Journal of Modern Applied Statistical Methods

Volume 4 | Issue 2 Article 6

11-1-2005

Testing Normality Against The Laplace Distribution

Taisuke Otsu
Yale University, taisuke.otsu@yale.edu

Journal of Modern Applied Statistical Methods
November, 2005, Vol. 4, No. 2, 402-407

1538 – 9472/05/$95.00

Testing Normality Against The Laplace Distribution

Taisuke Otsu
Cowles Foundation
Yale University

Some normality test statistics are proposed by testing non-nested hypotheses of the normal distribution and the Laplace distribution. If the null hypothesis is normal, the proposed non-nested tests are asymptotically equivalent to Geary's (1935) normality test. The proposed test statistics are compared by the method of approximate slopes and Monte Carlo experiments.

Key words: Normality test; non-nested hypothesis; Cox test; Atkinson test

Introduction

In statistical analysis, many models and methods rely upon the assumption of normality, which should be examined by some adequate tests. However, in several data (e.g. economic and financial data), the existence of outliers is much frequent, and the observations or disturbances may have some leptokurtic distributions, where the kurtosis is larger than three. In order to detect such leptokurtic non-normal distributions, we apply the method of non-nested testing which has high sensitivity (power) for an explicit alternative hypothesis.

Based on Cox (1961, 1962) and Atkinson (1970), it this article non-nested test statistics between the normal distribution and the Laplace (or double-exponential) distribution, which is a typical leptokurtic distribution are proposed. All of the proposed test statistics are asymptotically normal. When the null hypothesis is normal, these test statistics are asymptotically equivalent to Geary's (1935) normality test statistic.

In the context of regression models, the maximum likelihood estimator with the Laplace distribution error is the least absolute deviation (LAD) estimator. Therefore, these test statistics are also useful to decide whether the LAD regression or the conventional OLS regression should be applied.

By applying Pesaran's (1987) strict definition of non-nested hypotheses, we find that the normal distribution and the Laplace distribution are globally non-nested, and that the power analysis using Pitman-type local alternatives is not available. Therefore, these non-nested test statistics are compared by the method of approximate slope (or Bahadur efficiency) developed by Bahadur (1960, 1967). Furthermore, Monte Carlo simulations are carried out to compare the small sample properties of the proposed tests and other conventional normality tests. Simulation results indicate that these tests show reasonable performances in terms of the size and power.

Non-nested Test Statistics

Throughout this article, demeaned observations are considered, i.e., the mean is assumed to be zero. Let $Y = (Y_1, \ldots, Y_n)$ be independently and identically distributed (iid)

Taisuke Otsu is an Assistant Professor of Economics. His research interests are in empirical likelihood, nonparametric and semiparametric methods, and microeconometrics. Contact him at 30 Hillhouse Ave., Rm. 36,Department of Economics, Yale University,P.O. Box 208281, New Haven, CT 06520-8281, or taisuke.otsu@yale.edu.

random variables. Consider the following non-nested hypotheses:

$$H_f : f(y;\alpha) = \frac{1}{\sqrt{2\pi\alpha}} \exp\left[-\frac{y^2}{2\alpha}\right], \quad (1)$$

$$H_g : g(y;\beta) = \frac{1}{2\beta} \exp\left[-\frac{|y|}{\beta}\right], \quad (2)$$

where H_f is the normal distribution with zero mean, and H_g is the Laplace distribution with zero mean. H_f and H_g belong to separate parametric families and are called non-nested hypotheses. In order to test non-nested hypotheses, Cox (1961, 1962) proposed a testing procedure based on a modified likelihood ratio. When H_f is the null hypothesis and H_g is the alternative hypothesis, the Cox test statistic is written as

$$T_f = L_f(\hat{\alpha}) - L_g(\hat{\beta}) - \mathrm{E}_{\hat{\alpha}}(L_f(\alpha) - L_g(\beta_\alpha)), \quad (3)$$

where $L_f(\alpha) = \sum_{i=1}^{n} \log f(y_i;\alpha)$ and $L_g(\beta) = \sum_{i=1}^{n} \log g(y_i;\beta)$ denotes the log likelihood functions of the hypotheses H_f and H_g, respectively, $\hat{\alpha}$ and $\hat{\beta}$ denote the maximum likelihood estimators under H_f and H_g, respectively, $\mathrm{E}_{\hat{\alpha}}(\cdot)$ is the expected value under H_f when α takes the value $\hat{\alpha}$, and $\beta_\alpha = \mathrm{plim}_\alpha \hat{\beta}$ is the probability limit of $\hat{\beta}$ under H_f as $n \to \infty$. Define

$$F_i = \log f(Y_i;\alpha), \; G_i = \log g(Y_i;\beta_\alpha),$$
$$F_{\alpha i} = \frac{\partial \log f(Y_i;\alpha)}{\partial \alpha}. \quad (4)$$

Cox (1961, 1962) showed that T_f is asymptotically normal with zero mean and variance

$$\mathrm{V}_\alpha(T_f) = n\left[\mathrm{V}_\alpha(F_i - G_i) - \frac{\mathrm{C}_\alpha^2(F_i - G_i, F_{\alpha i})}{\mathrm{V}_\alpha(F_{\alpha i})}\right], \quad (5)$$

where $V_\alpha(\cdot)$ and $\mathrm{C}_\alpha(\cdot,\cdot)$ denote the variance and the covariance under H_f, respectively.

In the same manner, set the Laplace distribution H_g as the null hypothesis and set the normal distribution H_f as the alternative hypothesis. In this case, the Cox test statistic T_g is written as

$$T_g = L_g(\hat{\beta}) - L_f(\hat{\alpha}) - \mathrm{E}_{\hat{\beta}}(L_g(\hat{\beta}) - L_f(\alpha_\beta)), \quad (6)$$

where $\mathrm{E}_{\hat{\beta}}(\cdot)$ is the expected value under H_g when β takes the value $\hat{\beta}$, and $\alpha_\beta = \mathrm{plim}_\beta \hat{\alpha}$ is the probability limit of $\hat{\alpha}$ under H_g as $n \to \infty$. T_g is also asymptotically normal with zero mean and variance $\mathrm{V}_\beta(T_g)$, which is defined in the same manner as (4). If $\mathrm{V}_\alpha(T_f)$ and $\mathrm{V}_\beta(T_g)$ are consistently estimated by $\mathrm{V}_{\hat{\alpha}}(T_f)$ and $\mathrm{V}_{\hat{\beta}}(T_g)$, respectively,

$$N_f = T_f / \sqrt{\mathrm{V}_{\hat{\alpha}}(T_f)}, \; N_g = T_g / \sqrt{\mathrm{V}_{\hat{\beta}}(T_g)} \quad (7)$$

can be used as test statistics which follow the standard normal limiting distribution.

In setup (1) and (2), obtain

$$\hat{\alpha} = \sum_i Y_i^2 / n, \; \hat{\beta} = \sum_i |Y_i| / n, \quad (8)$$

$$\beta_\alpha = \mathrm{plim}_\alpha \hat{\beta} = \mathrm{E}_\alpha(|Y_i|) = \sqrt{2\alpha/\pi},$$
$$\alpha_\beta = \mathrm{plim}_\beta \hat{\alpha} = \mathrm{E}_\beta(Y_i^2) = 2\beta^2. \quad (9)$$

Therefore, when the null hypothesis is normal and the alternative hypothesis is Laplace, the Cox test statistic is

$$T_f = n\log\left(\frac{\hat{\beta}}{\beta_{\hat{\alpha}}}\right) = n\log\left(\sqrt{\frac{\pi}{2}}\frac{\hat{\beta}}{\sqrt{\hat{\alpha}}}\right), \qquad (10)$$

with the asymptotic variance $V_\alpha(T_f) = \frac{\pi}{2} - \frac{3}{2}$. On the other hand, when the null hypothesis is Laplace and the alternative hypothesis is normal, the Cox test statistic is

$$T_g = \frac{n}{2}\log\left(\frac{\hat{\alpha}}{\alpha_{\hat{\beta}}}\right) = \frac{n}{2}\log\left(\frac{\hat{\alpha}}{2\hat{\beta}^2}\right), \qquad (11)$$

with the asymptotic variance $V_\beta(T_g) = \frac{1}{4}$.

Next, derive Atkinson's (1970) test. The Atkinson test procedure is derived from the comprehensive probability density function (pdf), which includes $f(y;\alpha)$ and $g(y;\beta)$ as special cases. When H_f is the null hypothesis and H_g is the alternative hypothesis, the Atkinson test statistic is written as

$$TA_f = L_f(\hat{\alpha}) - L_g(\beta_{\hat{\alpha}}) - E_{\hat{\alpha}}(L_f(\alpha) - L_g(\beta_\alpha)). \qquad (12)$$

Comparing (3) and (12), the difference between T_f and TA_f is their second terms. Because the Atkinson test TA_f and the Cox test T_f are asymptotically equivalent under H_f, the asymptotic variance of TA_f is same as (5) (see Pereira, 1977). Analogous results are obtained for the case where H_g is the null hypothesis and H_f is the alternative hypothesis. In order to conduct the Atkinson test, we can use

$$NA_f = TA_f/\sqrt{V_{\hat{\alpha}}(T_f)},\ NA_g = TA_g/\sqrt{V_{\hat{\beta}}(T_g)} \qquad (13)$$

as test statistics which follow the standard normal limiting distribution. When the null hypothesis is normal and the alternative hypothesis is Laplace, the Atkinson test statistic is:

$$TA_f = n\left(\frac{\hat{\beta}}{\beta_{\hat{\alpha}}} - 1\right) = n\left(\sqrt{\frac{\pi}{2}}\frac{\hat{\beta}}{\sqrt{\hat{\alpha}}} - 1\right), \qquad (14)$$

and when the null hypothesis is Laplace and the alternative hypothesis is normal, the Atkinson test statistic is

$$TA_g = \frac{n}{2}\left(\frac{\hat{\alpha}}{\alpha_{\hat{\beta}}} - 1\right) = \frac{n}{2}\left(\frac{\hat{\alpha}}{2\hat{\beta}^2} - 1\right). \qquad (15)$$

Because the computation of our non-nested test statistics (i.e., N_f, N_g, NA_f, and NA_g) needs only $\hat{\alpha}$ and $\hat{\beta}$, their implementation is quite easy.

T_f and TA_f are related to another normality test suggested by Geary (1935). The Geary test statistic is written as

$$G = \frac{\sum_i |Y_i|}{\sqrt{n\sum_i Y_i^2}} = \frac{\hat{\beta}}{\sqrt{\hat{\alpha}}}, \qquad (16)$$

From (10) and (14), the relationships among G, T_f, and TA_f are

$$T_f = n\log\left(\sqrt{\frac{\pi}{2}}G\right),\ TA_f = n\left(\sqrt{\frac{\pi}{2}}G - 1\right). \qquad (17)$$

Therefore, if the standardized test statistics is compared, it can be shown that under H_f the Cox test and the Atkinson test are asymptotically equivalent to the Geary test.

Power Comparison

This section considers theoretical properties of the proposed non-nested tests. We first investigate the consistency of the Cox test and the Atkinson test. Pereira (1977) showed that the Cox test is always consistent, but the Atkinson test is not always consistent. From (14) and (15):

$$\text{plim}_{\beta} n^{-1} TA_f = \sqrt{\pi}/2 - 1 \approx -0.1138, \qquad (18)$$

$$\text{plim}_{\alpha} n^{-1} TA_g = (1/2)(\pi/4 - 1) \approx -0.1073. \qquad (19)$$

Because both TA_f and TA_g converge to non-zero constants, the Atkinson test is consistent in our particular setup.

Using Pesaran's (1987) strict definition of the non-nested hypotheses, which is based upon the Kullback-Leibler information criterion (KLIC), next examine the relationship between the normal distribution (H_f) and the Laplace distribution (H_g). The KLIC for the pdf $f(y;\alpha)$ against the pdf $g(y;\beta)$ is defined as

$$I_{fg}(\alpha,\beta) = E_{\alpha}(\log f(y;\alpha) - \log g(y;\beta)). \qquad (20)$$

Assume that $I_{fg}(\alpha,\beta)$ has a unique minimum at $\beta_*(\alpha)$. Pesaran (1987) defined the closeness of H_g to H_f as

$$C_{fg}(\alpha) = I_{fg}(\alpha, \beta_*(\alpha)). \qquad (21)$$

Similarly, define the KLIC for $g(y;\beta)$ against $f(y;\alpha)$ (denote $I_{gf}(\beta,\alpha)$) and the closeness of H_f to H_g (denote $C_{gf}(\beta)$). Using $C_{fg}(\alpha)$ and $C_{gf}(\beta)$, Pesaran (1987) classified the relationship between two hypotheses into three categories, i.e., nested, globally non-nested, and partially non-nested. In the case of (1) and (2), $I_{fg}(\alpha,\beta)$ and $I_{gf}(\beta,\alpha)$ are written as

$$I_{fg}(\alpha,\beta) = -\frac{1}{2}\log(2\pi\alpha) + \log(2\beta) + \frac{1}{\beta}\sqrt{\frac{2\alpha}{\pi}} - \frac{1}{2}, \qquad (22)$$

$$I_{gf}(\beta,\alpha) = \frac{1}{2}\log(2\pi\alpha) - \log(2\beta) + \frac{\beta^2}{\alpha} - 1. \qquad (23)$$

Because $\beta_*(\alpha) = \sqrt{2\alpha/\pi}$ and $\alpha_*(\beta) = 2\beta^2$,

$$C_{fg}(\alpha) = \log\left(\frac{2}{\pi}\right) + \frac{1}{2} \approx 0.04842, \qquad (24)$$

$$C_{gf}(\beta) = \log(\sqrt{\pi}) - \frac{1}{2} \approx 0.07236. \qquad (25)$$

Because both $C_{fg}(\alpha)$ and $C_{gf}(\beta)$ are nonzero constants, H_f and H_g are globally non-nested and the power analysis using a local alternative is not available (see Pesaran (1987)).

Because the Pitman-type power analysis cannot be applied, compare the Cox test and the Atkinson test by the method of approximate slopes developed by Bahadur (1960, 1967). The method of approximate slopes compares the convergence rates of the significance levels of tests (to zero) under some fixed alternative hypothesis with some fixed power.

Thus, approximate slopes are useful to analyze the power properties of tests under globally non-nested hypotheses. Let $\tilde{\alpha}_n$ be the asymptotic significance level of some test with a given sample size n. The approximate slope is defined as $\lim(-2n^{-1}\log\tilde{\alpha}_n)$. If a test T_1 has a greater approximate slope than another test T_2, we call that T_1 is Bahadur efficient relative to T_2. Pesaran (1984) showed that the approximate slopes of the Cox test and the Atkinson test are given by $\text{plim}_{\beta}(n^{-1}N_f^2)$ and $\text{plim}_{\beta}(n^{-1}NA_f^2)$, respectively. Therefore, from (10), (11), (14), and (15),

$$\text{plim}_{\beta} n^{-1} N_f^2 = \frac{\left(\log\left(\frac{\sqrt{\pi}}{2}\right)\right)^2}{\frac{\pi}{2} - \frac{3}{2}} \approx 0.2061, \qquad (26)$$

$$\text{plim}_{\beta} n^{-1} NA_f^2 = \frac{\left(\frac{\sqrt{\pi}}{2} - 1\right)^2}{\frac{\pi}{2} - \frac{3}{2}} \approx 0.1828, \qquad (27)$$

$$\text{plim}_{\alpha} n^{-1} N_g^2 = \left(\log\left(\frac{\pi}{4}\right)\right)^2 \approx 0.05835, \qquad (28)$$

Table 1. Finite sample rejection frequencies of the null hypothesis at the one side 5% level

DGP	n	T_f	T_g	TA_f	TA_g	BS	SW	DA	AD
	20	0.0429	0.1812	0.0368	0.0239	0.0234	0.0469	0.0526	0.0512
Normal	50	0.0451	0.6167	0.0410	0.4438	0.0353	0.0494	0.0488	0.0509
	100	0.0498	0.9291	0.0469	0.8875	0.0434	0.0484	0.0525	0.0522
	20	0.3427	0.0311	0.3012	0.0014	0.2118	0.2498	0.3556	0.2663
Laplace	50	0.7072	0.0418	0.6945	0.0190	0.5107	0.4105	0.6927	0.5498
	100	0.9377	0.0460	0.9339	0.0254	0.7783	0.5386	0.9175	0.8265
	20	0.1184	0.0995	0.1066	0.0108	0.0931	0.1102	0.1497	0.1052
Logistic	50	0.2549	0.2859	0.2428	0.1678	0.2313	0.1459	0.2984	0.1682
	100	0.4072	0.5356	0.3957	0.4512	0.3673	0.1289	0.4531	0.2367

$$\text{plim}_{\alpha} n^{-1} NA_g^2 = \left(\frac{\pi}{4} - 1\right)^2 \approx 0.04605. \qquad (29)$$

In both cases (i.e., the null is normal, and the null is Laplace), the Cox test is Bahadur efficient relative to the Atkinson test. Thus, the Cox test has better global power property than the Atkinson test.

Results

In order to analyze the finite sample properties of the proposed tests, we conduct Monte Carlo simulation. In addition to the non-nested test statistics in (10), (11), (14), and (15), consider the normality tests by Bowman and Shenton (1975) (BS), Shapiro and Wilk (1965) (SW), D'Agostino (1971) (DA) and Anderson and Darling (1954) (AD), which is a modified Kolmogorov-Smirnov test, as alternative tests.

As the data generating process (DGP), employ the standard normal, standard Laplace, and standard logistic distribution. The sample sizes are set as $n = (20, 50, 100)$. The number of replications is 10000.

Table 1 shows finite sample rejection frequencies of the null hypothesis at the 5% level. From this table, the following may be seen. First, the Cox test T_f with the normal null hypothesis demonstrates better performances than the Atkinson test TA_f in terms of the size accuracy and power. This power superiority of T_f is consistent with the relative Bahadur efficiency of T_f. Second, comparing to the other normality tests, T_f has the highest power when the DGP is the standard Laplace distribution. Also T_f is second best when the DGP is the logistic distribution. Third, the Atkinson test TA_g with the Laplace null hypothesis shows enough power when the DGP is the standard normal distribution. Note that T_g and TA_g can provide additional information, which cannot be obtained by the conventional normality tests based on the normal null hypothesis.

Conclusion

By applying the Cox and Atkinson test, we propose the non-nested test statistics of the normal and the Laplace distribution. The proposed test statistics proposed are asymptotically normal, and are easily computed. Approximate slopes show that the Cox test has better power properties than the Atkinson test. In simulation, the Cox test with the normal null hypothesis shows higher power for leptokurtic distributions comparing to the other normality tests. The Atkinson test with the Laplace null hypothesis is also useful to analyze distributional forms of data.

References

Anderson, T. W. & D. A. Darling (1954). A test of goodness-of-fit. *Journal of the American Statistical Association 49*, 765-769.

Atkinson, A. C. (1970). A method for discriminating between models. *Journal of the Royal Statistical Society B32*, 323-353.

Bahadur, R. R. (1960). Stochastic comparison of tests. *Annals of Mathematical Statistics 31*, 276-295.

Bahadur, R. R. (1967). Rates of convergence of estimates and test statistics, *Annals of Mathematical Statistics 38*, 303-324.

Bowman, K.O. & Shenton, B. R. (1975). Omnibus test contours for departures from normality based on $\sqrt{b_1}$ and b_2. *Biometrika, 62*, 243-50.

Cox, D. R. (1961). Tests of separate families of hypotheses. In: *Proceedings of the fourth Berkeley symposium on mathematical statistics and probability, Vol. 1*. University of California Press, Berkeley, p. 105-123.

Cox, D. R. (1962). Further results on tests of separate families of hypotheses. *Journal of the Royal Statistical Society B24*, 406-424.

D'Agostino, R. B. (1971). An omnibus test of normality for moderate and large size samples. *Biometrika 58*, 341-348.

Geary, R. C. (1935). The ratio of the mean deviation to the standard deviation as a test of normality. *Biometrika 27*, 310-332.

McAleer, M. (1986). Specification tests for separate models: a survey, Ch. 9 in King, M.L. & Giles, D.E.A. (ed.), *Specification analysis in the linear model*, London: Routledge and Kegan Paul, p. 146-196.

Pereira, B. B. (1977). A note on the consistency and on the finite sample comparisons of some tests of separate families of hypotheses. *Biometrika 64*, 109-113.

Pesaran, M. H. (1984). Asymptotic power comparisons of tests of separate families by Bahadur's approach. *Biometrika 71*, 245-252.

Pesaran, M. H. (1987). Global and partial non-nested hypotheses and asymptotic local power. *Econometric Theory 3*, 69-97.

Shapiro, S. S. & Wilk, M. B. (1965). An analysis of variance test for normality. *Biometrika, 52*, 591-611.