Portable and powerful tests for normality

Carlos M. Urzúa

Tecnológico de Monterrey, Campus Ciudad de México*

This version: May 2007

ABSTRACT

Two new omnibus tests for normality are introduced in this paper. They are constructed using Pearson's measure of skewness together with Geary's measure of kurtosis. As shown by Monte Carlo simulation using a wide range of distributions, in most cases they overpower two popular omnibus tests, the Jarque-Bera and D'Agostino-Pearson tests. Furthermore, the new tests have satisfactory nominal sizes even for small samples.

Keywords: Normality test; Omnibus test; Univariate Normality;
JEL classification: C10, C20

*E-mail: curzua@itesm.mx. Address: Calle del Puente 222, 14380 Tlalpan, México, DF, México. Phone: (52-55) 5483-1878. Fax: (52-55) 5483-1882.

1. Introduction

If it is not size-corrected, the test for normality proposed by Jarque and Bera (1980) behaves badly in the case of small and medium size samples. This fact was already known to Bowman and Shenton (1975), who first introduced that test in the statistics literature. Although with some delay, the same fact has been widely documented in the economics literature as well (see, e.g., Lawford, 2005, Poitras, 2006, and Urzúa, 1996). Yet, almost all applied economists continue to use the Jarque-Bera (JB) test without any size corrections. Granted, many of them use it simply because it is the test of normality more readily available in econometrics softwares, but we surmise that there are also other reasons that can account for the popularity of the JB test. First, it is an omnibus test that detects departures from normality due to both skewness and kurtosis, and we economists usually don't hold priors about the alternatives to normality (except for some areas such as finance). Second, the JB test has optimum (asymptotic) power properties when the alternatives to the normal are members of several important families of distributions (see, e.g., Jarque and Bera, 1980, and Urzúa, 1997). Finally, and this is only a personal opinion, since the JB test is biased in favor of the null hypothesis of normality, few applied economists have to loose their sleep after applying it.

Two better-behaved omnibus tests for normality are introduced in this paper. They make use of Pearson's measure of skewness, as the JB test does, but in the case of kurtosis they use a different measure due to Geary (1935). As shown by Monte Carlo simulation using a range of distributions, in most cases the new tests overpower not only the JB test, but also the D'Agostino-Pearson K^2 test (see D'Agostino et al., 1990), arguably the most popular test for normality in the statistics literature. Furthermore, in contrast to those two, the new tests are not noticeable biased against or in favor of the null hypothesis of normality in the case of small samples.

2. New tests for normality

Karl Pearson (1895) was the first to recommend the use of the standardized third and fourth moments when testing for normality. Four decades later, Geary (1935) introduced a new measure of kurtosis: $\alpha = \tau / \sigma,$ where $\tau = E(|X - \mu|)$ and σ is the standard deviation of the population. This measure was meant to be an alternative to Pearson's kurtosis measure β_2.

Correspondingly, given a sample size of order *n*, Geary proposed a new test of normality using the statistic

$$a = \frac{\sum |x_i - \bar{x}|}{\sqrt{n \sum (x_i - \bar{x})^2}}$$

Under the null hypothesis of normality, *a* is approximately equal to $\sqrt{2/\pi}.$ Unfortunately, a number of simulation studies show that the power of that test is inferior, in the case of most alternatives, to Pearson's test statistic $b_2 = m_4 / m_2^2,$ where $m_i = \sum (x_j - \bar{x})^i / n,$ when this last test is properly size-adjusted.

Yet, the difference in power is not too much, while the convergence to normality of *a* is much faster than the one of b_2. Thus, we propose in this paper to consider new omnibus tests of normality based, on the one hand, on Geary's test (actually a transformation of it), together with Pearson's skewness test statistic $\sqrt{b_1} = m_3 / m_2^{3/2}$.

The following proposition, proved in the Appendix, presents two alternatives for that end. One is a common Chi-square test, while the other is a Tukey-type test (an alternative to the one first suggested by Cox and Hinckley, 1974, p. 72).

Proposition 1. *For a normal sample of size n, the statistic*

$$Q = z_1^2 + z_2^2 \,, \text{ where } z_1 = \frac{\sqrt{b_1}}{\sqrt{6/n}} \text{ and } z_2 = \frac{a - \sqrt{2/\pi}}{\sqrt{(1-3/\pi)/n}} \,,$$

is asymptotically distributed as a chi-square with two degrees of freedom.

Likewise,

$$T = \max\left\{ |z_1| \,, |z_2| \right\}$$

is asymptotically distributed as an arctangent with parameter equal to 1 (using the terminology in Pollastri and Tornaghi, 2004), with density

$$f(t) = \frac{4e^{-t^2/2}}{\pi} \int_0^t e^{-z^2/2} dz \,.$$

Proof. See Appendix.

To make operational the results in Proposition 1, one would like to transform both z_1 and z_2 to speed up the convergence to normality. In the case of the former, it is common to use a transformation due to D'Agostino (1970a). However, we have found in our context that it is equally effective and much simpler to just replace the asymptotic variance with the exact variance (given in, e.g., Urzúa, 1996). In the case of Geary's test, we have found that the logarithmic transformation due to Bonett and Seier (2002) is quite helpful.

Thus, the first omnibus test to be proposed in this paper is:

$$ab = \frac{\sqrt{b_1}^2}{d} + \frac{(w-3)^2}{e} \quad , \quad \text{where } d = \frac{6(n-2)}{(n+1)(n+3)} \; , \; e = \frac{3.54}{(n+2)}$$

and $w = -6\ln(a)/\ln(\pi/2)$. The *ab* test is approximately distributed as a chi-square with two degrees of freedom.

The second test, *abt*, is defined as

$$abt = \max\left\{\frac{\sqrt{b_1}}{\sqrt{d}}, \frac{(w-3)}{\sqrt{e}}\right\}$$

In order to calculate the critical points to reject the null hypothesis of normality, one uses the fact that the two components are independent as follows: Given an α percent level of significance for the *abt* test, one finds the critical point of the normal distribution that corresponds to a $100\left(1-\sqrt{1-\alpha}\right)$ percent level. For instance, for a level of significance of 5%, the critical point is 2.236, which corresponds, in the case of the normal, to a level of significance of $100\left(1-\sqrt{.95}\right)$ percent.

3. Power of the tests

This section presents a Monte Carlo exercise that compares the power of the *ab* and *abt* tests with the JB test and the K^2 test. There are six alternatives to the normal distribution: uniform, Laplace, Student's t (with five degrees of freedom), heteroscedastic normal, and stable (with parameters 1.9 and 1).[1] Table 1 presents the estimated power, using 10000

[1] In the case of the uniform distribution, we used for the simulation Marsaglia's KISS-Monster algorithm, as implemented in GAUSS. For the rest of the distributions (including the normal), the generation of pseudo-random numbers was made using the author's GAUSS procedures GRAN1-5 that are publicly available at

replications and without any size correction. As can be appreciated from there, the results are mostly in favor of the new tests. Given these results, it seems worth to extend the tests given here to cover the case of time series observations, as well as the multivariate extensions (Urzúa, 2007).

IDEAS (http://ideas.repec.org).

Appendix A

Proof of Proposition 1. To prove Proposition 1, we first note that the asymptotic distribution of both z_1 and z_2 is a standard normal. That is well known in the first case. In the latter case, Geary (1936) shows that the exact mean and variance of the asymptotic normal distribution is given by

$$E(a) = \sqrt{\frac{2}{\pi}} \exp\left(-\frac{1}{4n} + \frac{1}{24n^3} - \frac{1}{20n^5} + \frac{17}{112n^7} - \ldots \right),$$

$$\mathrm{var}(a) = \frac{1}{n+1}\left(1 + \frac{2}{\pi}\sqrt{n^2-1} + \frac{2}{\pi}\sin^{-1}\frac{1}{n} \right) - E(a)^2 .$$

By approximating these two expressions, the asymptotic mean and variance are found to be $\sqrt{2/\pi}$ and $(1-3/\pi)/n$ (this last result was first derived by D’Agostino, 1970b, but he did not express the asymptotic variance in terms of π).

Next we note that $\sqrt{b_1}$ and a are asymptotically independent. This is so because, using Theorem 1 in Gastwirth and Owens (1977), they are asymptotically distributed according to a bivariate normal with covariance equal to zero. Thus, Q is asymptotically distributed as a Chi-square with two degrees of freedom.

On the other hand, using once again the asymptotic independence and normality of the two random variables,

$$P\{T \le t\} = P\{-t \le X_1 \le t\}P\{-t \le X_2 \le t\} = 4P\{0 \le X_1 \le t\}^2 = \frac{2}{\pi}\left(\int_0^t e^{-z^2/2} dz \right)^2 .$$

After deriving this last expression with respect to t, one obtains the density in the text.

References

Bonnet, D.G. and E. Seier, 2002, A test of normality with high uniform power, Computational Statistical & Data Analysis 40, 435-445.

Bowman, K.O. and L.R. Shenton, 1975, Omnibus test contours for departures from normality based on $\sqrt{b_1}$ and b_2, Biometrika 62, 243-250.

Box, G.E.P., 1953, A note on regions for tests of kurtosis, Biometrika 40, 465-468.

Cox, D.R. and D.V. Hinkley 1974, Theoretical Statistics. (Chapman Hall, London).

D'Agostino, R.B., 1970a, Transformation to normality of the null distribution of g_1, Biometrika 57, 679-681.

D'Agostino, R.B., 1970b, Simple portable test of normality: Geary's test revisited, Psychological Bulletin 74, 138-140.

D'Agostino, R.B., A. Belanger and R.B. D'Agostino, Jr., 1990, A suggestion for using powerful and informative tests of normality, American Statistician 44, 316-321.

Gastwirth, J.L. and M.E.B. Owens, 1977, On classical tests of normality, Biometrika 64, 135-139.

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

Geary, R.C., 1936, Moments of the ratio of the mean deviation to the standard deviation for normal samples, Biometrika 28, 295-305.

Jarque, C.M. and A.K. Bera, 1980, Efficient tests for normality, homoscedasticity and serial independence of regression residuals, Economics Letters 6, 255-259.

Lawford, S., 2005, Finite sample quantiles of the Jarque-Bera test, Applied Economics Letters 12, 351-354.

Pearson, K., 1895, Contributions to the mathematical theory of evolution II: skew variation in homogeneous material, Philosophical Transactions of the Royal Society of London, A, 186, 343-414.

Poitras, G., 2006, More on the correct use of omnibus tests for normality, Economics Letters 90, 304-309.

Pollastri, A. y F. Tornaghi, 2004, Some properties of the arctangent distribution, Statistica & Applicazioni 2, 1-16.

Thode, H.C., 2002, Testing for Normality. (Marcel Dekker, New York).

Urzúa, C.M., 1996, On the correct use of omnibus tests for normality, Economics Letters 53, 247-251. Erratum 1997, vol. 54, 301.

Urzúa, C.M., 1997, Omnibus tests for multivariate normality based on a class of maximum entropy distributions, Advances in Econometrics 12, 341-358.

Urzúa, C.M., 2007, A test for multivariate normality of observations and time series, EGAP working paper, Tecnológico de Monterrey, Campus Ciudad de México.

Estimated Power of the Tests

(5% significance level)

	ab	*abt*	*JB*	*K*
Laplace				
n=20	**0.336**	0.331	0.214	0.302
n=50	**0.653**	0.652	0.512	0.507
n=100	**0.89**	0.89	0.778	0.735
Student's t				
n=20	**0.221**	0.216	0.161	0.22
n=50	**0.417**	0.412	0.392	0.403
n=100	**0.619**	0.617	0.629	0.597
Het. Normal				
n=20	**0.157**	0.152	0.102	0.157
n=50	**0.264**	0.261	0.226	0.238
n=100	**0.41**	0.409	0.386	0.356
Stable				
n=20	0.155	0.15	0.125	**0.171**
n=50	0.294	0.29	0.291	**0.31**
n=100	0.462	0.457	0.479	**0.485**