Journal of the Nigerian Association of Mathematical Physics
Volume **23 (March, 2013), pp 389 – 394**

The Performance of Laplace Likelihood Ratio Test for Heteroscedasticity over Normal Likelihood ratio test

Amalare A. A.
Department of Mathematics/Statistics,
Lagos state polytechnic, Ikorodu, Nigeria.

Abstract

In this paper, the likelihood ratio test for heteroscedasticity, assuming the Laplace distribution is discussed and shows to give good approximation results for Gaussian and fat-tailed data . The drawback of the likelihood ratio test , assuming normality, is that, it is highly sensitive to any deviation from normality, especially when the observations are from a distribution with fat tails. However, in this work, it is affirmed that the Laplace likelihood ratio test can also be used as a more robust test for a constant variance in residuals or time series if the data is partitioned into groups, than the Normal likelihood ratio test .

Keywords: : Normal likelihood ratio, Laplace distribution, Laplace likelihood ratio, equality of variances , GARCH, time series , beta random variable

1.0 Introduction

The likelihood ratio test for equal variances will be derived under the assumption of Laplace or double exponential distributed observations or residuals. The excess kurtosis of this distribution is three and it is leptokurtic. It is shown that the likelihood ratio test for the equality of variances when assuming the Laplace distribution for the residuals is more robust than the normal one. The distributional properties of this Laplace Likelihood ratio test are very similar to that when normality is assumed, but with a better approximation in the asymptotic chi-square approximation of the log-likelihood than the normal case.

One of the factors to be considered when checking the fit of a model in time series is to see if the residuals are white noise. The use of volatility models for log returns attracted a lot of attention in the last few years, and Autoregressive Conditional Heteroscedasticity (ARCH) and Generalized Autoregressive Conditional Heteroscedasticity (GARCH) models are fitted when heteroscedasticity is present [1] . A test for a constant variance after partitioning the residuals, suitable for observations from a distribution with fatter tails than a normal, can also be used to check for white noise.

The tests of Levene [2], Brown and Forsythe [3] and Gastwirth et al [4] are often used as a more robust tests than the normal likelihood ratio test for the equality of variances in general statistical tests, for examples, ANOVA's (Analysis of variance), Bootstraps methods can also be helpful to investigate the distribution of test statistics in these problems. A review and some suggestions are given in the paper of Boos and Brownie [5]. A simulation study comparing the methods shows much more robust performance of the Laplace Likelihood ratio than the normal likelihood ratio and Levene test.

Assume, that a total of n independent observations $\{u_t\}$, t =1,……,n , from a normal distribution, are available. There are k group with sample sizes n_1,………, n_k . When equal sample sizes are under consideration, it will be assumed that n = kn_0 . Let $n_j\hat{\sigma}_j{}^2=\sum_{j=1}^{n}\left(u_j-\bar{u}_j\right)^2$, j =1,………, k denote the estimated variances for each partition , where $\bar{u}_j$ denotes the sample mean of the jth group or partition. The term $n_j\hat{\sigma}_j{}^2/\sigma^2 \sim\chi^2_{n_j-1}$ is a gamma variable with parameters (n_j -1)/2 , ½ .

In the case of equal variances . The normal likelihood ratio test λ_N proposed by Mood, Graybill and Boes [6] for the hypothesis of the equality of variances , H_0 : $\sigma_1^2 =....=\sigma_k^2=\sigma^2$, for normal data is given as

$$\lambda_N = \frac{\prod_{j=1}^{k}\left(\hat{\sigma}_j^2\right)^{n_j/2}}{\left(\sum_{j=1}^{k} n_j\hat{\sigma}_j^2/n\right)^{n/2}} \qquad (1)$$

A weak point of the statistic (Normal likelihood ratio test , λ_N) is that it is very sensitive to deviations from normality. The statistic and its asymptotic chi-square approximation, -2log(λ_N) , was studied widely and many corrections were suggested to improve the approximation. The ideas and results of Bartlett [7] and Box [8,9] were the basis for many of the asymptotic corrections later derived for the statistic. They considered the statistic M_1 :

Corresponding author: ***H. Akewe,*** E-mail: amalareasm@yahoo.com-, Tel. +2348023998403

$M_1 = (n - k)\log(\sigma^2) - \sum_{j=1}^{k}(n_j - 1)\log(\sigma_j^2)$

$$\hat{\sigma}^2 = \frac{\sum_{j=1}^{k}(n_j-1)\hat{\sigma}_j^2}{(n-k)} \quad , \ (n_j - 1)\hat{\sigma}_j^2 = \sum_{j}^{n_j}(u_j - \bar{u}_j)^2 , \qquad j = 1,\ldots\ldots\ldots, k . \qquad (2)$$

Let $v_j = n_j - 1$, j = 1,……, k . M_1 can be denoted in the normal case as M_1 (0 ; v_1,……., v_k) , where the zero indicates excess kurtosis of zero. The statistic can be generalized to M_1 (γ_{21},…….,γ_{2k} ; v_1,………, v_k) , where each of the k variance estimates is from a population with a different kurtosis , γ_{2j} , j = 1 ,……,k . It is shown that if the kurtosis is equal to γ_2 for the k samples, Box [9] , the statistic is distributed as $\delta^{-1}M_1$ (0; δ_{v1},…………., δ_{vk}) , δ = (1 + $(1/2_{\gamma_2})^{-1}$, in large samples , for any distribution having finite cumulants . Or that the statistic is distributed as (1 + $(1/2)_{\gamma_2}\chi_{k-1}^2$ in large samples .

To find the moments of the log-likelihood ratio was no problem, but the exact distributions of λ_N and also $\log(\lambda_N)$ are both extremely complex and not practical to use. The multivariate version of the normal likelihood tests concerning covariance matrices is covered in detail in the book of Muirhead [10]. The Normal likelihood ratio λ_N is an interesting statistic, and $\lambda_N^{2/n}$ can be viewed as the ratio of the geometric mean of the estimated variances to the arithemetic mean of gamma variables, which is equal to one, only when the individual terms are independent and equal. The ratio of the geometric mean to the arithmetic mean of gamma variables was studied by Glaser [11]. Another way to look at λ_N is to notice that it can be written as the product of Dirichlet random variables, or in this paper, it will be considered as the product of the beta random variables. Let

$$\omega_j = \frac{n_j\hat{\sigma}_j^2}{\sum_{j=1}^{k} n_j\hat{\sigma}_j^2}$$

$$= \frac{\hat{\sigma}_j^2}{\sum_{j=1}^{k}\hat{\sigma}_j^2} \qquad \text{for equal sample sizes .} \qquad (3)$$

This ratio has a beta distribution with parameters $v_j = \frac{n_j-1}{2}$, v = (n – k)/2 . The normal likelihood ratio test can be expressed in terms of the product of beta random variables and

$$\lambda_N^2 = \frac{\prod_{j=1}^{k}(\hat{\sigma}_j^2)^{n_j}}{(\sum_{j=1}^{k} n_j\hat{\sigma}_{j/n}^2)^n}$$

$$= \frac{(n^n/\prod_{j=1}^{k} n_j^{\,n_j})\prod_{j=1}^{k}\left(n_j\hat{\sigma}_j^2\right)^{n_j}}{(\sum_{j=1}^{k} n_j\hat{\sigma}_j^2)^n} \qquad (4)$$

$$= \frac{k^n\prod_{j=1}^{k}\left(\hat{\sigma}_j^2\right)^{n_j}}{\left(\sum_{j=1}^{k}\hat{\sigma}_j^2\right)^n}$$

$$\lambda_N^2 = k^n\prod_{j=1}^{k}\omega_j^{\,n_0}$$

The product of beta and Dirichlet random variable was studied by Springer and Thompson [12] and Rogers and Young [13]. The resulting density is complicated and expressed in terms of the Meijer's G and H-functions proposed by Gradshteyn and Ryzhik [14]. It will be shown that the likelihood test derived from residuals which have the Laplace distribution can also be expressed as a product of the beta random variables for large sample sizes.

2.0 : Methodology

2.1 : **Laplace Likelihood Ratio for Distributed Variables**

The Laplace or double exponential density is given by

$$P(x) = \frac{1}{2\emptyset}\exp\left(-\frac{|x-\theta|}{\emptyset}\right) , \quad -\infty \,□\, x \,□\, \infty , \qquad \emptyset > 0 \qquad (5)$$

The variance is $2\emptyset$ and the median of the observations is the maximum likelihood estimate of the mean θ. The maximum likelihood estimate of $\emptyset$ is $\sum_{j=1}^{n}(|x_j - \hat{\theta}|)/n$ for a sample of size n , where $\hat{\theta}$ denotes the estimated median. For θ known , $\sum_{j=1}^{n}(|x_j - \theta|)/n$ is distributed as a $(2n)^{-1}□\chi_{2n}^2$ variable. The properties of the Laplace distribution are reviewed in the book of Johnson et al [15]. The variance of the median is O(n^{-1}), and the absolute deviations $|x_j - \hat{\theta}|$, j = 1,….., n , are asymptotically independent [16].

For the series u_1, ... , u_n partition into k parts, the Laplace likelihood ratio λ_L , for the test H_0 : □$_1$ = ……..= □$_k$ = □ , is

$$\lambda_L = \frac{\prod_{j=1}^{k}\hat{\phi}_j^{\,n_j}}{\hat{\phi}^n} \quad , \qquad (6)$$

With $\hat{\phi}_j = \sum_{j=1}^{k}(|x_j - \hat{\theta}_j|)/n_j$ and $\hat{\phi} = (1/n)(n_1\hat{\phi}_1 + \cdots + n_k\hat{\phi}_k)$.

For equal sample sizes , n_0 , the ratio is simplified to

$$\lambda_L = k^n \prod_{j=1}^{k} \left(\frac{\hat{\phi}_j}{\sum_{j=1}^{k} \hat{\phi}_j} \right)^{n_0} , \qquad (7)$$

and $\lambda^{1/n}$ is proportional to the geometric mean of the ratios. If θ was known, the ratio $\hat{\phi}_j / \sum_{j=1}^{k} \hat{\phi}_j$ has a beta distribution with parameters n_j, n - n_j . This variance is approximately half that of the beta variable for the normal case . Terms involving the distribution of the sum of the log of powers of beta random variables are found in the normal and Laplace likelihood ratios . The moment-generating function of the log of a beta variable with parameters n_j and n - n_j to a power is

$$\begin{aligned} \varphi(t) &= E(\log(e^{t \log x^h})) \\ &= E(x^{th}) \\ &= \frac{\Gamma(n)\,\Gamma(n_j+ht)}{\Gamma(n+ht)\,\Gamma(n_j)} , \end{aligned} \qquad (8)$$

And the log of the moment- generating function of the sum of k , such variables is

$$E(\log(\varphi(t))) = k \log(\Gamma(n)) - k \log(\Gamma(n_j)) + \sum_{j=1}^{k} \log (\Gamma(n_j + ht)) - k \log(\Gamma(n+ht)) , \qquad (9)$$

Showing that the expected value of -2log(λ_L) found from the cumulant generating function is

$$\begin{aligned} E(-2\log(\lambda_L)) &= -2n\log(k) - 2\sum_{j=1}^{k} E(\log (x_j^{n_j})) \\ &= -2n\log(k) - 2\sum_{j=1}^{k}(n_j\varphi(n_j) - n_j\varphi(n)) \end{aligned} \qquad (10)$$

Where x_j denotes a beta variable with parameters n_j , n - n_j and φ is the digamma function, the derivative of the log of the gamma function . Then, making use of the approximation of φ suggested by Gradshteyn and Ryzhik [14] , φ(n) ≈ log (n) - (1/n)(1/2+1/12n)

$$\begin{aligned} E(-2\log(\lambda_L)) &= k - 1 - 2n\log(k) + 2\sum_{j=1}^{k}(n_j \log(n) - n_j \log(n_j)) + \frac{\sum_{j=1}^{k} 1}{(1+6n_j} - \frac{1}{(1-6n} \\ &\approx k - 1 , \end{aligned} \qquad (11)$$

For n = kn_0 , $n_0 = n_1 =$ -----= n_k , and large sample sizes , the expected value -2log(λ_L) is equal to that of χ^2_{k-1} random variable . For the normal ratio , it would be $v_j = (n_j - 1)/2$ in place of the n_j's , showing that the large sample χ^2 approximation for the Laplace likelihood ratio would be a better approximation for the same sample size n , assuming the distributional assumption concerning θ

The assumption of an expected value of zero for log returns is often made in financial time series of returns, but in most problems , the expected value would be unknown , and the median of the observations of a specific partition would be used. The median is a maximum likelihood estimator and good approximation can be expected for reasonable large sample sizes.

2.2: **Statement of the problem**

Consider a simulation study of sample size n = 200 partitioned into k = 5 equal parts such that comparism were made amongst the estimated: i). expected mean ii) variance, with their respective theoretical values to show if the generated simulation results is close to Gaussian curve in nature .

3.0: Results and Discussion

In Figure 1 , the histogram of simulated and expected frequencies of 1000 ratio is shown . The ratios were calculated using Gaussian white noise series . The $\hat{\phi}$'s are estimated using the median as the estimator of θ . , $\hat{\phi}_1 / \sum_{j=1}^{k} \hat{\phi}_j$, is shown in Figure 1 . The expected values are from a beta distribution with parameters $n_1 = 40$, n - $n_1 = 160$. The expected mean was 0.199875 compare with the theoretical value of 0.2 , and the estimated variance 7.9740e-004 compared with the theoretical value of 7.9602e-004 . While Table 1 , revealed the descriptive statistics summary from a beta distribution . However, since the estimated expected mean and variance were close to the respective theoretical values . Consequently, Fiqure 1, shows a better approximation resuits for the simulated data with a well displaced normal (Gaussian) curve .

The sample size of n = 200 is not very large for a time series. It can be seen that the ratios of the estimated ϕ_j's to the sum of the ϕ_j's are approximately beta distributed for this sample size

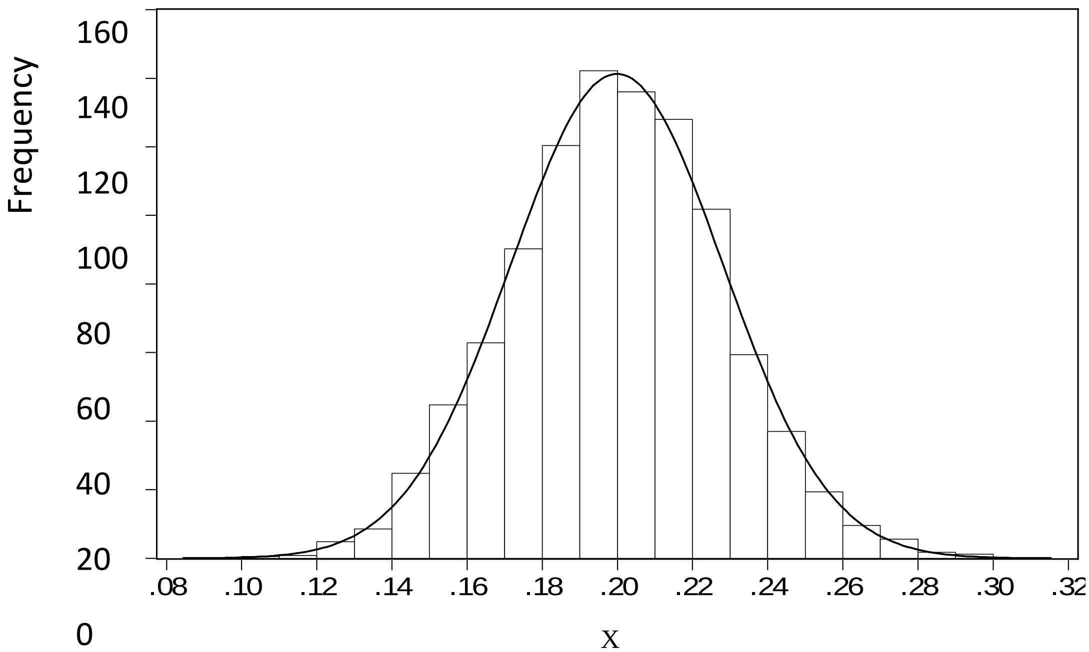

Figure 1: Simulated ratios and the expected frequencies of the beta density

Table 1: Descriptive Statistics Summary from a Beta distribution

	X
Mean	0.1998750
Median	0.1999720
Maximum	0.2946600
Minimum	0.0944350
Std. Dev.	0.0282400
Variance	0.0007974
Skewness	0.0077680
Kurtosis	3.0322540
Jarque-Bera	0.2670220
Probability	0.8750180
Observations	1000

In addition, the generated simulation data obtained in the given problem in section (2.2) were used to test and compare the asymptotic χ^2 approximation of λ_N, λ_L of the likelihood ratios and Brown-Forsythe variation of the Levene tests at $\propto = 0.05$ level of significance . Since series of n = 200 observations , partitioned into k = 5 , were simulated 1000 times . The expected values for the normal and Laplace χ^2 approximation are k – 1 = 4 , and for the Levene statistic [2] , the expected value of $F_{4;196}$ variable is 1.0103 . The following data was generated in Table 2: normal white noise , white noise from the Laplace distribution , and independent values from the stable distribution with index $\propto = 1.9$ and 1.5 .

Time series were also generated to check the results when the test is used for checking a constant variance in residuals . The series generated were an AR(1) and a Garch series . The disturbance terms are normally distributed . The Garch series is the IGARCH(1,1) fitted by Tsay [17] to excess returns. The results are shown in Table 2.

Table 2: Results of simulation study comparing the likelihood ratio tests and the Levene test

	Laplace		Normal		Levene	
	Proportion Rejected	Mean of -2log(λ_L)	Proportion Rejected	Mean of -2log(λ_N)	Proportion Rejected	Mean of Levene W)
Gaussian White Noise	0.0020	2.3277	0.0500	4.0783	0.0430	0.9455
Laplace White Noise	0.0570	4.0606	0.4250	9.7604	0.0840	1.0904
t-distribution (df = 4) White Noise	0.0560	4.0253	0.5240	13.8947	0.1120	1.2181
Stable (α =1.5)	0.4880	19.2967	0.9370	107.9499	0.3410	3.5335
Stable (α =1.9)	0.0530	4.3012	0.3950	23.5564	0.1190	1.5190
AR(1) ρ_1 = 0.1	0.0040	2.3710	0.0630	4.1258	0.0450	0.9491
AR(2) ρ_1 = 0.5	0.0290	3.5989	0.2140	6.4760	0.1590	1.4636
IGARCH(1,1)	0.7890	34.6704	0.9410	83.8057	0.6970	7.9725

Both tests are sensitive to heteroscedasticity , but the Laplace likelihood ratio test (λ_L) is less sensitive when testing for a constant variance for non-Gaussian white noise , since the normal likelihood ratio test (λ_N) can be very sensitive and effective only in the normal data case. It is interesting to note that both tests are sensitive to large autocorrelation in the series and also when the variance is infinite for stable data with $\propto < 200$. The stable noise with index $\propto = 1.9$ is close to Gaussian , but theoretically only E($x^{\propto}$) is finite . All the tests detect that the series with $\propto = 1.5$ is not second-order stationary . The tests are not very sensitive when autocorrelation is present in the AR(1) models, but the sensitivity increases as the first-order autocorrelation increases . All the tests especially Laplace Likelihood ratio test (λ_L) easily detect the heteroscedasticity in the IGARCH(1,1) model .

4.0 Conclusion

The Laplace likelihood test performs much better on data with heavier tails ,and better than the Brown-Forsythe variation of the Levene test , for example , in the case of the t-distributed series . The normal likelihood ratio test is only effective in the normal data case . It seems that these tests are sensitive for serial correlation and heteroscedasticity in series and can be used as a check for white noise in residuals. The use of filtering as proposed by Lumsdaine and Ng [18] can be applied to investigate and improve results when testing for heteroscedasticity where serial correlation is present in time series models. An investigation into the use of size-adjustment can improve the effectiveness of Laplace likelihood ratio test λ_L , especially when working with GARCH type series.

References

[1] Bollerslev , T. (1986) ‘’ Generalized Autoregressive Conditional Heteroscedasticity, ‘’ Journal of Econometrics . vol.31 , 307-327 .

[2] Levene, H. (1960) ‘’ Robust tests for equality of variances in contributions to Probability and Statistics, ‘’ pp. 278- 292 , Stanford University Press, Palo Alto, CA, USA.

[3] Brown , M. B and Forsythe , A. B (1974) ‘’ Robust test for equality of variances, ‘’ Journal of the American Statistical Association , vol. 69, no. 346 , pp. 364-367 .

[4] Gastwirth, J. L, Gel, Y. R, and Miao, W. (2009) ‘’ The impact of Levene’s test of equality of variances on statistical theory and practice, ‘’ Statistical Science , vol. 24, no. 3, pp. 343-360 .

[5] Boos, D.D and Brownie, C. (2004) ‘’ Comparing variances and other measures of dispersion, ‘’ Statistical Science, vol. 19 , no. 4, 571-578 .

[6] Mood, A. M, Graybill, F. A and Boes D. C (1974) ‘’ Introduction to the Theory of Statistics, ‘’ McGraw-Hill, New York, USA.

[7] Bartlett , M.S, (1937) ‘’ Properties of sufficiency and Statistical tests,’’ Proceedings of the Royal Society A, vol.160 , no. 901 , pp. 262-282 .

[8] Box, G. E. (1949) ‘’ A general distribution theory for a class likelihood criteria, ‘’ Biometrika , vol. 36, pp. 317-346 .

[9] Box , G. E. (1953) ‘’ Non-normality and tests on variances, ‘’ Biometrika, vol. 40, pp. 318-335 .

[10] Muirhead, R. J (1982) ‘’ Aspects of Multivariate Statistical Theory, ‘’ John Wiley & Sons, New York, USA, Wiley Series in Probability and Mathematical Statistics.

[11] Glaser, R. E. (1976) ‘’ The ratio of the geometric mean to the arithmetic mean for a random sample from a gamma distribution, ‘’ Journal of the American Statistical Association, vol. 71, no. 354, pp. 480-487 .

[12] Springer, M. D and Thompson, W. E (1970) ‘’ The distribution of products of beta, gamma and Gaussian random variables, ‘’ SIAM Journal on Applied Mathematics, vol. 18, pp. 721-737.

[13] Rogers, G. S and Young, D. L (1973) ‘’ On the products of powers of generalized Dirichlet components with an application, ‘’ The Canadian journal of Statistics, vol. 1, no. 2, pp. 159-169.

[14] Gradshteyn, I. S and Ryzhik, M. Y. (1980) ‘’ Table of Integrals, Series and Products, ‘’ Academic Press, New York, USA.

[15] Johnson, N. L, Kotz, S. and Balakrishnan, N.,(1995) ‘’ Continuous Univariate Distribution, vol. 2 of Wiley series in Probability and Mathematical Statistics: Applied Probability and Statistics, ‘’ John Wiley & Sons, New York, 2nd edition, USA

[16] Kendall, M. , Stuart, A and Ord, J. K. (1987) ‘’ Kendall’s Advanced Theory of Statistics, ‘’ vol. 1, Charles Griffin and Company, London, UK.

[17] Tsay, R. S (2010) ‘’ Analysis of Financial Time Series, ‘’ Wiley Series in Probability and Statistics, John Wiley & Sons, New York, 3rd edition, USA .

[18] Lumsdaine, R. L and Ng, S. (1999) ‘’ Testing for ARCH in the presence of a possibly misspecified conditional mean, ‘’ Journal of Econometrics, vol. 93, no. 2 , pp. 257-279