0.007	0.007	0.008	0.008	0.007	0.009
Logistic	(15,15,15)	0.006	0.005	0.007	0.009	0.008	0.010
	(30,30,30)	0.009	0.008	0.009	0.005	0.006	0.006
	(60,60,60)	0.011	0.012	0.011	0.009	0.009	0.010
	(10,15,20)	0.014	0.015	0.013	0.005	0.005	0.006
	(25,30,35)	0.008	0.007	0.008	0.008	0.007	0.008
	(50,60,70)	0.013	0.012	0.011	0.007	0.007	0.007
Lognormal	(15,15,15)	0.219*	0.159*	0.120*	0.004*	0.004*	0.003*
	(30,30,30)	0.204*	0.140*	0.110*	0.001*	0.002*	0.003*
	(60,60,60)	0.227*	0.138*	0.014	0.001*	0.004*	0.004*
	(10,15,20)	0.261*	0.174*	0.125*	0.003*	0.004*	0.004*
	(25,30,35)	0.241*	0.169*	0.120*	0.003*	0.003*	0.003*
	(50,60,70)	0.253*	0.162*	0.012	0.001*	0.002*	0.004*

- At significance level (α =0.01), the test statistics is called robustness when the probability of type I error has to fall between (0.005, 0.015). represents the instances where the probability falls outside the Type I error control rank.*

Table 4. Probabilities of rejection when H_0 is true (Type I error) of the test statistics $L_{AB}^{OLS}, L_{AB}^{LAD}, L_{AB}^{S}, L_{SQ}^{OLS}, L_{SQ}^{LAD}$ and L_{SQ}^{S} in the case of testing the equality of variance of three groups ($\sigma_1^2 = \sigma_2^2 = \sigma_3^2$) based on 1,000 simulations, $\alpha = 0.05$

distribution	sample size	test statistics					
		L_{AB}^{OLS}	L_{AB}^{LAD}	L_{AB}^{S}	L_{SQ}^{OLS}	L_{SQ}^{LAD}	L_{SQ}^{S}
Normal	(15,15,15)	0.053	0.054	0.053	0.053	0.054	0.053
	(30,30,30)	0.049	0.050	0.051	0.036	0.040	0.038
	(60,60,60)	0.052	0.052	0.051	0.045	0.045	0.046
	(10,15,20)	0.057	0.056	0.054	0.065	0.063	0.062
	(25,30,35)	0.055	0.054	0.056	0.059	0.058	0.058
	(50,60,70)	0.058	0.057	0.058	0.067	0.065	0.065
Logistic	(15,15,15)	0.054	0.053	0.054	0.046	0.046	0.045
	(30,30,30)	0.058	0.058	0.055	0.044	0.044	0.045
	(60,60,60)	0.048	0.049	0.050	0.045	0.045	0.045
	(10,15,20)	0.063	0.063	0.062	0.055	0.054	0.055
	(25,30,35)	0.053	0.053	0.052	0.059	0.058	0.056
	(50,60,70)	0.053	0.053	0.052	0.041	0.040	0.041
Lognormal	(15,15,15)	0.634*	0.245*	0.192*	0.019*	0.019*	0.020*
	(30,30,30)	0.593*	0.214*	0.188*	0.009*	0.013*	0.021*
	(60,60,60)	0.544*	0.188*	0.068	0.002*	0.009*	0.014*
	(10,15,20)	0.504*	0.198*	0.154*	0.017*	0.020*	0.024*
	(25,30,35)	0.564*	0.203*	0.183*	0.014*	0.018*	0.020*
	(50,60,70)	0.551*	0.190*	0.073	0.008*	0.012*	0.015*

- At significance level (α =0.05), the test statistics is called robustness when the probability of type I error has to fall between (0.025, 0.075). represents the instances where the probability falls outside the Type I error control rank.*

Table 5. Probabilities of rejection when H_0 is not true (power of the test) of the test statistics $L_{AB}^{OLS}, L_{AB}^{LAD}, L_{AB}^{S}, L_{SQ}^{OLS}, L_{SQ}^{LAD}$ and L_{SQ}^{S} in case the differences between the ratios of variances are low (1:2:3), based on 1,000 simulations, $\alpha = 0.01$

distribution	sample size	test statistics					
		L_{AB}^{OLS}	L_{AB}^{LAD}	L_{AB}^{S}	L_{SQ}^{OLS}	L_{SQ}^{LAD}	L_{SQ}^{S}
Normal	(15,15,15)	0.142	0.145	0.152	0.096	0.100	0.098
	(30,30,30)	0.398	0.394	0.400	0.329	0.330	0.331
	(60,60,60)	0.825	0.823	0.830	0.814	0.814	0.812
	(10,15,20)	0.088	0.090	0.089	0.041	0.040	0.051
	(25,30,35)	0.325	0.324	0.330	0.229	0.228	0.231
	(50,60,70)	0.787	0.787	0.790	0.730	0.732	0.740
Logistic	(15,15,15)	0.121	0.121	0.119	0.045	0.048	0.054
	(30,30,30)	0.290	0.303	0.298	0.181	0.180	0.187
	(60,60,60)	0.696	0.710	0.707	0.525	0.520	0.528
	(10,15,20)	0.051	0.048	0.059	0.018	0.030	0.024
	(25,30,35)	0.258	0.260	0.271	0.126	0.124	0.132
	(50,60,70)	0.667	0.658	0.680	0.426	0.428	0.435
Lognormal	(15,15,15)	0.265	0.328	0.514	0.004	0.008	0.010
	(30,30,30)	0.320	0.450	0.510	0.004	0.009	0.014
	(60,60,60)	0.466	0.487	0.530	0.003	0.007	0.011
	(10,15,20)	0.168	0.248	0.497	0.001	0.003	0.009
	(25,30,35)	0.247	0.304	0.499	0.001	0.003	0.010
	(50,60,70)	0.335	0.405	0.510	0.001	0.003	0.014

Table 6. Probabilities of rejection when H_0 is not true (power of the test) of the test statistics $L_{AB}^{OLS}, L_{AB}^{LAD}, L_{AB}^{S}, L_{SQ}^{OLS}, L_{SQ}^{LAD}$ and L_{SQ}^{S} in case the differences between the ratios of variances are low (1:2:3), based on 1,000 simulations, $\alpha = 0.05$

distribution	sample size	test statistics					
		L_{AB}^{OLS}	L_{AB}^{LAD}	L_{AB}^{S}	L_{SQ}^{OLS}	L_{SQ}^{LAD}	L_{SQ}^{S}
Normal	(15,15,15)	0.362	0.382	0.400	0.278	0.272	0.280
	(30,30,30)	0.665	0.664	0.680	0.623	0.613	0.630
	(60,60,60)	0.942	0.940	0.938	0.968	0.968	0.972
	(10,15,20)	0.280	0.284	0.282	0.174	0.180	0.188
	(25,30,35)	0.637	0.640	0.640	0.561	0.560	0.565
	(50,60,70)	0.944	0.940	0.949	0.934	0.930	0.935
Logistic	(15,15,15)	0.283	0.281	0.291	0.223	0.232	0.230
	(30,30,30)	0.580	0.574	0.582	0.444	0.484	0.480
	(60,60,60)	0.893	0.890	0.900	0.809	0.814	0.820
	(10,15,20)	0.217	0.218	0.224	0.112	0.118	0.210
	(25,30,35)	0.517	0.510	0.521	0.393	0.400	0.403
	(50,60,70)	0.862	0.860	0.872	0.746	0.750	0.750
Lognormal	(15,15,15)	0.830	0.854	0.884	0.020	0.031	0.024
	(30,30,30)	0.839	0.860	0.884	0.014	0.020	0.020
	(60,60,60)	0.903	0.921	0.928	0.021	0.024	0.025
	(10,15,20)	0.412	0.430	0.480	0.005	0.009	0.014
	(25,30,35)	0.755	0.780	0.800	0.010	0.012	0.014
	(50,60,70)	0.818	0.834	0.848	0.005	0.010	0.017

Table 7. Probabilities of rejection when H_0 is not true (power of the test) of the test statistics $L_{AB}^{OLS}, L_{AB}^{LAD}, L_{AB}^{S}, L_{SQ}^{OLS}, L_{SQ}^{LAD}$ and L_{SQ}^{S} in case the differences between the ratios of variances are moderate (1:3:5), based on 1,000 simulations, $\alpha = 0.01$

distribution	sample size	test statistics					
		L_{AB}^{OLS}	L_{AB}^{LAD}	L_{AB}^{S}	L_{SQ}^{OLS}	L_{SQ}^{LAD}	L_{SQ}^{S}
Normal	(15,15,15)	0.345	0.343	0.350	0.203	0.200	0.212
	(30,30,30)	0.811	0.814	0.821	0.667	0.670	0.668
	(60,60,60)	0.995	0.994	0.997	0.996	0.995	0.996
	(10,15,20)	0.190	0.241	0.288	0.074	0.088	0.101
	(25,30,35)	0.725	0.724	0.730	0.553	0.552	0.571
	(50,60,70)	0.994	0.995	0.995	0.978	0.978	0.981
Logistic	(15,15,15)	0.223	0.222	0.230	0.111	0.118	0.200
	(30,30,30)	0.691	0.690	0.700	0.413	0.430	0.428
	(60,60,60)	0.986	0.990	0.990	0.866	0.868	0.873
	(10,15,20)	0.132	0.154	0.172	0.041	0.040	0.051
	(25,30,35)	0.594	0.600	0.614	0.285	0.287	0.293
	(50,60,70)	0.968	0.974	0.973	0.810	0.810	0.818
Lognormal	(15,15,15)	0.256	0.295	0.334	0.003	0.003	0.005
	(30,30,30)	0.368	0.400	0.412	0.001	0.004	0.005
	(60,60,60)	0.421	0.479	0.501	0.001	0.004	0.006
	(10,15,20)	0.139	0.198	0.243	0.001	0.002	0.003
	(25,30,35)	0.228	0.294	0.354	0.000	0.000	0.001
	(50,60,70)	0.330	0.387	0.413	0.000	0.001	0.002

Table 8. Probabilities of rejection when H_0 is not true (power of the test) of the test statistics $L_{AB}^{OLS}, L_{AB}^{LAD}, L_{AB}^{S}, L_{SQ}^{OLS}, L_{SQ}^{LAD}$ and L_{SQ}^{S} in case the differences between the ratios of variances are moderate (1:3:5), based on 1,000 simulations, $\alpha = 0.05$

distribution	sample size	test statistics					
		L_{AB}^{OLS}	L_{AB}^{LAD}	L_{AB}^{S}	L_{SQ}^{OLS}	L_{SQ}^{LAD}	L_{SQ}^{S}
Normal	(15,15,15)	0.660	0.663	0.670	0.523	0.520	0.500
	(30,30,30)	0.948	0.946	0.951	0.930	0.922	0.932
	(60,60,60)	0.999	0.998	0.999	1.000	0.998	0.999
	(10,15,20)	0.497	0.502	0.500	0.294	0.300	0.298
	(25,30,35)	0.932	0.940	0.938	0.823	0.828	0.830
	(50,60,70)	0.999	0.999	1.000	1.000	0.998	0.999
Logistic	(15,15,15)	0.566	0.570	0.564	0.373	0.360	0.378
	(30,30,30)	0.872	0.870	0.878	0.740	0.740	0.747
	(60,60,60)	0.998	0.998	0.999	0.984	0.980	0.982
	(10,15,20)	0.404	0.400	0.410	0.203	0.213	0.210
	(25,30,35)	0.870	0.870	0.880	0.660	0.660	0.668
	(50,60,70)	0.996	0.996	0.997	0.967	0.970	0.974
Lognormal	(15,15,15)	0.883	0.890	0.888	0.009	0.009	0.008
	(30,30,30)	0.915	0.915	0.920	0.006	0.006	0.008
	(60,60,60)	0.948	0.950	0.956	0.009	0.009	0.010
	(10,15,20)	0.373	0.384	0.400	0.002	0.003	0.002
	(25,30,35)	0.830	0.834	0.840	0.004	0.004	0.005
	(50,60,70)	0.872	0.878	0.880	0.002	0.002	0.001

Table 9. Probabilities of rejection when H_0 is not true (power of the test) of the test statistics $L_{AB}^{OLS}, L_{AB}^{LAD}, L_{AB}^{S}, L_{SQ}^{OLS}, L_{SQ}^{LAD}$ and L_{SQ}^{S} in case the differences between the ratios of variances are high (1:5:10) based on 1,000 simulations, $\alpha = 0.01$

distribution	sample size	test statistics					
		L_{AB}^{OLS}	L_{AB}^{LAD}	L_{AB}^{S}	L_{SQ}^{OLS}	L_{SQ}^{LAD}	L_{SQ}^{S}
Normal	(15,15,15)	0.664	0.660	0.670	0.369	0.358	0.363
	(30,30,30)	0.990	0.990	0.992	0.927	0.928	0.928
	(60,60,60)	1.000	1.000	1.000	1.000	0.997	0.995
	(10,15,20)	0.448	0.444	0.450	0.170	0.168	0.174
	(25,30,35)	0.981	0.980	0.988	0.845	0.840	0.850
	(50,60,70)	1.000	1.000	0.999	1.000	0.998	0.998
Logistic	(15,15,15)	0.548	0.545	0.554	0.238	0.234	0.241
	(30,30,30)	0.971	0.972	0.979	0.696	0.700	0.698
	(60,60,60)	1.000	0.999	1.000	0.987	0.988	0.985
	(10,15,20)	0.324	0.320	0.321	0.089	0.088	0.100
	(25,30,35)	0.946	0.949	0.947	0.575	0.575	0.581
	(50,60,70)	1.000	0.999	1.000	0.958	0.954	0.955
Lognormal	(15,15,15)	0.236	0.240	0.261	0.002	0.002	0.002
	(30,30,30)	0.299	0.310	0.358	0.001	0.002	0.001
	(60,60,60)	0.382	0.399	0.423	0.002	0.002	0.003
	(10,15,20)	0.121	0.220	0.318	0.000	0.000	0.001
	(25,30,35)	0.187	0.258	0.342	0.000	0.000	0.000
	(50,60,70)	0.253	0.283	0.310	0.000	0.000	0.001

Table 10. Probabilities of rejection when H_0 is not true (power of the test) of the test statistics $L_{AB}^{OLS}, L_{AB}^{LAD}, L_{AB}^{S}, L_{SQ}^{OLS}, L_{SQ}^{LAD}$ and L_{SQ}^{S} in case the differences between the ratios of variances are high (1:5:10) based on 1,000 simulations, $\alpha = 0.05$

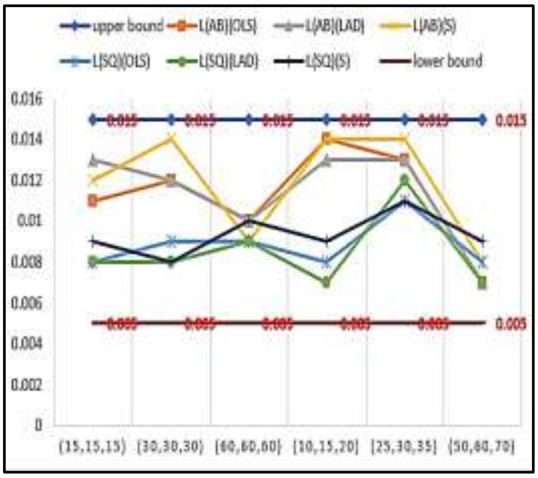
distribution	sample size	test statistics					
		L_{AB}^{OLS}	L_{AB}^{LAD}	L_{AB}^{S}	L_{SQ}^{OLS}	L_{SQ}^{LAD}	L_{SQ}^{S}
Normal	(15,15,15)	0.910	0.908	0.912	0.760	0.771	0.768
	(30,30,30)	1.000	0.998	1.000	0.990	0.994	0.992
	(60,60,60)	1.000	0.999	1.000	1.000	0.999	0.998
	(10,15,20)	0.779	0.784	0.812	0.547	0.550	0.552
	(25,30,35)	0.999	0.999	0.997	0.989	0.979	0.988
	(50,60,70)	1.000	0.997	1.000	1.000	0.998	0.999
Logistic	(15,15,15)	0.842	0.848	0.844	0.583	0.600	0.588
	(30,30,30)	0.997	0.997	0.999	0.937	0.940	0.938
	(60,60,60)	1.000	0.998	1.000	0.999	0.999	0.998
	(10,15,20)	0.711	0.718	0.717	0.325	0.320	0.330
	(25,30,35)	0.994	0.995	0.995	0.892	0.890	0.910
	(50,60,70)	1.000	0.999	1.000	0.996	0.994	0.995
Lognormal	(15,15,15)	0.928	0.930	0.929	0.011	0.011	0.019
	(30,30,30)	0.943	0.948	0.950	0.005	0.005	0.007
	(60,60,60)	0.962	0.962	0.971	0.012	0.012	0.015
	(10,15,20)	0.346	0.350	0.400	0.001	0.001	0.001
	(25,30,35)	0.894	0.898	0.898	0.004	0.003	0.005
	(50,60,70)	0.924	0.930	0.931	0.000	0.000	0.001

Table 11. The means, standard deviations, and distribution were derived from the Anderson-Darling test.

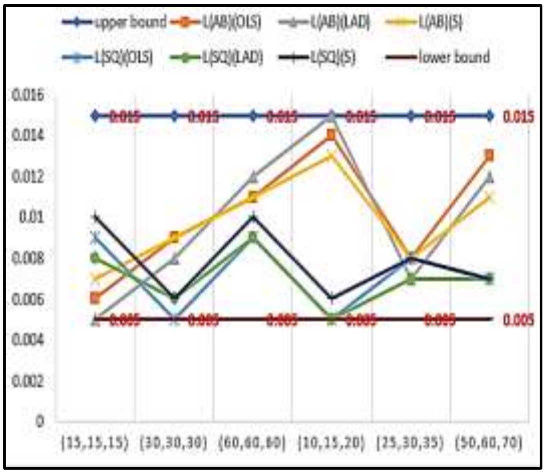
No.	data	$\overline{X}$	$S.D.$	distribution of data	AD test	P-value
1	The average household expenditure in 17 Northern Thai provinces					
	- the year 2009	11,693	2,461	Normal	0.417	0.293
				Logistic	0.262	>0.250
				Lognormal	0.835	0.025
	- the year 2010	12,398	2,173	Normal	0.444	0.251
				Logistic	0.246	>0.250
				Lognormal	0.893	0.017
	- the year 2011	13,290	2,190	Normal	0.530	0.150
				Logistic	0.276	>0.250
				Lognormal	0.982	0.010
2	The average marriage registration in 75 provinces across Thailand					
	- years 2009	3,616	2,417	Normal	2.389	<0.005
				Logistic	1.756	<0.005
				Lognormal	0.248	0.742
	- year 2010	3,429	2,220	Normal	2.237	<0.005
				Logistic	1.568	<0.005
				Lognormal	0.167	0.935
	- year 2011	3,284	2,158	Normal	2.384	<0.005
				Logistic	1.771	<0.005
				Lognormal	0.202	0.876

Table 12. The result of testing the equality of variance of two actual datasets each comprising three subsets of data from the test statistics $L_{AB}^{OLS}, L_{AB}^{LAD}, L_{AB}^{S}, L_{SQ}^{OLS}, L_{SQ}^{LAD}$ and L_{SQ}^{S} .

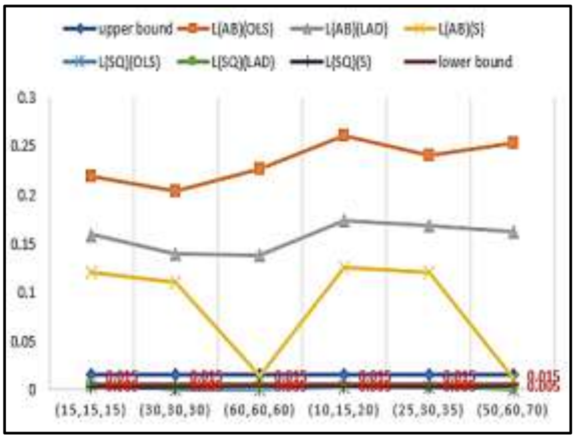
data	test statistics value (p)					
	L_{AB}^{OLS} (p)	L_{AB}^{LAD} (p)	L_{AB}^{S} (p)	L_{SQ}^{OLS} (p)	L_{SQ}^{LAD} (p)	L_{SQ}^{S} (p)
1	0.089 (0.915)	0.083 (0.921)	0.073 (0.930)	0.101 (0.905)	0.102 (0.904)	0.095 (0.910)
2	0.256 (0.775)	0.220 (0.803)	0.126 (0.882)	0.778 (0.463)	0.754 (0.474)	0.741 (0.480)



(a)

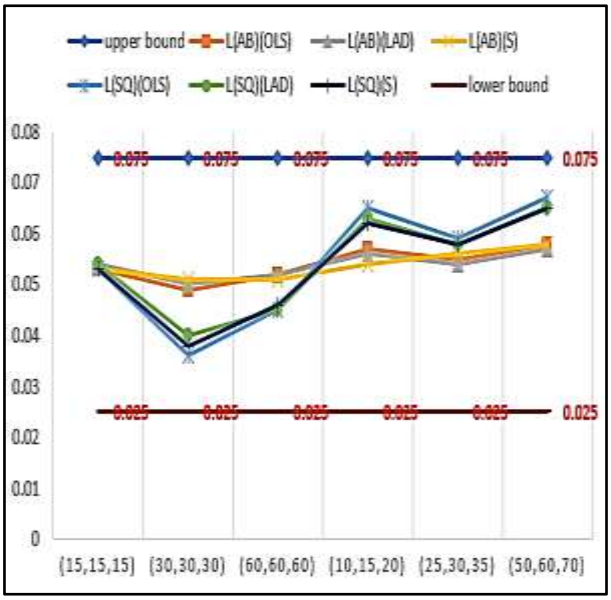


(b)

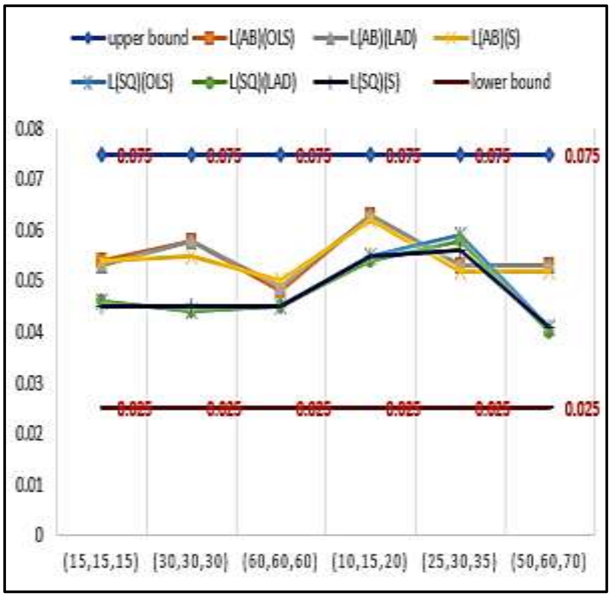


(c)

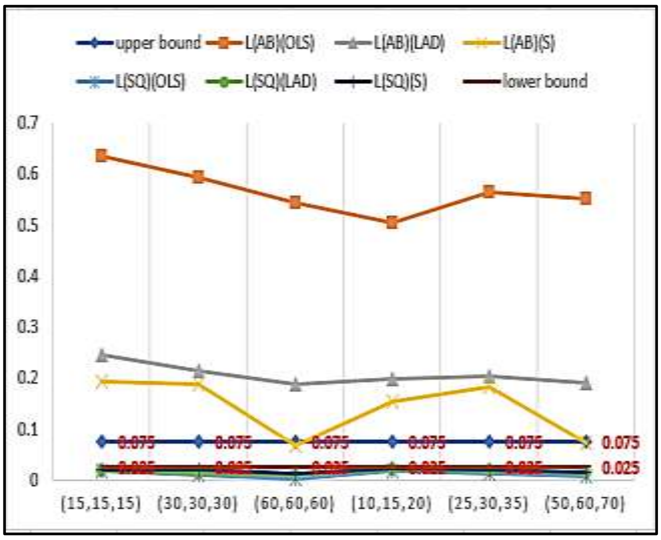
Fig. 4: Illustrate the ability of the test statistics $L_{AB}^{OLS}, L_{AB}^{LAD}, L_{AB}^{S}, L_{SQ}^{OLS}, L_{SQ}^{LAD}$ and L_{SQ}^{S} to control Type I error for testing the equality of the variance of three datasets ; (a) three datasets are normal distribution, $\alpha = 0.01$. (b) three datasets are logistic distribution, $\alpha = 0.01$. (c) three datasets are lognormal distribution, $\alpha = 0.01$



(a)

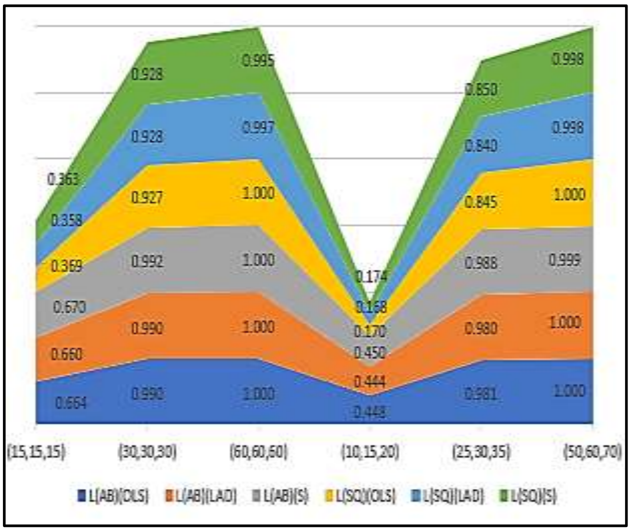


(b)

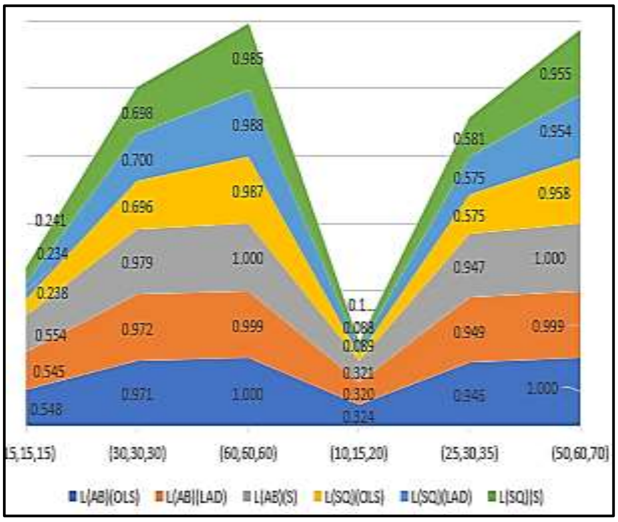


(c)

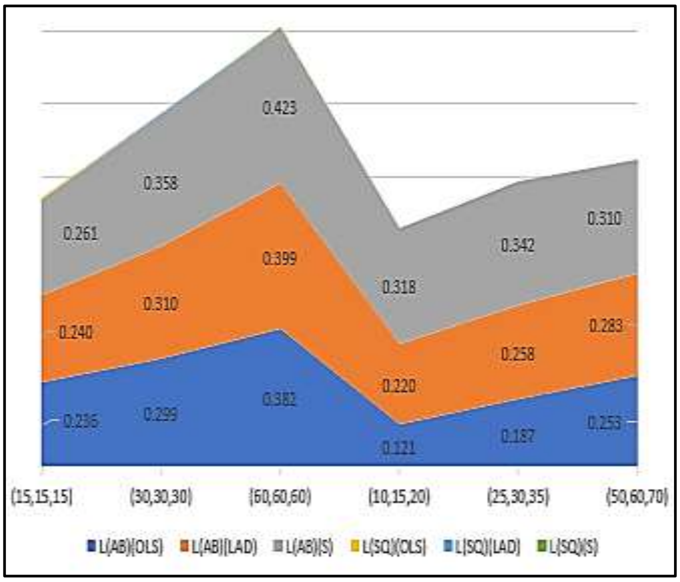
Fig. 5: Illustrate the ability of the test statistics $L_{AB}^{OLS}, L_{AB}^{LAD}, L_{AB}^{S}, L_{SQ}^{OLS}, L_{SQ}^{LAD}$ and L_{SQ}^{S} to control Type I error for testing the equality of the variance of three datasets ; (a) three datasets are normal distribution, $\alpha = 0.05$. (b) three datasets are logistic distribution, $\alpha = 0.05$. (c) three datasets are lognormal distribution, $\alpha = 0.05$



(a)

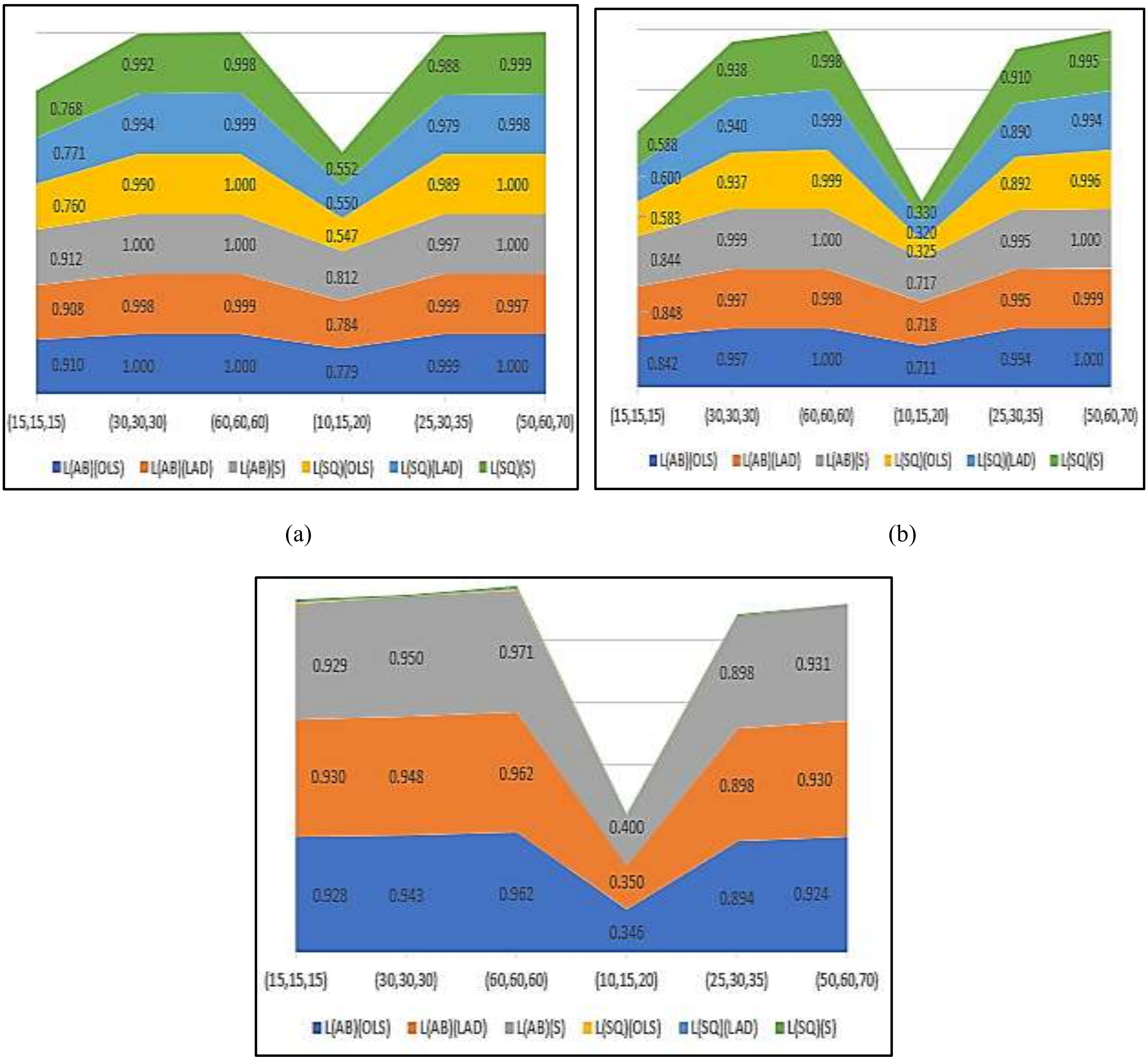


(b)



(c)

Fig. 6: (a) Area plot of power of the test of the test statistics $L_{AB}^{OLS}, L_{AB}^{LAD}, L_{AB}^{S}, L_{SQ}^{OLS}, L_{SQ}^{LAD}$ and L_{SQ}^{S} in the case of normal distribution, the differences between the ratios of variances are highly $(1:5:10)$, and $\alpha = 0.01$. (b) Area plot of power of the test of the test statistics $L_{AB}^{OLS}, L_{AB}^{LAD}, L_{AB}^{S}, L_{SQ}^{OLS}, L_{SQ}^{LAD}$ and L_{SQ}^{S} in the case of logistic distribution, the differences between the ratios of variances are highly $(1:5:10)$, and $\alpha = 0.01$. (c) Area plot of power of the test of the test statistics $L_{AB}^{OLS}, L_{AB}^{LAD}, L_{AB}^{S}, L_{SQ}^{OLS}, L_{SQ}^{LAD}$ and L_{SQ}^{S} in the case of the lognormal distribution, the differences between the ratios of variances are highly $(1:5:10)$, and $\alpha = 0.01$



(c)

Fig. 7: (a) Area plot of power of the test of the test statistics $L_{AB}^{OLS}, L_{AB}^{LAD}, L_{AB}^{S}, L_{SQ}^{OLS}, L_{SQ}^{LAD}$ and L_{SQ}^{S} in the case of normal distribution, the differences between the ratios of variances are highly $(1{:}5{:}10)$, and $\alpha = 0.05$. (b) Area plot of power of the test of the test statistics $L_{AB}^{OLS}, L_{AB}^{LAD}, L_{AB}^{S}, L_{SQ}^{OLS}, L_{SQ}^{LAD}$ and L_{SQ}^{S} in the case of logistic distribution, the differences between the ratios of variances are highly $(1{:}5{:}10)$, and $\alpha = 0.05$. (c) Area plot of power of the test of the test statistics $L_{AB}^{OLS}, L_{AB}^{LAD}, L_{AB}^{S}, L_{SQ}^{OLS}, L_{SQ}^{LAD}$ and L_{SQ}^{S} in the case of the lognormal distribution, the differences between the ratios of variances are highly $(1{:}5{:}10)$, and $\alpha = 0.05$

DOI: 10.37394/23206.2025.24.12

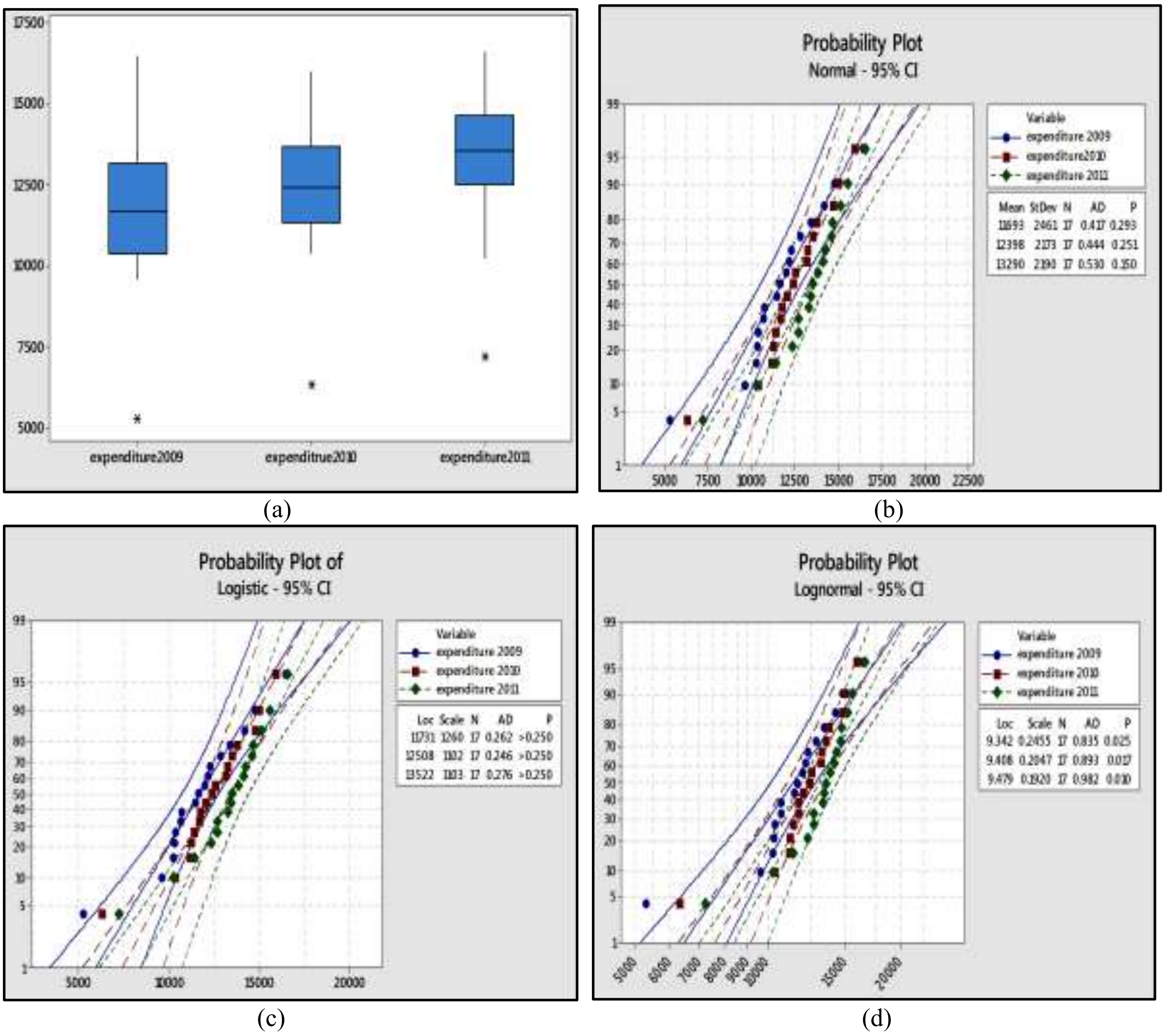


Fig. 8: (a) Box Plot of the average household expenditure in 17 Northern Thai provinces covering the years 2009, 2010, and 2011. (b) Normal probability plot of the average household expenditure in 17 Northern Thai provinces covering the years 2009, 2010, and 2011, and distribution of the three data sets derived from the Anderson-Darling test. (c) Logistic probability plot of the average household expenditure in 17 Northern Thai provinces covering the years 2009, 2010, and 2011, and distribution of the three data sets derived from the Anderson-Darling test. (d))Lognormal probability plot of the average household expenditure in 17 Northern Thai provinces covering the years 2009, 2010, and 2011, and distribution of the three data sets derived from the Anderson-Darling test

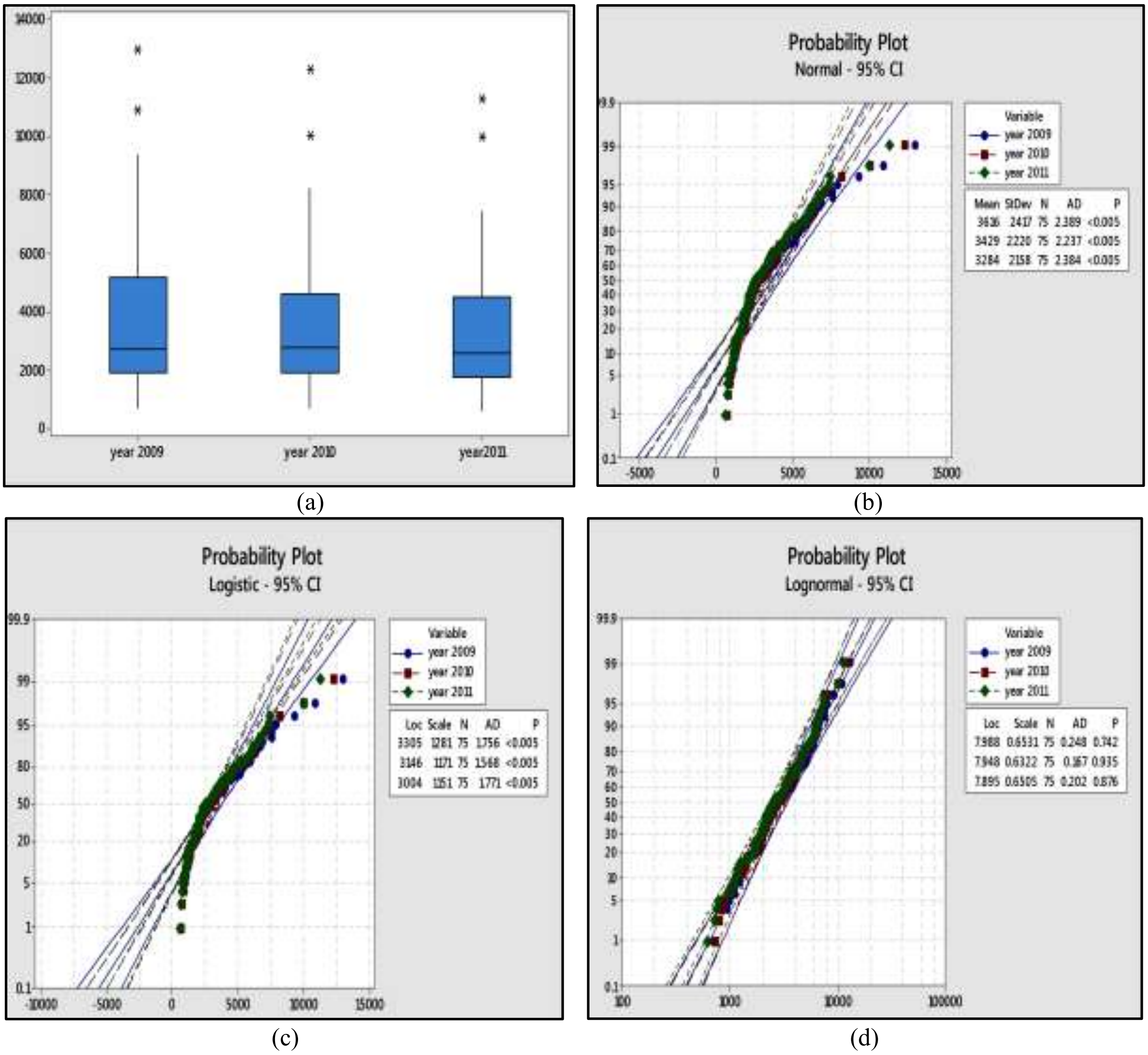


Fig. 9: (a) Box Plot The average marriage registration in 75 provinces across Thailand covering the years 2009, 2010, and 2011. (b) Normal probability plot of the average marriage registration in 75 provinces across Thailand covering the years 2009, 2010, and 2011, and distribution of the three data sets derived from the Anderson-Darling test. (c) Logistic probability plot of the average marriage registration in 75 provinces across Thailand covering the years 2009, 2010, and 2011, and distribution of the three data sets derived from the Anderson-Darling test. (d)) Lognormal probability plot of the average marriage registration in 75 provinces across Thailand covering the years 2009, 2010, and 2011, and distribution of the three data sets derived from the Anderson-Darling test