# A Normality Test for the Mean Estimator

Michael K. Andersson
*Working Paper Series in Economics and Statistics No. 310*
Stockholm School of Economics

March 4, 1999

**Abstract**

This note proposes a tool to investigate and demonstrate the adequacy of the central limit theorem in small samples. The suggested testing procedure provides a method to investigate if the mean estimator is approximately normally distributed, given data and sample size at hand. This is important when deciding upon inference procedure, i.e. is parametric inference possible or do we have to settle for nonparametrics. The procedure is well-suited for teaching in undergraduate statistics classes.
KEY WORDS**:** Central limit theorem; Bootstrap; Empirical distribution.

## 1 Introduction

To make statistical inference, several assumptions about the data must be fulfilled. When performing a test for the mean we often assume that the sample mean estimator is approximately normally distributed, due to large enough sample size and the central limit theorem. This note introduces a procedure to test, given a sample of $n$ observations, if this is the case. We provide not only a test of the validity of the central limit theorem, the procedure and its results may be employed to demonstrate the distributional convergence. Furthermore, the test is very useful in empirical applications and is well-suited for teaching. The paper unfolds as follows: Section 2 explains the testing procedure and some examples are given in Section 3.

## 2 The Testing Procedure

The role of the central limit theorem in statistical inference is an important one. The central limit theorem states the following; assume that the stochastic variables $X_1, ..., X_n$ are independently and identically distributed with mean $\mu < \infty$ and variance $\sigma^2 < \infty$. The distribution of the ratio

$$Z_n = \frac{\bar{X} - \mu}{\hat{\sigma}/\sqrt{n}}, \tag{1}$$

where the random variables $\bar{X}$ and $\hat{\sigma}$ are the sample mean and standard deviation, then approaches a standard normal when the sample size $n$ goes to infinity. The result of the central limit theorem motivate the use of parametric inference, but if the sample is not big enough and the variable is nonnormal, we must do nonparametrics. For practical purposes however, the theorem says nothing about the sample size required for approximate normality.

A test for normality, given a sample on $X$, may be performed by the Jarque-Bera statistic

$$JB = n\left\{\frac{(skew)^2}{6} + \frac{(kurt-3)^2}{24}\right\}, \tag{2}$$

which distribution is tabulated by Bera and Jarque (1981). The test can not be used on the sample mean since we cannot calculate the skewness and kurtosis from the only available sample point. There are of course other tests for normality, e.g. the Kolmogorov-Smirnov test and the $\chi^2$ goodness of fit test, but they are also not feasible. However, the following procedure may easily be applied to any normality test.

Our task is to estimate the distribution of a test for normality of the sample mean. Assume that

$$\mathbf{x} = (x_1, x_2..., x_n)$$

are realizations of a random sample and let $\hat{F}$ denote the empirical distribution of $\mathbf{x}$. A bootstrap resample $\mathbf{x}^*$ is then defined to be a random sample drawn from $\hat{F}$, that is the $n$ members of $\mathbf{x}^*$ are drawn with replacement from the collection $\mathbf{x}$. We can create a large number, say $B$, of bootstrap samples and calculate the statistic of interest $\left(\text{here } \bar{X}\right)$ for each of the resamples.

Thus, we may retrive $B$ estimates of the sample mean $(\bar{x}_1^*, \bar{x}_2^*, ..., \bar{x}_B^*)$, which in turn allow us to estimate the skewness and kurtosis of $\bar{X}$, denoted respectively by $sk^*(\bar{x})$ and $k^*(\bar{x})$. Furthermore, the bootstrap procedure provides a value of the statistic (2) for the sample mean accordingly,

$$JB(\bar{x}) = B\left\{\frac{\left[sk^*(\bar{x})^2\right]}{6} + \frac{\left[k^*(\bar{x}) - 3\right]^2}{24}\right\}.$$

For further readings on the bootstrap please consult Efron and Tibshirani (1993).

# 3 An example

In this example we construct three artificial samples of size $m = 500$. The first sample is drawn from a standard normal distribution, the second from a $t(5)$ distribution, having excess kurtosis, and the third sample is generated by a skewed $\chi^2(1)$ distribution.[1] For each sample we use the last $n$ observations, for which we create $B = 1000$ bootstrap samples and calculate the skewness, kurtosis and $JB$ statistic. By letting $n = 5, ..., 500$ we compute the quantities using different amount of information (sample sizes).

The sample skewness and kurtosis is presented in Figure 1. For normally distributed data the sample mean is also normally distributed, and subsequently the skewness and kurtosis fluctuate about the (expected) values zero and three. The normality is also supported by the $JB$ statistic, see results in Figure 2. When the data are symmetric, but exhibit excess kurtosis, the skewness and kurtosis of the sample mean are also close to zero and three. Despite a few rejections, the $JB$ statistic indicates normality of the mean estimator already at small sample sizes. The convergence of the mean to a normal distribution is much slower if the observations origin from a skewed distribution. The skewness is inherited by the sample mean and the kurtosis is (for small samples) notably larger than three. The behavior of the $JB$ statistic supports the slow convergence and stress the need for large samples when the true data generating process is skewed.

The results above are known to statisticians, but the procedure provides a method for the testing and demonstration of one of the most fundamental

[1] A Fortran program for the excercise is available on request from the author.

assumptions in statistical inference. The procedure is so simple that it can easily be used in under-graduate statistics classes.

# Acknowledgements

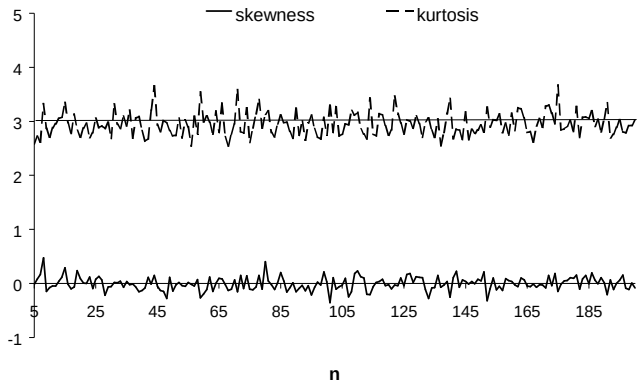
Financial support from the HSFR, the Swedish Research Council for the Humanities and Social Sciencies, is gratefully acknowledged.

# References

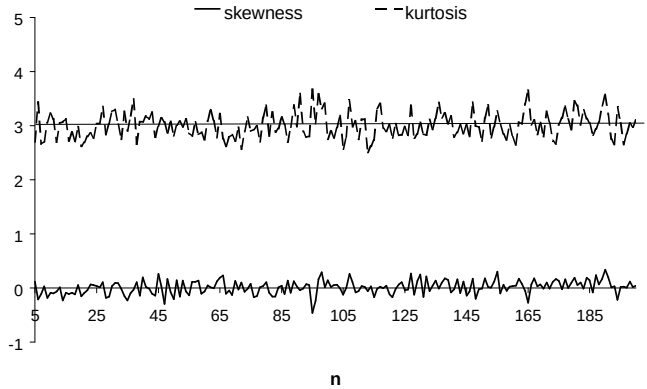
Bera, A.K. and C.M. Jarque C.M. (1981) An efficient large-sample test for normality of observations and regression residuals, *Working Papers in Economics and Econometrics,* 40, Australian National University.

Efron, B. and R.J. Tibshirani, (1993), *An introduction to the bootstrap.* New York: Chapman and Hall.

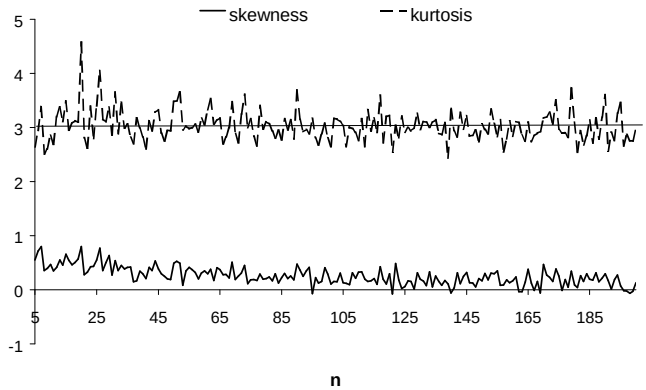
Figure 1: Estimated skewness and kurtosis of the sample mean.




The data are N(0,1) distributed.



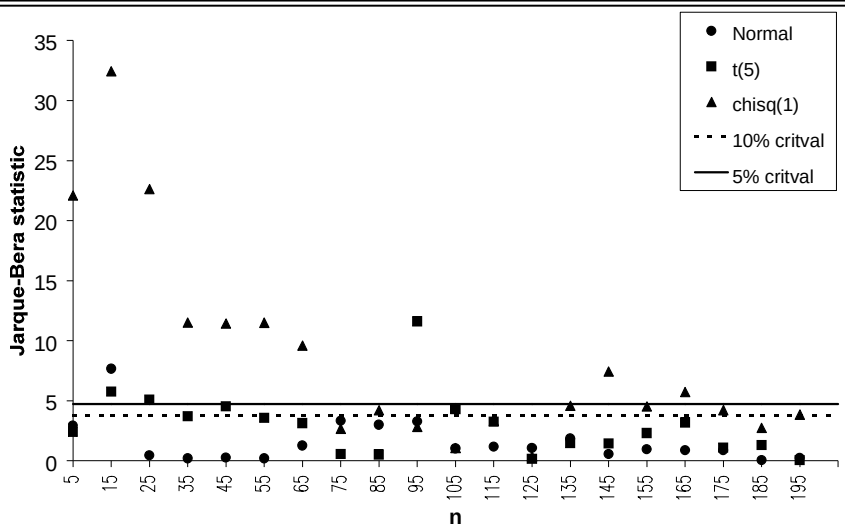

The data are $t\,(5)$ distributed.




The data are $\chi^2\,(1)$ distributed.

The figures present the estimated skewness and kurtosis of the sample mean. The bootstrap involves 1,000 replicates.

Figure 2: Bootstrap Jarque-Bera normality test statistics for the sample mean.




The figure presents the value of the $JB$ statistic associated with the skewness and kurtosis reported in Figure 1.