11-1-2005

Testing Normality Against The Laplace Distribution

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

Follow this and additional works at: http://digitalcommons.wayne.edu/jmasm
Part of the Applied Statistics Commons, Social and Behavioral Sciences Commons, and the Statistical Theory Commons

Recommended Citation
DOI: 10.22237/jmasm/1130803500
Available at: http://digitalcommons.wayne.edu/jmasm/vol4/iss2/6

This Regular Article is brought to you for free and open access by the Open Access Journals at DigitalCommons@WayneState. It has been accepted for inclusion in Journal of Modern Applied Statistical Methods by an authorized editor of DigitalCommons@WayneState.
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)
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

\[ T_f = L_f(\hat{\alpha}) - L_g(\hat{\beta}) - E\hat{\alpha}(L_f(\alpha) - L_g(\beta)), \quad (3) \]

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. \( E\hat{\alpha}(\cdot) \) is the expected value under \( H_f \) when \( \alpha \) takes the value \( \hat{\alpha} \), and \( \beta_\alpha = \text{plim}_\beta \hat{\beta} \) is the probability limit of \( \hat{\beta} \) under \( H_f \) as \( n \to \infty \). Define

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

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

\[ V_\alpha(T_f) = n \left[ V_\alpha(F_i - G_i) - \frac{C_\alpha(F_i - G_i, F_{\alpha i})}{V_\alpha(F_{\alpha i})} \right], \quad (5) \]

where \( V_\alpha(\cdot) \) and \( C_\alpha(\cdot, \cdot) \) 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

\[ T_g = L_g(\hat{\beta}) - L_f(\hat{\alpha}) - E\hat{\beta}(L_g(\beta) - L_f(\alpha)), \quad (6) \]

where \( E\hat{\beta}(\cdot) \) is the expected value under \( H_g \) when \( \beta \) takes the value \( \hat{\beta} \), and \( \alpha_\beta = \text{plim}_\alpha \hat{\alpha} \) 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 \( V_\beta(T_g) \), which is defined in the same manner as (4). If \( V_\alpha(T_f) \) and \( V_\beta(T_g) \) are consistently estimated by \( \hat{V}_\alpha(T_f) \) and \( \hat{V}_\beta(T_g) \), respectively,

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

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

In setup (1) and (2), obtain

\[ \hat{\alpha} = \sum Y_i^2 / n, \quad \hat{\beta} = \sum |Y_i| / n, \quad (8) \]

\[ \beta_\alpha = \text{plim}_\alpha \hat{\alpha} = E\alpha(|Y_i|) = \sqrt{2\alpha \pi}, \quad \alpha_\beta = \text{plim}_\beta \hat{\alpha} = E\beta(Y_i^2) = 2\beta^2. \quad (9) \]
Therefore, when the null hypothesis is normal and the alternative hypothesis is Laplace, the Cox test statistic is

\[ T_f = n \log \left( \frac{\hat{\beta}}{\beta_0} \right) = n \log \left( \frac{\pi}{\sqrt{2}} \frac{\hat{\beta}}{\sqrt{\alpha}} \right), \]  

(10)

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

\[ T_g = \frac{n}{2} \log \left( \frac{\hat{\alpha}}{\alpha_\beta} \right) = \frac{n}{2} \log \left( \frac{\hat{\alpha}}{2\beta^2} \right), \]  

(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

\[ T_A = L_f(\hat{\alpha}) - L_g(\beta_0) - E_0(L_f(\alpha) - L_g(\beta_0)). \]  

(12)

Comparing (3) and (12), the difference between \( T_f \) and \( T_A \) is their second terms. Because the Atkinson test \( T_A \) and the Cox test \( T_f \) are asymptotically equivalent under \( H_f \), the asymptotic variance of \( T_A \) 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

\[ N_f = T_A / \sqrt{V_\alpha(T_f)}, \quad N_g = T_A / \sqrt{V_\beta(T_g)} \]  

(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:

\[ T_A = n \left( \frac{\hat{\beta}}{\beta_0} - 1 \right) = n \left( \frac{\sqrt{\pi}}{2} \frac{\hat{\beta}}{\sqrt{\alpha}} - 1 \right), \]  

(14)

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

\[ T_A = \frac{n}{2} \left( \frac{\hat{\alpha}}{\alpha_\beta} - 1 \right) = \frac{n}{2} \left( \frac{\hat{\alpha}}{2\beta^2} - 1 \right). \]  

(15)

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

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

\[ G = \frac{\sum_i |Y_i|}{\sqrt{n \sum_i Y_i^2}} = \frac{\hat{\beta}}{\sqrt{\alpha}}, \]  

(16)

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

\[ T_f = n \log \left( \frac{\sqrt{\pi} G}{2} \right), \quad T_A = n \left( \sqrt{\frac{\pi}{2}} G - 1 \right). \]  

(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):
plim\(n^{-1}TA_f = \sqrt{\frac{\pi}{2}} + 1 \approx -0.1138,\) \hspace{1cm} (18)

plim\(n^{-1}TA_g = (1/2)(\pi/4 - 1) \approx -0.1073.\) \hspace{1cm} (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

\[ I_{fs}(\alpha, \beta) = E_{\alpha} \left( \log f(y; \alpha) - \log g(y; \beta) \right). \] \hspace{1cm} (20)

Assume that \(I_{fs}(\alpha, \beta)\) has a unique minimum at \(\beta_0(\alpha)\). Pesaran (1987) defined the closeness of \(H_g\) to \(H_f\) as

\[ C_{fs}(\alpha) = I_{fs}(\alpha, \beta_0(\alpha)). \] \hspace{1cm} (21)

Similarly, define the KLIC for \(g(y; \beta)\) against \(f(y; \alpha)\) (denote \(I_{sf}(\beta, \alpha)\)) and the closeness of \(H_f\) to \(H_g\) (denote \(C_{gf}(\beta)\)). Using \(C_{fs}(\alpha)\) 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), \(I_{fs}(\alpha, \beta)\) and \(I_{sf}(\beta, \alpha)\) are written as

\[ I_{fs}(\alpha, \beta) = -\frac{1}{2} \log(2\pi\alpha) + \log(2\beta) + \frac{1}{\beta}\sqrt{\frac{2\alpha}{\pi}} - 0.5, \] \hspace{1cm} (22)

\[ I_{sf}(\beta, \alpha) = \frac{1}{2} \log(2\pi\alpha) - \log(2\beta) + \frac{\beta^2}{\alpha} - 1. \] \hspace{1cm} (23)

Because \(\beta_0(\alpha) = \sqrt{2\alpha/\pi}\) and \(\alpha_0(\beta) = 2\beta^2\),

\[ C_{fs}(\alpha) = \log \left( \frac{2}{\pi} \right) + 1 - 0.04842, \] \hspace{1cm} (24)

\[ C_{gf}(\beta) = \log \left( \frac{\sqrt{\pi}}{2} \right) - 0.5 = 0.07236. \] \hspace{1cm} (25)

Because both \(C_{fs}(\alpha)\) 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 \(\bar{\alpha}_n\) be the asymptotic significance level of some test with a given sample size \(n\). The approximate slope is defined as \(\lim(-2n^{-1}\log\bar{\alpha}_n)\). 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 \(\text{plim}_n(n^{-1}N_j^2)\) and \(\text{plim}_n(n^{-1}NA_j^2)\), respectively. Therefore, from (10), (11), (14), and (15),

\[ \text{plim}_n(n^{-1}N_j^2) = \left( \frac{\log \left( \frac{\sqrt{\pi}}{2} \right)}{\frac{\pi}{2} - \frac{3}{2}} \right)^2 \approx 0.2061, \] \hspace{1cm} (26)

\[ \text{plim}_n(n^{-1}NA_j^2) = \left( \frac{\sqrt{\pi} - 1}{\frac{\pi}{2} - \frac{3}{2}} \right)^2 = 0.1828, \] \hspace{1cm} (27)

\[ \text{plim}_n(n^{-1}N_g^2) = \left( \frac{\log \left( \frac{\pi}{4} \right)}{\pi + \frac{3}{2}} \right)^2 \approx 0.05835. \] \hspace{1cm} (28)
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 $T_A$ 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

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

$$\text{plim}_n n^{-1} \text{NA}^2_g = \left( \frac{\pi}{4} - 1 \right)^2 \approx 0.04605. \quad (29)$$

$$\text{plim}_n n^{-1} \text{NA}^2_g = \left( \frac{\pi}{4} - 1 \right)^2 \approx 0.04605. \quad (29)$$
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


