Determining the Relatively Efficient Test Statistic Measure between the Traditional Chow test and the Milek Permutation Test for detecting Structural break in Linear Models

Aronu, C. O.[1], and Nworuh, G. E.[2]

[1]*Department of Statistics, Chukwuemeka Odumegwu Ojukwu University, Anambra State, Nigeria*

[2]*Department of Statistics, Federal University of Science and Technology (FUTO), Owerri, Imo State, Nigeria*

***Abstract*:-This study examines the performance of the traditional Chow test and the Milek permutation test for detecting structural break based on relative efficiency of their test statistic value. The aim of this study was to determine the method that performs best in terms of relative efficiency of the test values for the Standard normal distribution, Gamma distribution, and the Exponential distribution. The Milek permutation method for structural break was found to have a better relative efficiency of the test statistic value than the Chow test for the standard normal distribution, Gamma distribution, and the Exponential distribution.**

***Keywords:* Chow Test, Exponential Distribution, Structural Break, Permutation method**

I. INTRODUCTION

Efficiency of a test statistic is the measure of the quality of a particular estimator of a hypothesis testing method. A more efficient test statistic requires fewer observations than a less efficient test statistic in other to obtain a given performance.

The relative efficiency of two test procedures is the ratio of their efficiencies especially in a situation where comparison is made between a given procedure and a best possible procedure. The efficiency of the test procedures often depends on the sample size available for the given procedure. But some researchers employ the asymptotic relative efficiency which is the limit of the relative efficiencies as the sample size increases. This study examines the relative efficiency of the traditional Chow test and the Milek (2015) permutation method for the Standard normal distribution, Gamma distribution and Exponential distribution.

II. LITERATURE REVIEW

Fay and Proschan (2010) opined that most times researcher may wish to assess whether there exist significant difference between two groups using either the t-test or the Wilcoxon-Mann-Whitney (WMW) test. Although both the t-tests and the Wilcoxon-Mann-Whitney (WMW)tests are usually related with some degree of variation in the hypotheses, the decision rule and p-value from either test could be associated with many different sets of assumptions, which can be referred to as perspectives. It is important to have many of the different perspectives to which a decision rule may be applied collected in one place, since each perspective allows a different interpretation of the associated p-value.

According to Lebanon (2006), the relative efficiency of two unbiased estimators can be measures by the ratio of their variances while the quality of two estimators can be compared by examining the ratio of their Mean Square Error (MSE) (Nwakuya and Nwabueze, 2016).

Wang *et al*. (2012) examined two test statistics which can used for testing the Poisson distribution against the zero-inflated Poisson distributions. They noted that the two test statistic are asymptotically equivalent under null hypothesis with relative efficiency equal to 1. They observed that the two test has significantly different behaviors for small and medium sample sizes. Findings of the study showed that T_1 has a reasonable empirical size (under null hypothesis) and power (under alternative hypothesis) for small and medium sample sizes while T_2 showed some erratic behaviors even for medium sample sizes which may lead to misleading inference in practical situations.

Umeh and Eriobu (2016) examined the relative efficiency and sensitivity of four test statistic. The findings of their study showed that the median test was relatively more efficient than the modified median test for both symmetric and asymmetric distributions. Also, they added that for asymmetric distribution, the Modified Mann- Whitney U test was found to be more relatively efficient than Mann-Whitney U test (MMWU), while for the symmetric distribution, the Mann-Whitney U test (MMWU) was found to be more relatively efficient than Modified Mann-Whitney at sample size of 5 while the Modified Mann-Whitney U test (MMWU) was more relatively efficient than the Mann-Whitney for other sample sizes considered in their study.

Nwakuya and Nwabueze (2016) examined the relative efficiency of estimates in Multiple Imputation analysis with regards to percentages of missing data using 3 different imputation numbers; 7, 5 and 3 on four different simulated data sets with 50%, 45%, 25% and 10% missing values. They calculated the variance of each of their data set with different percentages of missing value for each imputation number. The findings of their study showed that their proposed method yield's a lower variances compared to an existing method. Also, further findings of their study showed that when the missingness was 50% the estimates from data set gotten from imputation number 7 was most efficient when compared to estimates from data sets gotten from imputation numbers 5 and 3. Similarly, when the missingness was set at 10% and 25% the estimates from data set gotten from imputation number 5 were found to be most efficient followed by estimates from data sets gotten from imputation number 7 and then 3. The relative efficiency for 40% missingness compared among the 3 imputation numbers showed that estimates from imputation number 3were most efficient.

III. MATERIAL AND METHODS

3.1 Method of Data Collection

The source of data used for this study is simulated data from standard normal distribution, Gamma distribution and Exponential distribution for sample size 15, 20, 25, 30, 40, 50, 60, 70, 80, 90 and 100.

3.2 Chow Test

The Chow test is often used to determine whether there exist different subgroups in a population of interest. The single/full model of a Chow test is written as:

$$Y_t = X_t\ \beta + \varepsilon_t \qquad (1)$$

Where,

Y_t is a random variable called the response or dependent variable.

β represents constants or parameters whose exact value are not known and thus must be estimated from the experimental data.

X_t represents the mathematical variable called regressor or covariate or predictor independent non-random variable whose value are controlled or at least accurately observed by the experimenter t_b represents the break point.

To estimate the regression parameters β properly using the least-squares estimation, the assumption that n > p holds and X is of full rank. Here n is the number of observation while p is the number of regression coefficients.

The null hypothesis, tested by Chow, states that two disjoint models with the sum of squares residual is:

$$Y_{1t} = X_{1t}\ \beta_1 + \varepsilon_{1t} \qquad (2)$$

$$Y_{2t} = X_{2t}\ \beta_2 + \varepsilon_{2t} \qquad (3)$$

Y_{1t} and Y_{2t} represents the random variable called the response or dependent variable for the first group and second group respectively.

β_1 and β_2 represents constants or parameters whose exact value are not known and thus must be estimated from the experimental data for the first group and second group respectively.

X_{1t} and X_{2t} represents the mathematical variable called regressor or covariate or predictor independent non-random variable whose value are controlled or at least accurately observed by the experimenter for the first group and second group respectively.

This suggest that model (2) applies before the break at time t, while model (3) applies after the structural break.

The Chow test basically tests whether the single regression line or the two separate regression lines fit the data best (Chow, 1960).

Taking advantage of the various F-test (Mood, 1950; Davis, 1952), to test for presence of structural break in a given set of data, a special and useful application of the F test procedure is found in the

The Chow test statistic follows the F-distribution with $n_1 + n_2 - 2k$ degree of freedom. $n_1$is the number of observations before structural break and n_2 is the number of observations after the break point (Mood, 1950).

The Chow test statistic is given as

$$F = \frac{(RSS_T - (RSS_1 + RSS_2))/k}{(RSS_1 + RSS_2)/(n_1 + n_2 - 2k)} \text{ this follows the } F(k, n_1+n_2-2k) \qquad (4)$$

where,

RSS_T represents the residual sum of squares for the full model

RSS_1 represents the residual sum of squares for the first sub sample or first reduced model

RSS_2 represents the residual of the second sub sample or second reduced model,

k is the number of parameters,

n_1 and n_2 represents the length of the two subsamples.

Decision Rule

H_0 is to reject at the significance level α if

$$F \geq F(k, n_1+n_2-2k)$$

The other criterion equivalent to the decision rule above is to compare the p-value for F-statistics with α and reject H_0 if

$$\Pr(F) \leq \alpha$$

It is important to note that a parametric test of significance of Chow test can be carried out using an F-statistic under the assumption of normality. If this condition is not met, a permutation method becomes an alternative to perform the test. Under normality, one expects a permutation test to produce approximately the same results as the parametric F-test.

3.3 Milek (2015) Permutation Method for Structural Break

Miłek (2015) proposed a permutation method for structural break; the study considered the reduced models as

$$y_{1t} = a_1 + b_1 t_1 + \varepsilon_{1t} \quad (5)$$

$$y_{2t} = a_2 + b_2 t_2 + \varepsilon_{2t} \quad (6)$$

where, y_{1t} and y_{2t} : represents the random vectors called the response or dependent vector for the first group and second group respectively,

a_1 and a_2: represents the intercept for the first group and second group respectively,

b_1 and b_2: represents the slope for the first group and second group respectively,

ε_{1t} and ε_{2t} : represents the random error for the first group and second group respectively, and

t_1 and t_2: represents the independent variable for the first group and second group respectively.

The test statistic was given as:

$$T = |b_2 - b_1| + \left|\frac{a_2}{a_1}\right| \quad (7)$$

where, $a_1 = \dfrac{\sum_{t=1}^{k}(t-\bar{t}_{1t})(y_t-\bar{y}_{1t})}{\sum_{t=1}^{k}(t-\bar{t}_{1t})^2}$,

$a_2 = \dfrac{\sum_{t=k+1}^{n}(t-\bar{t}_{2t})(y_t-\bar{y}_{2t})}{\sum_{t=k+1}^{k}(t-\bar{t}_{2t})^2}$, $b_1 = \bar{y}_{1t} - a_1\bar{t}_{1t}$,

$b_2 = \bar{y}_{2t} - a_2\bar{t}_{2t}$, $\bar{y}_{1t} = \dfrac{\sum_{t=1}^{k} y_t}{k}$, $\bar{y}_{2t} = \dfrac{\sum_{t=k+1}^{n} y_t}{n-k}$,

$\bar{t}_{1t} = \dfrac{\sum_{t=1}^{k} t}{k}$, and $\bar{t}_{2t} = \dfrac{\sum_{t=k+1}^{n} t}{n-k}$

3.4 Relative Efficiency

The relative efficiency of the test value is the ratio of their precision (inverse of variance or standard deviation) (Nikulin, 2001). This implies that the efficiency of the first test to the second would be the variance/standard deviation of the second divided by the variance of the first. Hence, the method with the least variance/standard deviation of the test value is considered less efficient.

Expressed mathematically as:

$$e(T_1, T_2) = \frac{E(T_2-\theta_2)^2}{E(T_1-\theta_1)^2} \quad (8)$$

where T_1 and T_2 are the test values while θ_1 and θ_2 are the respective means.

IV. DATA ANALYSIS AND RESULT

This section presents the summary result of the test statistic value or reference value for the Chow test and the Milek permutation method.

Table 1: Summary of Test value for Chow test methods across the sample sizes for Standard Normal Distribution, Gamma Distribution and Exponential Distribution

Distribution	Test		Sample 15	Sample 20	Sample 25	Sample 30	Sample 40	Sample 50	Sample 60	Sample 70	Sample 80	Sample 90	Sample 100
Standard Normal	Chow	Mean	211.65	235.03	190.63	674.80	340.91	431.69	849.04	907.84	1576.83	1236.72	211.65
		SD	651.28	822.92	571.93	3362.72	1044.51	1092.98	4311.17	1820.96	3174.02	2485.84	651.28
	Milek	Mean	68.09	67.24	66.90	67.46	148.616	440.05	456.11	1288.79	456.99	385.21	404.34
		SD	124.70	124.81	124.97	124.71	555.55	705.73	698.815	5820.31	661.0644	631.3	626.38
Gamma	Chow	Mean	4455.17	1588.32	2329.64	1506.20	589.86	1288.98	925.73	290.96	1057.52	503.94	32.47
		SD	3458.12	1922.29	2500.69	2159.41	1218.27	2078.21	1634.40	940.71	2150.92	1585.66	74.79
	Milek	Mean	4.896	17.6	17.5	19.41	51.8	424	111.658	190.89	232.2	1680.9186	505.095
		SD	3.35	27.6	13.4	22.55	105	1250	168.568	337.94	415.7	6184.449	1062.66
Exponential	Chow	Mean	3869.24	201.37	1758.19	1279.09	416.85	857.20	300.02	131.091	790.39	84.893	116.72
		SD	3397.16	853.99	2377.49	2115.67	1051.31	1724.6	1014.8	636.31	1926.2	535.81	630.03
	Milek	Mean	5.90	22.503	18.03	21.27	60.17	722.18	147.325	252.21	257.39	1810.744	458.729
		SD	3.80	27.53	10.06	22.23	122.34	1586.572	167.44	250.86	391.63	6161.55	977.12
All Distribution	Chow	Mean	8536.06	2024.72	4278.46	3460.09	1347.62	2577.87	2074.79	1329.891	3424.74	1825.553	360.84
		SD	7506.56	3599.2	5450.11	7637.8	3314.09	4895.79	6960.37	3397.98	7251.14	4607.31	1356.1
	Milek	Mean	78.886	107.343	102.43	108.14	260.586	1586.23	715.093	1731.89	946.58	3876.873	1368.164
		SD	131.85	179.94	148.43	169.49	782.89	3542.302	1034.823	6409.11	1468.394	12977.3	2666.16

V. CONCLUSION

This study compares the performance of the traditional Chow test and the Milek permutation test based on relative efficiency of their test statistic value. The two test are used for detecting structural break in linear models. It was found that the Milek permutation method for structural break has better relative efficiency test statistic value than the Chow test for the standard normal distribution since the test values recorded a smaller standard deviation of 10198.34 over the Chow test with a standard deviation value of 19989.61. Findings showed that the Milek permutation method for structural break has better relative efficiency test statistic value efficiency than the Chow test for the Gamma distribution since the test values recorded a lesser standard deviation of 9591.22 over the Chow test with a standard deviation value of 19723.47. Similarly, the Milek permutation method for structural break was found to have better relative efficiency than the Chow test for the Exponential distribution since the test values recorded a lesser standard deviation of 9721.13 over the Chow test with a standard deviation value of 16263.37. Further result showed that the Milek permutation method for structural break has better relative efficient test statistic value across the distributions since it recorded the least standard deviation of 29510.69 against the Chow test with standard deviation of 55976.45. Hence, we conclude that the Milek permutation method is more relatively efficient than the Chow test based on their test statistic values. In view of the findings, it is recommended that users who have interest in relative efficiency of the test value should employ the Milek permutation method over the traditional Chow test for detecting structural break in linear models.

REFERENCES

[1]. Chow, G. C. (1960). Tests of Equality between Sets of Coefficients in Two Linear Regressions. *Econometrica*, 28.3: 591–605.

[2]. Fay, M. P. and Proschan, M. A. (2010). Wilcoxon-Mann-Whitney or t-test? On assumptions for hypothesis tests and multiple interpretations of decision rules. *Stat Surv*., 4: 1–39

[3]. Miłek, M. (2015). Detecting structural changes in the time series in the use of simulation methods. The 9th Professor Aleksander Zelias International Conference on Modelling and Forecasting of Socio-Economic Phenomena, Cracow University of Economics and the Committee on Statistics and Econometrics of the Polish Academy of Sciences, Zakopane, Poland. Pp:145-153.

[4]. Mood, A. M. (1950). *Introduction to the theory of statistics*. New York, NY: McGraw Hill.

[5]. Nikulin, M. S. (2001). "Efficiency of a statistical procedure". In Hazewinkel, Michiel, *Encyclopedia of Mathematics*, Springer Science Business Media B.V. / Kluwer Academic Publishers.

[6]. Nwakuya, M. T. and Nwabueze, J. C. (2016). Relative Efficiency of Estimates Based on Percentages of Missingness Using Three Imputation Numbers in Multiple Imputation Analysis. *European Journal of Physical and Agricultural Sciences*, 4(1): 63-69

[7]. Umeh, E. U. and Eriobu, N. O. (2016). Comparison of Two Sample Tests Using Both Relative Efficiency and Power of Test. *Open Journal of Statistics*, 6: 331-345

[8]. Wang, H., Feng, C., Tu, X., and Kowalski, J. (2012). A Comparison of Two Test Statistics for Poisson Over-dispersion / Under-dispersion. *Applied Mathematics*, 3: 795-799