SEMI–MOMENTS BASED TESTS OF NORMALITY AND THE EVOLUTION OF STOCK RETURNS TOWARDS NORMALITY

FRANÇOIS DESMOULINS-LEBEAULT

Abstract. Testing for normality is of paramount importance in many areas of science since the Gaussian distribution is a key hypothesis in many models. As the use of semi–moments is increasing in physics, economics or finance, often to judge the distributional properties of a given sample, we propose a test of normality relying on such statistics. This test is proposed in three different versions and an extensive study of their power against various alternatives is conducted in comparison with a number of powerful classical tests of normality. We find that semi–moments based tests have high power against leptokurtic and asymmetric alternatives. This new test is then applied to stock returns, to study the evolution of their normality over different horizons. They are found to converge at a "log-log" speed, as are moments and most semi–moments. Moreover, the distribution does not appear to converge to a real Gaussian.

1. Introduction

The problem of normality testing is well known and has generated plenty of attention from researchers, see Mardia (1980) and D'Agostino and Stephens (1986). This is because a lot of classical optimal procedures were developed based on the normality assumption. However, researchers soon realised that this assumption was not always satisfied. Three approaches can be taken to deal with non-normality of data. The first approach is transforming the data to normality so that the classical procedures could still be used. The second approach is the use of nonparametric procedures. The third is to use robust procedures that are less sensitive to deviation from normality, especially tail behaviour. Each of the three comes with strengths and weaknesses and there is no consensus on which is the best approach in a given case.

The role of normality testing is not just to see if the data are well approximated by the normal distribution; but also to provide information on the deviation from normality. This information would then guide the researchers to the best approach in dealing with the non-normality of their data. One of the illustrations of such an approach can be found in the study of stock returns distributions, which first came into focus with Mandelbrot (1963) and Fama (1965). These returns were found generally non-Gaussian, yet seem to converge to normality.

In this research, it will be assumed that one is testing normality because the user wishes to fit a financial model (either a classical Gaussian model or an alternative model using asymmetry and/or heavy tails). Suppose the data collected, $x_1, x_2, \cdots, x_n$, represent an independent and identically distributed (iid) random sample of size n from a population with probability density function $f(x)$ and cumulative density function (cdf) $F(x)$. Let Φ be the cdf of x that is normally distributed with

CEREG (CNRS UMR 7088), Université de Paris Dauphine.

unknown mean and variance. The null hypothesis in this problem of testing for normality is $H_0 : F(x) = \Phi(x)$ and the alternative hypothesis simply states H_0 is false. Hence only omnibus tests will be considered in this article. Here, omnibus refers to the ability of a test to detect any deviation from normality with an adequate sample size. However, as we consider financial returns, which are supposed to be continuously distributed, our definition of omnibus tests is restricted to continuous distributions, hence the alternatives cannot be a discrete distribution, like the binomial.

In this problem, the focus is on failing to reject H_0 so that the conclusion is that the data come from a normal distribution. As noted by D'Agostino and Stephens (1986), this distinguishes normality testing from most statistical tests. Also, with a vague alternative hypothesis, they commented that 'the appropriate statistical test will often be by no means clear and no general Neyman-Pearson type (test) appears applicable'. Hence, it will be unlikely to have a single test that will have power superior to their alternatives. This fact is the origin of the vast number of tests of normality.

Indeed, there is an impressive number of normality tests in the literature. Major power studies done by Shapiro et al. (1968) and Pearson et al. (1977) have not arrived at a definitive answer; but a general consensus has been reached about which tests are powerful. Pearson (1900) chi-squared test, which is possibly the oldest, is not very sensitive. Data are grouped and compared to the expected counts under normality. Since information is lost in the grouping and this test is not specially tailored for the normal distribution, the conclusion is not surprising.

At the opposite end of the spectrum of normality tests, the Bowman-Shenton test is quite powerful. Although it is not a strictly speaking omnibus test (a distribution may have the same third and fourth moment than a Gaussian and yet be quite different) it has an advantage in terms of practice. This test is based on moments and such statistics may be used to judge the type of departure from normality and to model an alternate distribution, e.g. with an Edgeworth expansion or a Cornish-Fisher expansion.

Using the same sort of thought, we propose three tests based on semi–moments. These tests may be used to detect departures from normality in data but at the same time provide a mean to understand more about these departures. The use of semi–moments in financial markets analysis seems particularly justified, especially in the light of such findings as those of Fishburn (1977) and Tversky and Kahneman (1974) and Kahneman and Tversky (1979). In the following sections we will introduce the main families of normality tests, present our semi–moments based tests and some of their properties, then conduct a detailed power study of some of the most powerful tests in the literature. Eventually, we apply our tests to stock returns and study the influence of time scales on their normality.

2. Existing Normality Tests

There are many available tests of normality, probably more than 50[1]. Testing for normality has been an important question in statistics since the beginning of the XXth century, for normality is hypothesised in many models. We will present in this section the most important existing tests, focusing on those which have been found most powerful. To clarify our exposition of the existing tests, we regroup them in broad families. Normality tests can be classified in three categories: those based on

[1]For a detailed survey of the literature, see Mardia (1980) and D'Agostino and Stephens (1986).

the empirical distribution function (EDF), those based on regression techniques, and finally, the tests based on moments.

2.1. **Distance/EDF Tests - Anderson-Darling A^2.** This category of tests can be traced back to the χ^2 goodness of fit test developed by Pearson (1900). EDF or distance tests are a broad class of normality tests that are based on a comparison between the EDF, $F_n(x_{(i)}) = \frac{i}{n}$, and the hypothesised distribution under normality, Z_i , as defined by:

$$Z_i = \Phi\left(\frac{x_{(i)} - \bar{x}}{s}\right),$$

where $\bar{x} = (1/n)\sum_{i=1}^{n} x_i$ and $s^2 = 1/(n-1)\sum_{i=1}^{n}(x_i - \bar{x})^2$. Stephens (1974) provided versions of the EDF tests with unknown μ and σ^2. EDF tests can be further classified into those involving either the supremum or the square of the discrepancies, $F_n(x_{(i)}) - Z_i$. The most well known EDF tests involving the supremum is the Kolmogorov-Smirnov statistic $K = \max(D^+, D^-)$ where $D^+ = \sup_i(i/n - Z_i)$ and $D^- = \sup_i(Z_i - (i-1)/n)$.

EDF tests involving the square of the discrepancies are known as those from the Cramér-von Mises family with the general form:

$$CvM = n\int\left(\frac{i}{n} - Z_i\right)^2 \psi(Z_i)\mathrm{d}Z_i,$$

where $\psi(Z_i)$ is the weighting function. If $\psi(Z_i) = 1$, that is the Cramér-von Mises statistic itself, ω^2. For the Anderson- Darling statistic, A^2,

$$\psi(Z_i) = \frac{1}{Z_i(1 - Z_i)}.$$

This choice of $\psi(Z_i)$ gives emphasis to tail values and the computational form is given by:

$$A^2 = -\frac{1}{n}\sum_{i=1}^{n}((2i-1)[\ln(Z_i) + \ln(1 - Z_{n-i+1})]) - n.$$

Stephens (1974) extensively studied these tests. Moreover, he gave corrections increasing the power of many different EDF tests, and found that A^2 has the highest power among them when used with his corrections. The asymptotic distribution is known and it was found that the critical values for finite samples quickly converge to their asymptotic values for $n > 5$.

2.2. **Regression/Correlation Tests - Shapiro-Wilk W.** The main idea behind these tests is normal probability plotting. Normal probability plotting is a graphical technique to determine the normality of the data by looking for linearity in a plot of the ordered observations $x_{(i)}$ against the expected values of standard normal order statistics, m_i. Formal determination of the linearity uses regression or correlation techniques, hence the name of this group of tests. If $x_{(i)}$ is indeed normal, then the slope would give the standard deviation of x_i, σ, and the intercept, the mean of the x_i's, μ. Since the ordered observations are not independent, let $\mathbf{V} = (v_{ij})$ be the $n \times n$ covariance matrix, $x' = (x_1, x_2, \cdots, x_n)$ and $m' = (m_1, m_2, \cdots, m_n)$. The best linear unbiased estimators of the slope and intercept, using generalised least squares, are:

$$\hat{\sigma} = \frac{m'\mathbf{V}^{-1}x}{m'\mathbf{V}^{-1}m} \text{ and } \hat{\mu} = \bar{x}.$$

The usual symmetric estimate of the variance regardless of the distribution of x_i is given by s^2 . Then the Shapiro and Wilk (1965) W statistic is defined as:

$$W = \frac{K\hat{\sigma}^2}{(n-1)s^2} = \frac{a'x}{(n-1)s^2} = \frac{\left(\sum_{i=1}^{n} a_i x_{(i)}\right)^2}{\sum_{i=1}^{n}(x_{(i)} - \bar{x})},$$

where

$$a' = (a_1, a_2, \cdots, a_n) = m'\mathbf{V}^{-1}\left[(m'\mathbf{V}^{-1})(\mathbf{V}^{-1}m)\right]^{-\frac{1}{2}},$$

$$K = \frac{m'\mathbf{V}^{-1}m}{m'\mathbf{V}^{-1}\mathbf{V}^{-1}m}.$$

W compares the ratio of two estimates of variance,$\hat{\sigma}^2$ and s^2, apart from a normalising constant, K, and $(n-1)$. If the distribution of x_i is normal, then W will be close to 1. Otherwise, W is less than 1. The critical values of W are tabulated up to sample sizes of 50. However, values for $\{a_i\}$ are also needed to carry out this test. For larger sample sizes, Shapiro and Francia (1972) noted that the ordered observations, as n increases , may be treated as independent (i.e. $v_{ij} = 0$ for $i \neq j$) . Treating $\mathbf{V}$ as an identity matrix, W can be extended for n larger than 50 by

$$W' = \frac{\left(\sum_{i=1}^{n} m_i x_i\right)^2}{\sum_{i=1}^{n}(x_{(i)} - \bar{x})\sum_{i=1}^{n} m_i^2}$$

Values of $\{m_i\}$ are available from Harter (1961) up to sample sizes of 400. However, two tables are still needed to carry out this test. A further modification was suggested by Weisberg and Bingham (1975) that uses this approximation:

$$m_i \approx \Phi^{-1}\left(\frac{i - \frac{3}{8}}{n + \frac{1}{4}}\right)$$

due to Blom (1958). This approximation was shown to be close even in small samples, and the null distribution of W was practically identical to W'. This simplifies the computation of the test statistics since separate values for m_i need not be kept. Royston (1982) used another approximation suggested by Shapiro and Wilk (1965) for $\{a_i\}$ and applied the following normalising transformation to W:

$$y = (1 - W)^{\lambda} \text{ and } z = (y - \mu_y)/\sigma_y$$

where z is standard normal and λ , μ_y and σ_y are functions of n. λ is estimated by maximising the correlation between certain empirical quantiles of W and the corresponding standard normal equivalent with weights given according to the variance of a normal quantile. The relation between μ_y and σ_y and n is then determined by applying λ to simulated values W. The normalising transformation producing W^* does away with any special tables, besides the standard ones, needed to find the critical values of W. However, the first version of this approximation was not entirely flawless and was therefore corrected and enhanced by Royston (1993b), who published as well a second version of his algorithm the same year.

2.3. **Moments Tests - Bowman-Shenton K^2.** Since the concepts of skewness and kurtosis can be used to differentiate between distributions, one of the most important classes of normality tests is based on these moments. The standardised coefficients of skewness, $\sqrt{\beta_1}$, and kurtosis, β_2 are defined as

$$\sqrt{\beta_1} = \frac{\mu_3}{\sigma^3} \text{ and } \beta_2 = \frac{\mu_4}{\sigma^4}$$

where μ_i is the ith central moment.

Skewness refers to the symmetry of a distribution. For a symmetric distribution like the normal, $\sqrt{\beta_1} = 0$. A distribution that is skewed to the right has $\sqrt{\beta_1} > 0$ while one that is skewed to the left has $\sqrt{\beta_1} < 0$.

Kurtosis refers to the flatness or 'peakedness' of a distribution. The normal distribution has $\beta_2 = 3$ and is used as a reference for other distributions. A leptokurtic distribution is one that is more peaked and with heavier tails than the normal, resulting in $\beta_2 > 3$. A platykurtic distribution has a flatter distribution with shorter tails than the normal, hence $\beta_2 < 3$.

The sample skewness, $\sqrt{b_1}$, and kurtosis, b_2 , are defined as:

$$\sqrt{b_1} = \frac{m_3}{m_2^{3/2}} \text{ and } b_2 = \frac{m_4}{m_2^2}$$

where m_i is the ith sample moment. Since the moments of $\sqrt{b_1}$ and b_2 are known, their distributions have been approximated using Pearson or Johnson curves. The critical values for the normality tests of skewness and kurtosis are tabulated in Pearson and Hartley (1972) for selected values of $n \geq 25$ at $\alpha = 0.02$ and 0.10. Normalising transformations have been found for $\sqrt{b_1}$ and b_2 by D'Agostino (1970) and D'Agostino and Pearson (1973), respectively. $Z(\sqrt{b_1})$ and $Z(b_2)$ denote the resulting approximate standardised normal variables. D'Agostino and Pearson suggested combining $\sqrt{b_1}$ and b_2 in the following way:

$$K^2 = Z^2(\sqrt{b_1}) + Z^2(b_2)$$

where K^2 is distributed as χ_2^2 since it is the sum of the squares of 2 standardised normal equivalent deviates. However, they assumed that the squared standardised normal equivalent deviates were independent, which is only asymptotically true. The error can be important, especially for small sample sizes. Using simulations, Bowman and Shenton (1975) obtained 90%, 95% and 99% contours for K^2, for sample sizes between 20 and 1000. Carrying out this test would then only require calculating $\sqrt{b_1}$ and b_2, selecting the appropriate contour, and determining if $(\sqrt{b_1}, b_2)$ falls within the contours. If it does not, then normality is rejected.

This same approach, albeit without the normalising transformations and the tabulation of the contours taking dependency into account, is used again in the Jarque and Bera (1981) test. Therefore, this last test is only efficient for large samples.

3. Semi-Moments Based Tests of Normality

3.1. **The Test Statistics.** Since semi–moments complement moments and provide additional information, at odd and even orders they contain informations on the shape and location of the distribution, it seems interesting to use them in normality tests. Their main advantage over classical moments, in the context of testing for normality, comes from the fact that they allow to distinguish tail asymmetries at a relatively low order. Therefore we expect that the use of semi–moments in tests of normality can increase the power of the moments based tests, especially against asymmetric leptokurtic alternatives.

We propose three different test statistics, one focusing more on the left tail asymmetries, one more dedicated to right tail asymmetries. The last statistic is more neutral as to the location of the asymmetries.

As we have noticed in our study of the sampling properties of the semi–moments, there is a degree of dependence between the estimators of the semi–moments. Therefore, in order to minimise the perturbations caused by this dependence, we chose to use cross selected semi–moments. That is, we combine the left (right) semi–skewness to the right (left) semi–kurtosis, respectively. Moreover, to limit noises in the estimations of the critical values, we standardised the estimators of the semi–moments used.

The first statistic proposed here is expected to perform best against positively skewed alternatives. It is as follows :

$$\Delta_1 = \frac{\left(\widehat{m}_3^- + [n/(n+1)]\sqrt{2/\pi}\right)^2}{\sigma_{ss}} + \frac{\left(\widehat{m}_4^+ - [(n-1)/(n+1)]3/2\right)^2}{\sigma_{sk}}, \tag{3.1}$$

with:

$$\sigma_{ss} = \frac{1}{n}1.743795 - \frac{1}{n^2}10.062152,$$

and

$$\sigma_{sk} = \frac{1}{n}21.558373 - \frac{1}{n^2}209.049576.$$

In these expressions, as before, $\widehat{m}_3^-$ is the estimator of the left semi–skewness, $\widehat{m}_4^+$ the estimator of the right semi–kurtosis and n is the number of observations in the sample tested. Following Fisher (1929) we believe that the variance of the estimators of semi–moments on Gaussian samples are a polynomial in $1/n$, the inverse number of observations, of order two. The estimation of these variances, σ_{ss} for semi–skewness and σ_{sk} for semi–kurtosis, is detailed in Appendix A.

Using the same notation, the second statistic, expected to perform best against negatively skewed alternatives, is given by:

$$\Delta_2 = \frac{\left(\widehat{m}_3^+ - [n/(n+1)]\sqrt{2/\pi}\right)^2}{\sigma_{ss}} + \frac{\left(\widehat{m}_4^- - [(n-1)/(n+1)]3/2\right)^2}{\sigma_{sk}}, \tag{3.2}$$

The last statistic we propose has a structure that makes it generally equal to either Δ_1 or Δ_2, depending on the properties of the sample. However, since it uses maximums, its distribution is further away from the underlying χ^2 and it may have, in some cases, a structure with parallel semi–moments (e.g. right semi–skewness and right semi–kurtosis), and therefore it may prove less powerful against certain alternatives. This statistic is as follows:

$$\begin{aligned}\Delta_3 = & \frac{\left(\max(\left|\widehat{m}_3^+\right|, \left|\widehat{m}_3^-\right|) - [n/(n+1)]\sqrt{2/\pi}\right)^2}{\sigma_{ss}} + \\ & + \frac{\left(\max(\left|\widehat{m}_4^+\right|, \left|\widehat{m}_4^-\right|) - [(n-1)/(n+1)]3/2\right)^2}{\sigma_{sk}},\end{aligned} \tag{3.3}$$

As explained before, the exact sampling distribution of semi–moments is as yet unknown, even the moments of it. Moreover, the non linear dependency structure between semi–moments implies that, to construct the distribution of a statistic comprising semi–skewness and semi–kurtosis, we need the multivariate distribution of all these estimators. Therefore, we tabulated critical values for the different test statistics we propose. These values are based on Monte Carlo simulations of 100 000 samples of the 34 sizes already used in the study of the sampling properties of the semi–moments. The resulting critical values for the three statistics are presented in tables 1, 2 and 3[2].

Table 1. Percentage points of the distribution of Δ_1

	Confidence Levels							
sample size	0.20	0.15	0.10	0.05	0.025	0.01	0.005	0.001
25	2.5860	3.3347	4.6210	7.4369	11.1187	17.5347	23.7369	40.3180
50	2.6633	3.4104	4.5707	7.1009	10.3100	15.9821	21.4919	39.4102
75	2.7630	3.4913	4.6714	7.1884	10.1869	15.3459	20.3122	36.7530
100	2.8178	3.5546	4.6726	6.9474	9.7999	14.5279	19.1375	34.0987
250	2.9744	3.6933	4.8042	6.9339	9.5498	13.3710	16.3059	26.4350
500	3.0833	3.7890	4.8417	6.8699	9.0618	12.2753	15.3635	25.2615
1000	3.1359	3.8159	4.8489	6.7446	8.8277	11.7667	14.3612	22.0249
5000	3.1884	3.8566	4.8602	6.6731	8.5529	11.2425	13.5729	18.6675

Table 2. Percentage points of the distribution of Δ_2

	Confidence Levels							
sample size	0.20	0.15	0.10	0.05	0.025	0.01	0.005	0.001
25	2.6050	3.3547	4.5963	7.4030	11.1573	17.4214	23.1066	43.1937
50	2.6658	3.3952	4.5744	7.0782	10.2906	15.8581	21.7465	41.5505
75	2.7565	3.4757	4.6475	7.1586	10.1378	15.4407	20.8761	39.9846
100	2.8129	3.5454	4.6754	7.0189	9.7665	14.2651	18.9436	31.9495
250	2.9848	3.6849	4.7957	6.9735	9.4567	13.0405	16.5189	27.2613
500	3.0892	3.7634	4.8131	6.8358	9.1231	12.4003	15.3754	23.5589
1000	3.1108	3.8139	4.8568	6.7187	8.7672	11.7696	14.2845	20.7197
5000	3.1884	3.8485	4.8415	6.6460	8.5514	11.2946	13.5171	19.5563

Table 3. Percentage points of the distribution of Δ_3

	Confidence Levels							
sample size	0.20	0.15	0.10	0.05	0.025	0.01	0.005	0.001
25	3.1935	4.5256	6.7692	12.1857	19.0634	31.8223	41.9733	78.1272
50	3.1777	4.4377	6.6137	11.1382	17.1549	28.0070	38.2603	71.3414
75	3.2813	4.4962	6.6013	11.0275	16.7108	26.5282	36.3657	64.9030
100	3.3338	4.5160	6.4754	10.6383	15.7954	25.0846	32.1906	57.3123
250	3.4230	4.5159	6.3430	10.1040	14.5024	21.0188	27.3957	44.6579
500	3.3955	4.4892	6.2152	9.4247	13.2461	18.7645	23.7978	39.2705
1000	3.3933	4.4437	6.0026	8.9144	12.1760	17.3116	21.0481	32.7229
5000	3.3931	4.3236	5.7419	8.3262	11.0594	14.9927	18.0884	27.5194

As it could be expected, the critical values of Δ_1 and Δ_2 are very similar. In pure theory, they even should be equal, since the left and right semi–skewness have

[2]The results for all sample sizes are in annex

symmetrical distributions and the left and right kurtosis have the same distributions for a symmetrical distribution as the Gaussian. A further comment is that, since the distribution of Δ_3 for the Gaussian has more extreme values for the percentage points computed here, the test is statistic is further away from the classical χ^2 approximation.

3.2. **Sensitivity Surfaces.** As a first mean to analyse the performance of our tests, we used the sensitivity surface presented by Mudholkar et al. (1991). This procedure consists in studying the variation of the statistic's level when applied to 'profiles', that is 'ideal samples', from the Tukey lambda family of two parameters distributions.

Profiles are samples that are perfectly representative of their parent distribution. A size n profile P_n has the following definition:

Definition 3.1. A set $\{y_1, y_2, \cdots, y_n\}$ is a size n profile P_n of a given distribution F if $\sup_x |F_n P_n(x) - F(x)| \to 0$ when $n \to \infty$, where $F_n P_n$ is the empirical distribution function (EDF) of the $\{y_1, y_2, \cdots, y_n\}$, and $F_n P_n(x)$ is the value of this EDF at x.

Patel and Mudholkar (1983) show that for any continuous distribution F, the size n profile given by:

$$P_n^* = \left\{ F^{-1}\left(\frac{i - 0.5}{n}\right), \quad i = 1, 2, \cdots, n \right\},$$

is optimal among all size n profiles of F, in the sense that $\sup_x |F_n P_n(x) - F(x)|$ and $E_F[F_n P_n(X) - F(X)]^2$ both reach their minimum for $P_n = P_n^*$.

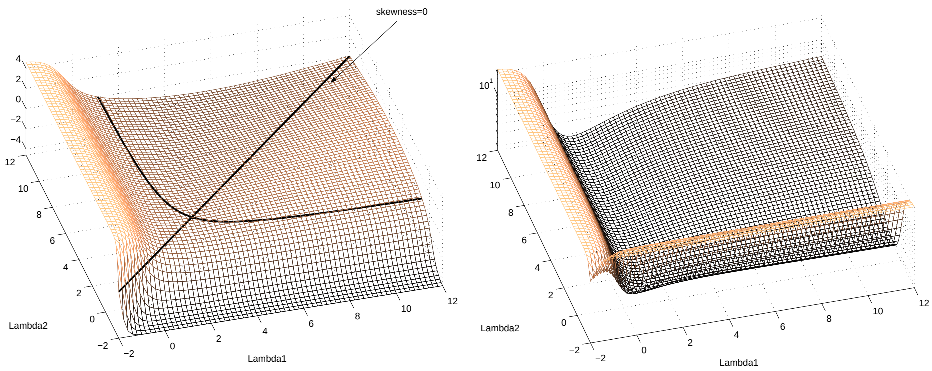
These optimal size n profiles are employed for the distributions of the two shape parameters Tukey lambda family. These distributions are the distribution of random variables X_{λ_1,λ_2} such that:

$$X_{\lambda_1,\lambda_2} = F^{-1}(U) = \frac{U^{\lambda_1} - 1}{\lambda_1} - \frac{(1-U)^{\lambda_2} - 1}{\lambda_2},$$

with $(\lambda_1, \lambda_2) \in \mathbb{R}^2$, using L'Hospital rule when λ_1 and λ_2 are zero. This distribution proves almost undistinguishable from a normal $\mathcal{N}(0, \sigma)$, with $\sigma = 1.46357$ at $\lambda_1 = \lambda_2 = 0.1349$ and exhibits similar moments at $\lambda_1 = \lambda_2 = 5.2$ although with truncated tails. To help interpret the surfaces, the members of this family can be classified as follows:

- **Class I :** $(\lambda_1 < 1, \lambda_2 < 1)$, contains most of the distributions occurring in statistical and economical practice, they are unimodal with continuous tails. When both parameters are below zero, the k^{th} moment exists when $\lambda_1, \lambda_2 \geq -1/k$.
- **Class II :** $(\lambda_1 > 1, \lambda_2 < 1)$, these are distributions similar to the exponential. When $\lambda_1 = \infty$ and $\lambda_2 = 0$, it is indeed the exponential distribution.
- **Class III :** $(2 < \lambda_1 < 1, 2 < \lambda_2 < 1)$, contains U-shaped distributions, with both tails truncated.
- **Class IV :** $(\lambda_1 > 2, 2 < \lambda_2 < 1)$, distributions have a density with one mode and one anti mode, truncated at both tails, the right tail rising sharply.
- **Class V :** $(\lambda_1 > 2, \lambda_2 > 2)$, regroups unimodal distributions with both tails truncated.

To get an idea of the moments properties of these distributions, figure 1 presents their approximate skewness and kurtosis. They are symmetric along the $\lambda_1 = \lambda_2$ axis and approximately along semi lines following $\lambda_1 = 1$ and $\lambda_1 = 2$. They are platykurtic in the neighbourhood of these semi lines and leptokurtic elsewhere.



(A) Skewness (B) Kurtosis

FIGURE 1. Approximate moments of the Tukey Lambda distributions

Sensitivity surfaces were obtained, with λ_1 and λ_2 between -2 and 12, for the three statistics presented in 3.1, 3.2 and 3.3. This surfaces are presented in figures 2, 3 and 4, respectively.

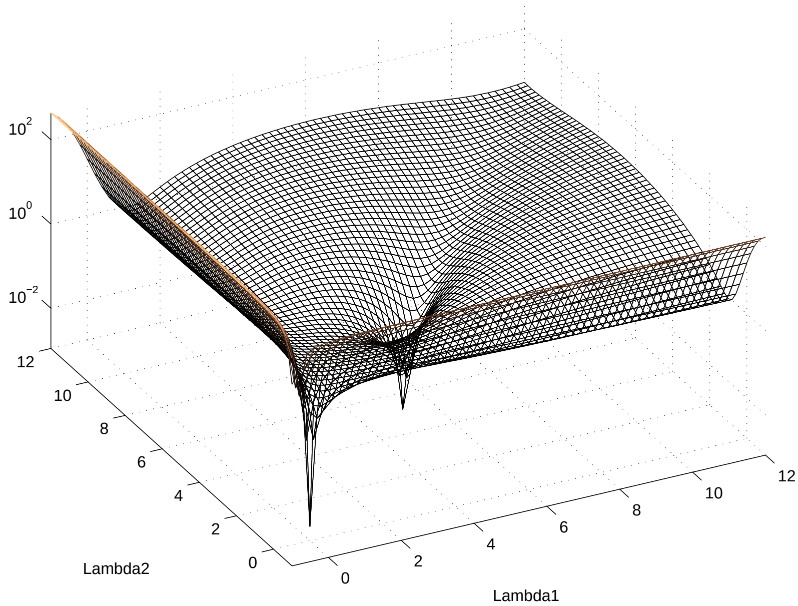


FIGURE 2. Sensitivity Surface of Δ_1 for $n = 25$

For Δ_1 and Δ_2, the sensitivity surface confirm the characteristics anticipated. Both perform very well against asymmetric distributions, (away from the $\lambda_1 = \lambda_2$ axis). They are however slightly asymmetric, Δ_1 being stronger against leptokurtic right skewed distributions, while Δ_2 performs better against leptokurtic left skewed distributions. However, the difference is only limited and situated mostly in areas where the rejection of normality is very strong. The rest of there sensitivity surfaces are roughly identical, implying that there powers will be very similar against most alternatives. On the other hand Δ_3 presents a quite different sensitivity surface. Indeed,

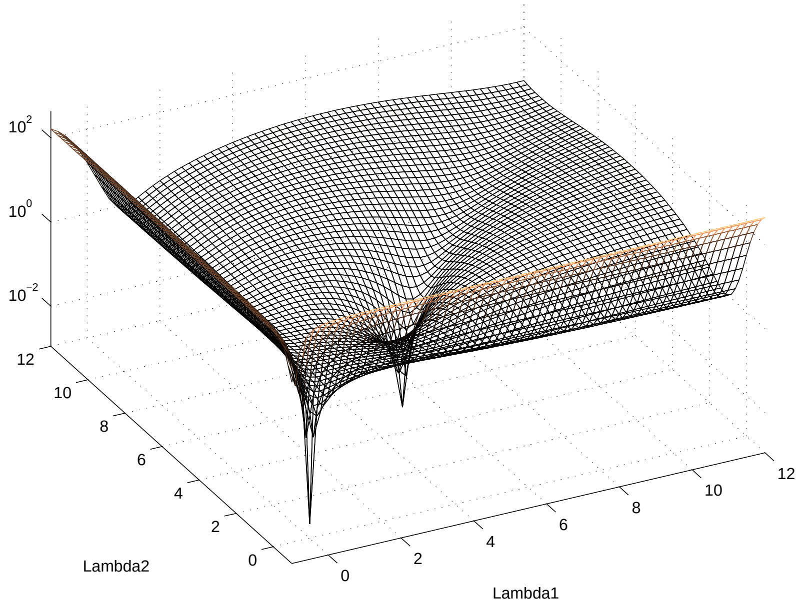


FIGURE 3. Sensitivity Surface of Δ_2 for $n = 25$

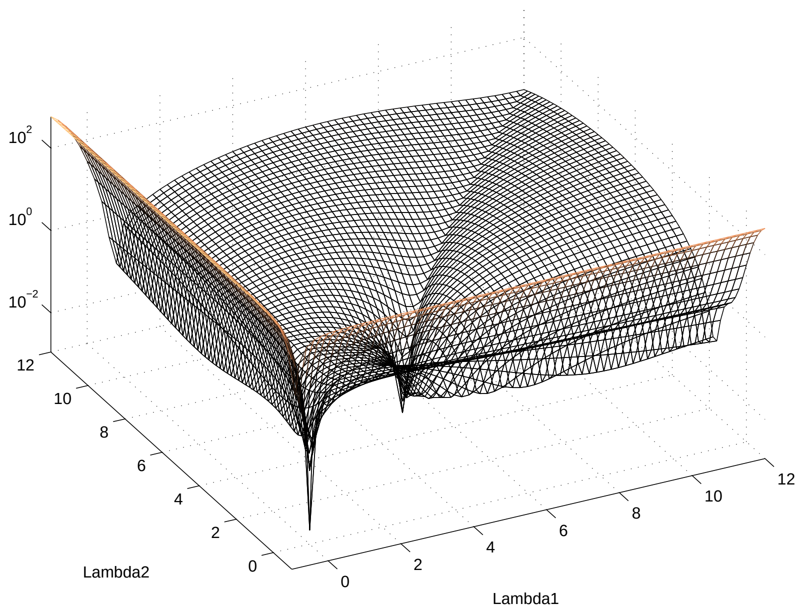


FIGURE 4. Sensitivity Surface of Δ_3 for $n = 25$

as expected, it performs as the best of the two previous tests for strongly asymmetric leptokurtic distributions, yet at other points of the surface it presents deep rifts indicating a limited power. The main such rift originates at the (5.2,5.2) point of quasi normality and follows the lines of $\lambda_2 = 1$, an other rift is along the $\lambda_1 = 1$. Δ_1 and Δ_2 also present such zones of weakness, yet in their case they are far less deep, indicating a much higher sensitivity against such alternatives. Another marking feature of the sensitivity surfaces is that there is, in all three cases a noticeable recess along the axis of symmetrical distributions, marking a lesser power against those alternatives. However, this feature is much stronger in the case of Δ_3 which therefore should have less power than its siblings against symmetric alternatives.

4. Powers Comparison

The power of Δ_1, Δ_2 and Δ_3 against different alternatives and compared to the main existing normality test is estimated through a simulation study. For a review of other major power studies, see Shapiro et al. (1968) and Pearson et al. (1977) or Saniga and Miles (1979), which focuses on moments tests and stables alternatives. Refer to section 2 for some general conclusions that these major power studies have reached.

4.1. **Simulation Set-up.**

The simulation study was carried out with $n = 25$, 50 and 100[3] with 5000 samples drawn from 26 non-normal distributions specified in Table 4. The distributions considered are classified according to the following groups:

Group I : symmetric, leptokurtic,
Group II : symmetric, platykurtic,
Group III : asymmetric, leptokurtic,
Group IV : asymmetric, platykurtic.

The distributions within each group are arranged in order of increasing departure from normality as measured by their standardised coefficient of skewness $\sqrt{b_1}$ and their standardised coefficient of kurtosis, b_2. In group I, the distributions include $SC(\varepsilon, \sigma_\varepsilon)$ which is the scale-contaminated normal with $100 * \varepsilon\%$ of $\mathcal{N}(0, \sigma_\varepsilon^2)$ being the contaminant. Similarly, $LC(\varepsilon, \mu_\varepsilon)$ is the location-contaminated normal with $100 * \varepsilon\%$ of $\mathcal{N}(\mu_\varepsilon, 1)$ being the contaminant in Group III.

Table 4. Properties of the Distributions Used in the Power Study

distribution	var$X$	$\sqrt{\beta_1}$	β_2	distribution	var$X$	$\sqrt{\beta_1}$	β_2
Group I. Symmetric-Leptokurtic				Group III. Asymmetric-Leptokurtic			
$\mathcal{N}(0,1)$	1	0	3	Weibull(2)	0.21	0.63	3.25
t_{10}	1.25	0	4	$LC(0.10, 2)$	1.36	0.36	3.36
Logistic	3.29	0	4.2	$LC(0.20, 6)$	6.76	1.18	3.18
$SC(0.05, 3)$	1.4	0	7.65	χ^2_{10}	20	0.89	4.20
$SC(0.1, 3)$	1.8	0	8.33	$LC(0.10, 6)$	4.24	1.78	5.98
t_4	2	0	∞	$LC(0.05, 6)$	2.71	2.07	8.99
$SC(0.05, 5)$	2.2	0	19.96	χ^2_2	4	2	9
$SC(0.1, 5)$	3.4	0	16.45	χ^2_1	2	2.83	15
t_2	∞	0	∞	Weibull(0.5)	20	6.62	87.72
				Lognormal(0,10)	4.67	6.18	113.94
Group II. Symmetric-Platykurtic							
$U(0,1)$	0.08	0	1.8	Group IV. Asymmetric-Platykurtic			
Beta(1.25,1.25)	0.07	0	1.91	Beta(3,2)	0.04	-0.29	2.36
Beta(1.5,1.5)	0.06	0	2	Beta(2,1)	0.06	-0.57	2.4
Beta(2,2)	0.05	0	2.14	Beta(2.5,1)	0.05	-0.73	2.76

[3] Due to the fact that, with samples of 100, the non-normalities of the samples already become "very" easy to detect for all tests, we limited our study to this size.

The tests used in this power study are the three tests presented in section 3.1, the Bowman-Shenton test, the Anderson-Darling test (in its Stephens (1974) guise), the Shapiro-Francia and the Royston's 1993 version of the Shapiro-Wilks.

The existing normality tests considered in this study include $W(W')$, $W*$ and A^2. Recall that W is the Shapiro and Wilk (1965) test and A^2 is Stephens (1974)'s version to the Anderson-Darling (1954) test. Where the sample size exceeds 50, Shapiro and Francia (1972) W' will be used in place of W since it extends the range of W from 50 and below to 400. $W*$, which is Royston (1993b) approximation to W(W'), will be considered a separate test as it will be informative to compare its power to W(W'). Bowman and Shenton (1975)'s K^2 is also included as it is structurally closest to our tests.

To differentiate between the tests to see if one test is superior to another, the practice in the literature has been to determine which test has the highest power based on the same set of pseudo-random numbers for each distribution. To generalise the results across different distributions, the averaged rank calculated for each test is sometimes used. The fact that a different set of pseudo-random numbers might give rise to a different ordering of the power is usually ignored. To account for this variability, a formal statistical test on the equality of the power of the tests is conducted in this power study.

As all the tests are subjected to the same set of pseudo-random numbers, the powers of the individual tests are correlated. Hence, Cochran's Q test (Cochran, 1950; Berger and Gold, 1973) is used to account for this correlation. In cases where the equal power hypothesis is rejected, McNemar's test with correction for continuity is used for pairwise comparisons to determine whether the test with the highest power is significantly different from the rest. To maintain the overall type I error rate at 0.05 in the presence of multiple testings, the idea from Fisher's Least Significance Difference is used here. This means that multiple comparisons are carried out only if the hypothesis of equal power using Cochran's Q is rejected. In addition, the same type I error rate is used for both Cochran's Q and McNemar's tests. For details of both tests, refer to Siegel and Castellan (1988).

The results from using Cochran's Q and McNemar's tests will be reflected as superscripts to the test with the highest power in this power study. The superscripts will denote the number of tests, including the one with the highest power, that are significantly better than the rest. Hence, a '1' would reflect that the test with the highest power has significantly higher power than the rest while a '7' would mean that all the tests have the same power. In practice, we do not use superscripts in the case where all tests have the same power. The empirical level of each test is also given based on a normal sample of 100 000. 95% confidence intervals on the empirical level of each test will be used to assess if they contain the relevant nominal levels. This information is useful since it acts as a check on possible inflation/deflation of the power estimates.

4.2. **Results of the Power Tests.** The results are presented in two tables for each sample size and significance level. The first table details the results for symmetric distributions, while the second is concerned with asymmetric distributions. Presented in table 5 and 6 are the results for a sample size of 50, with a confidence level of 0.1. These are probably the kind of parameters most used in empirical works. The detailed results for the other sizes and significance levels are available in annex C. The results presented with an exponent are significantly better than the rest (as indicated by the aforementioned test procedure). When no results have an exponent, all tests perform similarly.

TABLE 5. Power comparisons of normality tests based on 5000 samples at $\alpha = 0.10$ and $n = 50$, Symmetric Distributions

distribution	Δ_1	Δ_2	Δ_3	K^2	A^2	$W(W')$	$W*$
Group I. Symmetric-Leptokurtic							
$\mathcal{N}(0,1)$	0.0924	0.0914	0.0718	0.0448	0.1060^2	0.0978	0.1022^2
t_{10}	0.2446	0.2532^1	0.2268	0.1940	0.1962	0.1692	0.1956
Logistic	0.2880^2	0.2882^2	0.2726	0.2370	0.2438	0.2016	0.2228
$SC(0.05,3)$	0.4682^2	0.4652^2	0.4566	0.4314	0.3608	0.3636	0.4050
$SC(0.1,3)$	0.6536^2	0.6594^2	0.6496	0.6354	0.5508	0.5240	0.5824
t_4	0.5464^2	0.5420^2	0.5342	0.5030	0.5142	0.4620	0.4712
$SC(0.05,5)$	0.7120^2	0.7136^2	0.7054	0.6918	0.6434	0.6560	0.6762
$SC(0.1,5)$	0.8952^3	0.8948^3	0.8948^3	0.8886	0.8592	0.8550	0.8708
t_2	0.8728	0.8720	0.8704	0.8718	0.9006^1	0.8506	0.8648
Group II. Symmetric-Platykurtic							
$U(0,1)$	0.0036	0.0072	0	0.0818	0.7296	0.7541^2	0.7544^2
Beta(1.25,1.25)	0.0064	0.0056	0.0006	0.0232	0.5330^1	0.4294	0.4940
Beta(1.5,1.5)	0.0028	0.0026	0	0.0108	0.3846^1	0.3136	0.3182
Beta(2,2)	0.0042	0.0052	0.0008	0.0022	0.2310^1	0.1634	0.1536

TABLE 6. Power comparisons of normality tests based on 5000 samples at $\alpha = 0.10$ and $n = 50$, Asymmetric Distributions

distribution	Δ_1	Δ_2	Δ_3	K^2	A^2	$W(W')$	$W*$
Group III. Asymmetric-Leptokurtic							
Weibull(2)	0.4606^2	0.4202	0.3204	0.2712	0.4474^2	0.4354	0.4206
$LC(0.10,2)$	0.2634^1	0.2400	0.2232	0.1592	0.2290	0.1880	0.1972
$LC(0.20,6)$	0.9990	0.9732	0.7990	0.9988	1.0000^2	1.0000^2	1.0000^2
χ^2_{10}	0.6666^1	0.6276	0.5356	0.4924	0.6128	0.5862	0.5946
$LC(0.10,6)$	0.9956^4	0.9956^4	0.9920	0.9952^4	0.9922	0.9938	0.9948^4
$LC(0.05,6)$	0.9308^1	0.9282	0.9282	0.9244	0.9028	0.9086	0.9222
χ^2_2	0.9978	0.9926	0.9602	0.9848	0.9994^2	0.9882	0.9998^2
χ^2_1	1.0000^6	0.9998^6	0.9972	0.9998^6	1.0000^6	1.0000^6	1.0000^6
Weibull(0.5)	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Lognormal(0,10)	1.0000^2	0.9996^2	0.9938	0.9988	0.9998^2	0.9982	1.0000^2
Group IV.Asymmetric-Platykurtic							
Beta(3,2)	0.0648	0.0814	0.0290	0.0214	0.2864^1	0.2108	0.2032
Beta(2,1)	0.3616	0.4902	0.2312	0.2730	0.8490^2	0.8318	0.8464^2
Beta(2.5,1)	0.6174	0.7190	0.4798	0.4896	0.9094	0.9046	0.9168^1

The results call for some analysis and comments. Very obviously not all test have the same sensitivity to all sorts of departures from normality. Therefore, we will present our comments by class of distributions.

4.2.1. *Asymmetric-Leptokurtic Alternatives.* This class of distributions is quite important in economical practice, since the empirical distributions always exhibit leptokurtosis and present a certain degree of skewness. It is the most frequent distribution for phenomena that present smooth tails (with no truncation on extreme values).

The classification of the tests in terms of power on this class of distributions seems to depend largely on the size of the samples. In smaller samples the Anderson-Darling A^2 has more power, closely followed by the Δ_1 and Δ_2 tests. These two tests seem to gain relative power when the size increases and for larger samples they clearly are the most powerful tests against asymmetric leptokurtic alternatives. They perform particularly well against distributions that are not too far away from the normal, as their powers are more important for the distributions at the top of our list.

Moreover, we notice that there exists a certain level of deflation for the three Δ tests, that present on normals a power always inferior to the nominal level. This deflation is less severe for larger samples and at 10% confidence level, probably explaining the variations in relative power of the tests. This characteristic is shared by the Bowman-Shenton K^2, although at a less important level. This is probably caused by the difficulties arising when simulating small Gaussian samples for the estimation of the critical values.

Another important feature of the power tests against this class of alternatives is that, as expected, the Δ_2 test is consistently less powerful than the Δ_1. The explanation of such a feature of the results is quite obvious: the skewness of the distributions in our panel of asymmetric-leptokurtic are all positive. It is to notice that the power of Δ_3 is generally not even as good as the minimum of Δ_1 and Δ_2. However it still retains a certain level of power and its lack of power compared to its siblings may be caused by a certain tendency to deflation, shown by its small percentage of rejection of the normal.

We notice as well that on this class of alternatives, the Anderson-Darling A^2 seems to be generally better than the Shapiro tests.

4.2.2. *Symmetric-Leptokurtic Alternatives.* This class of distributions is also quite important in many empirical fields. Furthermore, it encompasses many theoretically important distributions like the t distribution. Though not as frequent as the asymmetric-leptokurtic, these distributions are quite common.

The first striking result is that, at 0.1 significance level, the Δ test are almost always more powerful than the classical tests. There is only one test that has a relatively close power against these alternatives, and it is the Bowman-Shenton K^2 test. However, there is one exception to this domination of the moments and semi–moments tests, the Anderson-Darling A^2 test has the best power against a t_2. It is not entirely important, since this distribution is the farther away from the Gaussian in our sample and all the tests are quite powerful against it.

As in the case of the asymmetric-leptokurtic alternatives, the Δ tests are not as powerful at 0.05 as at 0.1. Again, this comes from a tendency to deflation at this confidence level, a tendency that is still present, but not as important, at 0.1.

An other striking point is that, as expected, the Δ_3 test has powers that are quite close to those of Δ_1 and Δ_2. Indeed, it sometimes even is the most powerful test. Moreover, the Δ_1 and Δ_2 seem to have, as expected from construction and from the sensitivity surface analysis, a clearly identical power.

4.2.3. *Asymmetric-Platykurtic Alternatives.* Clearly, moments, or semi–moments based tests perform poorly against platykurtic alternatives. In the case of asymmetric-platykurtic distributions, the feature that permits to these tests to have a limited power is the skewness. The higher the skewness, the closer their power is to the power of the other classes of tests. It is to notice, however, that the Δ tests perform consistently better than the K^2. This is probably caused by their higher sensitivity to asymmetries. In this case again, Δ_3 has considerably less power than the other Δ tests, thus making it by far the least powerful test of our study against asymmetric-platykurtic alternative.

4.2.4. *Symmetric-Platykurtic Alternatives.* In this case of platykurtosis, there is no asymmetry to save the moments based tests. However, K^2 has far more power than the Δ tests in these cases especially for larger samples, and seems to be more suited to detect departure from normality in terms of peakedness.

5. The Normality of Stock Returns at Different Horizons

One of the key stylised facts about stock returns is that they are generally not normally distributed at short horizons and tend to normality as the observation period increases. However, the details of such a tendency towards normality are yet to be studied. We propose to evaluate both the speed of convergence towards normality for asset returns as the length of time over which they are computed increases and the specifics of such a convergence, notably in terms of semi–moments, as our tests permit us to do.

5.1. **The Data.** In order to study the evolution of the distribution of stock returns, we selected 239 stocks listed in the Standard and Poor's 500 index. These stocks were selected for two reasons : they are relatively liquid, in the sense that their prices changed in at least 88% of the quotation days during the selected period, and these stocks have been quoted at least since 1984.

This selection of sample induces a double bias : first a survivorship bias, the distribution of returns might be different for the stock of companies unlisted or which went bankrupt. However, in this case, we needed a sample of the same number of observation for every stock, in order to have the same confidence level for all stocks. The second bias is that thinly traded stock may have a different behaviour in terms of distribution. Again, we could not include such stocks since thin trading hides some of the evolutions of the price.

We thus obtained 4293 daily quotes for each stock, from January, 1st, 1985 to December, 31st, 2001. From these, we constructed daily log returns. The choice of the log returns comes from the fact that these returns allow us to easily construct returns for longer periods. Indeed, such returns, handy as they are, constitute an approximation that only holds in the neighbourhood of zero. More extreme returns are amplified.

From the original daily observations, we constructed the returns on longer periods of time, ranging from 5 days to 90 days, with a 5 days step. The limit of 90 days was selected so as to maintain a sufficient number of observations (the number of returns per stock in the case of the 90 days returns is 47). Moreover, it is common practice to consider returns on more than three months to be Gaussians, therefore this horizon should represent a sufficient limit. Here, a precision seems to be needed. The time horizons used in our study are not calendar time lengths but quoting days. When we mention a 60 days time horizon, we do not speak of roughly two months of real time but of a span close to three calendar months. We believe that the influence of time scale on the distribution of returns is somehow linked to the information revealed and thus our choice of length measure seems more appropriate in some ways than the calendar time. However, a measure in terms of volumes of transactions might yet improve the results.

The 18 additional returns horizons allow us to evaluate quite precisely the evolution of their distribution towards normality. However, it is quite clear that such a methodology has a few drawbacks. Indeed, the selected horizons correspond to a certain number of actual trading days, and therefore the length of calendar time over which the returns are evaluated is not exactly constant. This variability exists from one observation to another, and more so at certain horizons.

5.2. **Evolution towards Normality.** The first possible analysis of the evolution towards normality of the distributions of returns comes from the actual results of a test of normality. Indeed, we can analyse the proportion of stocks for which the test rejects normality. This would lead to a general understanding of the evolution of the distributions, as the indicator would be a mean taken across the different stocks of the sample. To test for normality, in the following sections, we use the Δ_2 test, presented before.

The results presented in table 7 are, indeed, concordant with our expectations. There is an overall decrease in the proportion of stocks for which distributions normality is rejected as the length of the returns' period increases. However, the decrease is not continuous, indicating that for certain stocks, the hypothesis of normally distributed returns is rejected for a given horizon while the test fails to reject it at some shorter horizons.

TABLE 7. Proportion of rejection of normality at 5%, by horizon

5 days	10 days	15 days	20 days	25 days	30 days	35 days	40 days
1.0000	0.9749	0.8661	0.8577	0.7824	0.6820	0.7071	0.6234

45 days	50 days	55 days	60 days	65 days	70 days	75 days	80 days
0.5105	0.5397	0.7113	0.5356	0.5188	0.3724	0.4435	0.3682

85 days	90 days
0.3808	0.2259

Clearly, this first approach tells us that the "convergence" towards normality of the returns' distributions is not the effect of calendar time alone. Indeed, there are probably matters of information being integrated into the price at different speeds.

To obtain a more precise image of the evolution of these distributions, we propose to use an other indicator of normality. Indeed, the results of the test are a simple binary response : either the test rejects normality or it fails to do so. A more defined picture of the situation could be obtained by analysing the level of the test statistic. Such a method would provide us with more details as to the degree of rejection. However, as the number of observations differs for the different horizons, the confidence level of our estimator will not be constant, implying that comparisons regarding normality are more delicate.

However, to smooth the results, we consider the average values of the statistic over the 239 stocks of the sample. The results are presented in table 8. We need to keep in mind while looking at these results that, if the statistics themselves are comparable, their indications about the normality of the data varies.

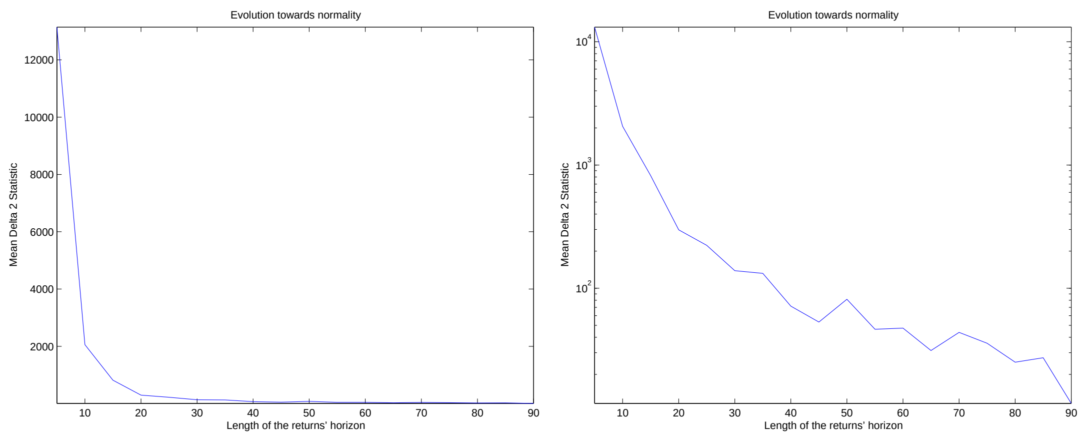
Table 8. Average level of the Δ_2 statistic, by horizon

5 days	10 days	15 days	20 days	25 days	30 days	35 days	40 days
13140	2059.2	820.84	297.93	223.14	138.92	132.12	71.585

45 days	50 days	55 days	60 days	65 days	70 days	75 days	80 days
53.189	81.379	46.51	47.469	31.236	43.869	35.795	25.128

85 days	90 days
27.315	11.646

The speed at which the mean Δ_2 statistic for the returns converges to zero (its level for a large Gaussian sample) is quite important for the first few horizons, then it slows quite a lot. This can be well observed on graph 5.



(A) Linear Scale (B) Log Scale

Figure 5. Convergence of the mean Δ_2 statistics

Our impression in front of such a graph as 5(a) is that the speed convergence is of a logarithmic type. Simply using a log-scale for the graph (as in 5(b)) indicates that the speed is still a bit higher for small values. This leads us to what we call the "log-log" hypothesis. Indeed, the logarithm of the statistic may be linearly related to the logarithm of the number of days in the returns period. To test for this hypothesis, we use a simple Ordinary Least-Squares regression. The results, shown in table 9 and graph 6, seem to confirm the hypothesis.

The "log-log" hypothesis we formulate for the speed of convergence towards normality can be written in the following way:

$$\ln(\Delta_2) = \alpha + \beta \ln(d) + \varepsilon, \quad (5.1)$$

where d is the length (in quotation days) of the period over which returns are computed. The question that remains is: does this equation represents the general speed of convergence to normality, and are the coefficients α and β constant, or do they vary across stocks and markets.

Table 9. Speed of convergence for the log of Δ_2 statistics

Ordinary Least-Squares Estimates			
R-squared	= 0.9742		
Rbar-squared	= 0.9726		
σ_e^2	= 0.0842		
Durbin-Watson	= 1.5188		
Variable	Coefficient	t-statistic	t-probability
Constant	12.546623	38.484845	0.000000
$\ln n$	-2.156534	-24.568006	0.000000

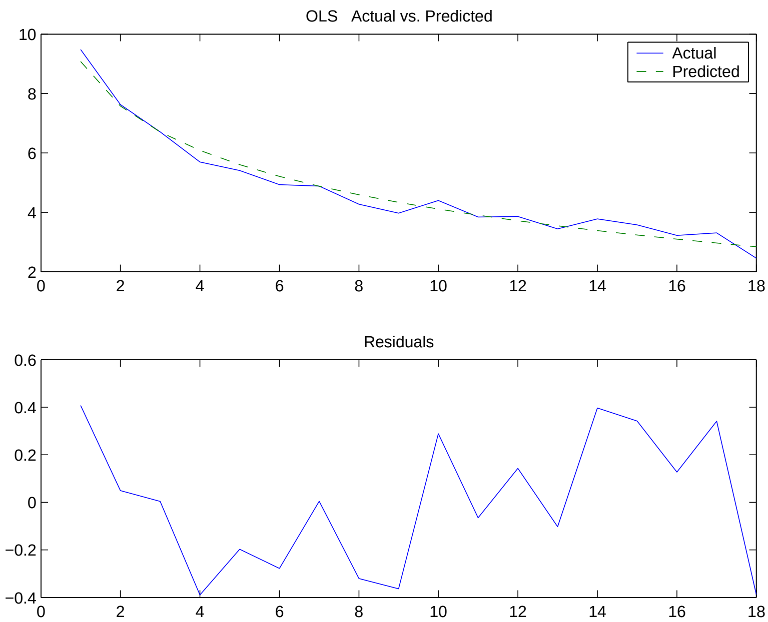


Figure 6. Results of the OLS regression for the convergence of Δ_2

Applying the same estimation to individual stocks is less conclusive. The average level of the R^2 is 63.05%, while the average level of the coefficients is 9.63 for the constant and -1.976 for the coefficient of the log number of days, not varying strongly across stocks and being closer to the results obtained on the general regression when the R^2 is high. However, only a limited number of stocks presents a small R^2 and on the 239 stocks of the sample, only 102 exhibit a R^2 less than or equal to 60%.

On the other hand, only for 21 stocks is the R^2 greater than or equal to 90%. This might be explained by two effects. First, it is possible that for less liquid stocks (they present the lowest R^2) quotation days are not an appropriate measure of the diffusion of information. Second, the limited number of observations probably makes our estimation of the level of normality at certain horizons quite noisy, especially for longer horizons which imply a more limited number of observations.

Confirming the “log-log” hypothesis and the level of the parameters can probably be done by testing out of sample, on a sample with more observations, and perhaps based on actual calendar days or number of trades and not only quotation days. We plan to further our study in this direction in a near future.

5.3. **Details of the Evolution Towards Normality.** As we pointed out in the introduction, the goal of a test of normality is not only to reject or fail to reject the null hypothesis, but also to provide additional informations on the deviations from normality so as to enable the researcher to tackle the distributional properties of the sample.

Indeed, the tests of normality we proposed, and mostly Δ_2, which is most appropriate against the distributions commonly found in economics and finance, are based on semi–moments, which permit a closer inspection of the sources of non-normal behaviour. We will now use them, as well as classical moments to study the details of the evolution towards normality of the returns' distributions.

The estimators of the first moments of the distributions, averaged over the 239 stocks, are presented in table 10. The mean and the variance, not reported here as they are not related to normality, increase with the length of the period, as expected. The mean 5 days mean return is 0.0022368, while for the 90 days return it is 0.038415. The mean 5 days returns variance is 0.045763, at the 90 days horizon, it is 0.17434.

TABLE 10. Mean moments of the stock returns' distribution

horizon	skewness	left skew.	right skew.	kurtosis	left kurt.	right kurt.
5 days	-0.52928	-1.4157	0.88731	9.6447	7.2852	2.3371
10 days	-0.51122	-1.3283	0.8189	7.5243	5.5302	1.959
15 days	-0.44815	-1.2421	0.79625	6.3442	4.4743	1.8255
20 days	-0.49331	-1.25	0.76017	5.902	4.2041	1.6429
25 days	-0.55235	-1.2673	0.71975	5.7234	4.1568	1.4999
30 days	-0.47949	-1.2001	0.72564	5.2261	3.6663	1.487
35 days	-0.59332	-1.2346	0.64858	5.0147	3.7281	1.2047
40 days	-0.5416	-1.2019	0.66789	4.7951	3.4246	1.2814
45 days	-0.49064	-1.1393	0.65641	4.3486	3.059	1.1985
50 days	-0.57244	-1.1939	0.63156	4.6043	3.3679	1.1287
55 days	-0.69934	-1.2888	0.60288	4.9657	3.7763	1.0629
60 days	-0.56739	-1.1899	0.6345	4.5106	3.2293	1.1551
65 days	-0.58779	-1.1864	0.61189	4.3803	3.1633	1.0853
70 days	-0.47656	-1.108	0.64316	4.1887	2.9043	1.1482
75 days	-0.55389	-1.1521	0.61271	4.2168	3.0128	1.0574
80 days	-0.54917	-1.1534	0.61966	4.2563	2.9968	1.1005
85 days	-0.52141	-1.1277	0.62186	4.1368	2.8556	1.1175
90 days	-0.37971	-1.0111	0.64345	3.6165	2.3378	1.1265

The first striking fact about these mean moments is that skewness fluctuates a bit, yet does not clearly change. On the other hand, the components of this third central moment are evolving. The left semi–skewness increases while the right semi–skewness decreases in the same proportion. The Gaussian level for the left and right semi–skewness is $-\sqrt{2/\pi} \approx -0.7979$ and $\sqrt{2/\pi} \approx 0.7979$, respectively. It is striking that the right skewness seems to decrease to the Gaussian level and then keeps on decreasing, opposing the common idea of a general convergence of the distribution of returns to normality.

Even more interesting is the case of kurtosis. The classical kurtosis decreases quite fast initially and then more slowly. However, it remains over 3, the Gaussian level. The semi–kurtosis bring much more information. Indeed, the left semi–kurtosis represents the major part of the total kurtosis, and decreases, again, fast for the first steps, then much slower. This left semi–kurtosis, corresponding to the thickness of the left tail of the distribution (the probability of extremely negative returns) is a key dimension of the risk. Indeed, risk adverse investors should be adverse to kurtosis, and more so to the left semi–kurtosis. Moreover, Tibiletti (2002) shows that semi–moments constitute coherent measures of risk. The Gaussian level of the left semi–kurtosis is 3/2 (as it is for the right semi–kurtosis). We easily see that the left semi–kurtosis remains largely above such a level.

On the other hand, the right semi–kurtosis has a more surprising behaviour. It starts at a rather high level, indicating leptokurtosis, then decreases. However, its decrease is not converging towards the Gaussian level but goes beyond. Indeed, stock returns, thought they are on average leptokurtic, do present a right platykurtosis (the probability of extreme positive events is less than in the Gaussian case).

Analysing both pairs of semi–moments leads us to the conclusion that the distribution towards which the returns' distribution converges is only apparently normal. Indeed, it seems to be an asymmetric distribution, with a thick left tail and a thin right tail. This constitutes quite a bad news, as most finance practitioners, when aware of the problems of normality, think that for periods longer than three months, returns are Gaussian. Of course, some of the individual stocks in the sample seem to converge to gaussianity (and the average distribution is closer to normality at long horizons than at short ones), yet the average moments presented in table 10, give a quite faithful image of the individual moments and semi–moments of the stocks in our sample.

Table 11. Speed of convergence for the log of kurtosis statistics

Ordinary Least-Squares Estimates				
R-squared	=	0.9608		
Rbar-squared	=	0.9583		
σ_e^2	=	0.0024		
Durbin-Watson	=	1.3154		
Variable		Coefficient	t-statistic	t-probability
Constant		2.685151	48.611986	0.000000
$\ln n$		-0.294328	-19.790533	0.000000

The speed of evolution of the estimators of moments, again seems, for most of them, to follow the "log-log" hypothesis. Indeed, for the average moments, the equation 5.1 holds well, with the moments or semi–moments in place of Δ_2. The coefficients, however, are quite different.As could be expected from the first analysis we carried, the model does not hold for the skewness. It has a correct R^2 of 69.78% for the right semi–skewness, while the coefficient of determination is above 92% for the right semi–skewness, kurtosis, left semi–kurtosis and right semi–kurtosis. The results of this regression are given in table 11, while the results of the regressions for semi–moments are summarised in table 12 .

These results provide important information for practitioners, as they allow investment analysts to forecast the evolution of the distribution of the returns on their investments as they modify their holding period. Such information should probably be included in a pricing model or portfolio selection model as the evolution of

Table 12. summary of the regressions to evaluate the log-log model for semi–moments

log of :	left skew.	right skew.	left kur.	right kur.
R-squared	0.6978	0.9253	0.9399	0.9249
Rbar-squared	0.6789	0.9200	0.9362	0.9202
σ_e^2	0.0018	0.0012	0.0036	0.0056
Durbin-Watson	1.4358	1.4271	1.2885	1.4152
constant	0.466732	0.109723	1.322742	2.433220
$\ln n$	-0.077892	-0.136785	-0.288292	-0.317395

the distribution for different returns periods possibly imply the selection of different portfolios depending on the intended investment horizon.

6. Concluding Remarks

The use of semi–moments in estimating the normality of distributions has led us to the construction of goodness-of-fit tests of normality based on these semi–moments. We estimated their critical values at different confidence levels and for different sample sizes through simulations.

These tests come in three different flavours: one of it is more powerful against negatively asymmetric alternatives, the second one is more powerful against positively asymmetric alternatives and the third one is more general. The tests of power conducted showed us that the first two tests perform extremely well against both symmetric and leptokurtic alternatives, being each more powerful against the type of asymmetries it was designed to handle yet retaining important power against the other asymmetries. The third test proved less interesting, not generally yielding improvement to the others and to existing tests of normality.

The Δ_1 and Δ_2 tests were found more powerful than the existing tests at detecting asymmetries especially in presence of leptokurtosis. However, they had difficulties detecting platykurtosis and more so in smaller samples. The results of our power tests show that the best omnibus test is still to be found between the Anderson-Darling A^2 test (modified by Stephens (1974)) and the Royston (1993b) version of the Shapiro-Wilks W test. However, in many fields most of the distributions are both asymmetric and leptokurtic, alternatives against which the semi–moments tests of normality are the most powerful tests. The presence of fat tails and asymmetries is particularly notable in financial returns and thus justifies the use of Δ_1 and Δ_2 in this field of empirical research.

Indeed, in many fields, normality in itself is not the issue, the normality of only one half of the distribution is required. For example, in studying a phenomenon significant only for negative values. More precisely, if as many economists advocates it, investors are only concerned with losses or sub-expected returns, the distributions of returns may not be normal but this is not important if the part of it below its mean closely ressembles that of a Gaussian.

In applying our Δ_2 test to stock returns, we have shown that their distributions do indeed evolve over time towards resembling more a Gaussian. Moreover, this speed of this evolution was determined to be of a "log-log" type, providing practitioners with a mean to forecast the evolution of the distribution when they change their investing or holding horizon. This speed of evolution is also found to describe most of the moments and semi–moments finance is concerned with. However, the evolution of both the right semi–skewness and right semi–kurtosis imply that the distributions of returns do not, in fact, tend to real Gaussians.

Appendix A. Estimation of the Variance of the Semi–Moments

The distribution of the estimators of both the semi–skewness and both the semi–kurtosis, although still unknown have the same variance and this variance is, for Gaussian samples, a polynomial in $1/n$, with n the sample size. We estimated these variances by linear regressions on the results of 100 000 Monte Carlo simulations for 34 sample sizes. The precision of the estimation was increased by simultaneously taking into account the estimators of the right and left semi–moments.

Table 13. Joint semi–skewness variance

Ordinary Least-Squares Estimates			
R-squared	= 1.0000		
Rbar-squared	= 1.0000		
σ_e^2	= 0.0000		
Durbin-Watson	= 2.4562		
Variable	Coefficient	t-statistic	t-probability
$1/n$	1.743795	899.694191	0.000000
$1/n^2$	-10.062152	-170.689160	0.000000

Table 14. Joint semi–kurtosis variance

Ordinary Least-Squares Estimates			
R-squared	= 0.9998		
Rbar-squared	= 0.9998		
σ_e^2	= 0.0000		
Durbin-Watson	= 1.8735		
Variable	Coefficient	t-statistic	t-probability
$1/n$	21.558373	450.424619	0.000000
$1/n^2$	-209.049576	-143.605448	0.000000

Appendix B. Detailed Percentage Points of the Δ Statistics

Table 15. Percentage points of the distribution of Δ_1

sample	Confidence Levels (Δ_1)							
size	0.20	0.15	0.10	0.05	0.025	0.01	0.005	0.001
25	2.5860	3.3347	4.6210	7.4369	11.1187	17.5347	23.7369	40.3180
50	2.6633	3.4104	4.5707	7.1009	10.3100	15.9821	21.4919	39.4102
75	2.7630	3.4913	4.6714	7.1884	10.1869	15.3459	20.3122	36.7530
100	2.8178	3.5546	4.6726	6.9474	9.7999	14.5279	19.1375	34.0987
125	2.8399	3.5614	4.7286	6.9151	9.4674	14.1611	18.4138	33.0372
150	2.9217	3.6100	4.7236	6.9558	9.5413	13.8505	17.7157	29.5374
175	2.9238	3.6400	4.7529	6.9298	9.4082	13.6847	17.7643	32.8185
200	2.9428	3.6473	4.7561	6.8762	9.5204	13.5982	17.8019	29.3939
225	2.9835	3.7177	4.8036	6.9659	9.4552	13.5124	17.2969	28.5735
250	2.9744	3.6933	4.8042	6.9339	9.5498	13.3710	16.3059	26.4350
275	2.9942	3.7030	4.8234	6.9232	9.3140	12.9555	16.3827	26.6910
300	2.9783	3.6766	4.7588	6.7714	9.1033	12.8632	16.1282	26.2977
325	2.9840	3.6784	4.7465	6.7625	8.9891	12.8210	15.7641	25.4076
350	3.0069	3.7015	4.7890	6.8673	9.2160	12.9741	16.6181	26.5715
375	3.0510	3.7689	4.8958	6.9837	9.3199	12.8663	15.8456	24.2661
400	3.0331	3.7263	4.8136	6.7675	9.0163	12.6378	15.8408	23.9371
425	3.0759	3.7923	4.8802	6.9460	9.2730	12.6696	15.6804	24.1399
450	3.0637	3.7815	4.8213	6.8959	8.9889	12.5402	15.2889	23.6060
475	3.0106	3.6913	4.7359	6.6924	8.8440	12.0947	14.7468	21.8206
500	3.0833	3.7890	4.8417	6.8699	9.0618	12.2753	15.3635	25.2615
600	3.0534	3.7552	4.8154	6.7649	8.9464	12.1421	14.7296	22.0635
700	3.1058	3.7849	4.8353	6.8395	9.0128	12.1820	15.0440	22.5222
800	3.0956	3.7587	4.7998	6.8008	8.7911	11.8771	14.2265	22.1190
900	3.0891	3.7633	4.8453	6.8121	8.8763	12.1256	14.6336	21.1512
1000	3.1359	3.8159	4.8489	6.7446	8.8277	11.7667	14.3612	22.0249
2000	3.1494	3.8508	4.8574	6.6965	8.8314	11.7369	14.0996	20.5451
3000	3.1924	3.8745	4.8781	6.6875	8.5610	11.3406	13.4696	18.3021
4000	3.1632	3.8593	4.8329	6.6698	8.6775	11.4369	13.5429	18.5359
5000	3.1884	3.8566	4.8602	6.6731	8.5529	11.2425	13.5729	18.6675
6000	3.1970	3.8874	4.8734	6.7085	8.6482	11.3030	13.2727	18.0659
7000	3.1527	3.8203	4.8190	6.6278	8.5136	11.0248	12.9509	18.5266
8000	3.1574	3.8529	4.8437	6.7532	8.6668	11.3266	13.1849	17.6968
9000	3.1699	3.8579	4.8693	6.7249	8.6977	11.3187	13.2613	17.8795
10000	3.1582	3.8468	4.8780	6.6595	8.5900	11.0104	12.7648	18.0180

TABLE 16. Percentage points of the distribution of Δ_2

sample	Confidence Levels (Δ_2)							
size	0.20	0.15	0.10	0.05	0.025	0.01	0.005	0.001
25	2.6050	3.3547	4.5963	7.4030	11.1573	17.4214	23.1066	43.1937
50	2.6658	3.3952	4.5744	7.0782	10.2906	15.8581	21.7465	41.5505
75	2.7565	3.4757	4.6475	7.1586	10.1378	15.4407	20.8761	39.9846
100	2.8129	3.5454	4.6754	7.0189	9.7665	14.2651	18.9436	31.9495
125	2.8417	3.5682	4.6958	6.8814	9.5681	14.0726	18.1886	32.6232
150	2.9342	3.6260	4.7287	6.8931	9.5212	13.8594	17.7403	29.9134
175	2.9176	3.6133	4.7461	6.9610	9.6285	13.8271	17.9156	32.3932
200	2.9381	3.6442	4.7200	6.8137	9.4696	13.7408	17.4793	29.1355
225	2.9818	3.7138	4.7747	6.9638	9.3607	13.5670	17.0663	26.6660
250	2.9848	3.6849	4.7957	6.9735	9.4567	13.0405	16.5189	27.2613
275	2.9937	3.7053	4.8145	6.9650	9.2539	12.7815	16.1093	24.4746
300	2.9700	3.6811	4.7352	6.8062	9.1769	12.8667	15.8667	24.2779
325	2.9780	3.6768	4.7551	6.7712	8.9643	13.0106	16.1122	26.0845
350	3.0007	3.7211	4.7828	6.8250	9.1978	12.8804	16.4489	26.9597
375	3.0680	3.7879	4.8572	7.0080	9.3145	13.0655	16.1224	25.6955
400	3.0281	3.7463	4.8181	6.8474	9.0028	12.2913	15.5131	23.6245
425	3.0585	3.7963	4.8989	6.9862	9.2138	12.7418	15.7763	24.9512
450	3.0639	3.7649	4.8656	6.8733	9.0246	12.4052	15.2599	25.2040
475	3.0063	3.7156	4.7354	6.6420	8.8145	12.1115	15.0627	22.8146
500	3.0892	3.7634	4.8131	6.8358	9.1231	12.4003	15.3754	23.5589
600	3.0450	3.7666	4.7987	6.8068	8.9468	12.0436	14.7924	22.1619
700	3.0845	3.7916	4.8296	6.8032	8.8905	12.1303	15.3605	23.0547
800	3.0765	3.7634	4.8164	6.7733	8.7573	11.8380	14.2312	21.3776
900	3.0792	3.7768	4.8298	6.7781	8.9511	12.0770	14.8505	22.9247
1000	3.1108	3.8139	4.8568	6.7187	8.7672	11.7696	14.2845	20.7197
2000	3.1410	3.8423	4.8664	6.6945	8.7790	11.6582	13.9541	20.3583
3000	3.1891	3.8671	4.8629	6.6485	8.6945	11.1820	13.4412	18.8296
4000	3.1576	3.8396	4.8464	6.6908	8.7834	11.6207	13.5986	18.8851
5000	3.1884	3.8485	4.8415	6.6460	8.5514	11.2946	13.5171	19.5563
6000	3.1861	3.8856	4.8804	6.6855	8.5992	11.2016	13.2766	18.2487
7000	3.1462	3.8109	4.8300	6.6531	8.4715	11.1036	12.9629	18.2625
8000	3.1467	3.8337	4.8551	6.7214	8.6662	11.2935	13.5073	18.1054
9000	3.1658	3.8561	4.9123	6.7223	8.6311	11.2713	12.9450	17.7985
10000	3.1737	3.8531	4.8322	6.6371	8.5697	11.1888	13.2231	17.6398

TABLE 17. Percentage points of the distribution of Δ_3

	Confidence Levels							
sample size	0.20	0.15	0.10	0.05	0.025	0.01	0.005	0.001
25	3.1935	4.5256	6.7692	12.1857	19.0634	31.8223	41.9733	78.1272
50	3.1777	4.4377	6.6137	11.1382	17.1549	28.0070	38.2603	71.3414
75	3.2813	4.4962	6.6013	11.0275	16.7108	26.5282	36.3657	64.9030
100	3.3338	4.5160	6.4754	10.6383	15.7954	25.0846	32.1906	57.3123
125	3.3020	4.4455	6.3945	10.3929	15.0275	23.8097	31.0722	53.4788
150	3.3727	4.5345	6.4306	10.2971	15.0628	23.0367	29.5584	51.2982
175	3.3682	4.5609	6.3969	10.2049	14.8309	22.6984	30.0527	54.5842
200	3.3449	4.4665	6.2874	10.0603	14.7165	22.6283	29.5158	52.2279
225	3.3919	4.5403	6.3630	10.0336	14.5195	21.9935	28.1860	45.0961
250	3.4230	4.5159	6.3430	10.1040	14.5024	21.0188	27.3957	44.6579
275	3.3912	4.5242	6.3249	9.8256	13.9931	20.8443	26.4470	41.9150
300	3.3731	4.4729	6.2255	9.6553	13.7991	20.3080	25.9205	43.3899
325	3.3792	4.4756	6.2121	9.5541	13.4395	19.9570	25.4475	43.6572
350	3.3948	4.5000	6.1731	9.6111	13.6919	20.3318	26.6943	45.0672
375	3.4574	4.5150	6.2848	9.8715	14.0439	20.1289	25.7969	40.7093
400	3.3936	4.5002	6.1836	9.4670	13.2310	19.0535	24.5229	39.2535
425	3.4172	4.5060	6.2034	9.5124	13.5633	19.5182	25.0405	41.5043
450	3.4048	4.5128	6.2611	9.5032	13.1813	19.0887	24.2048	39.1294
475	3.3435	4.4195	6.0699	9.2995	12.8805	18.4658	23.3001	35.5909
500	3.3955	4.4892	6.2152	9.4247	13.2461	18.7645	23.7978	39.2705
600	3.3584	4.3962	6.0473	9.2364	12.8461	18.3072	22.6187	34.8925
700	3.4328	4.4582	6.0936	9.1910	12.5977	18.1375	22.7198	36.0858
800	3.3779	4.4027	5.9695	9.0035	12.2217	17.1932	21.3698	34.1306
900	3.3848	4.4018	5.9553	9.0280	12.4061	17.5553	21.9064	34.0511
1000	3.3933	4.4437	6.0026	8.9144	12.1760	17.3116	21.0481	32.7229
2000	3.3773	4.3705	5.7865	8.6229	11.6550	16.0253	19.7226	28.2308
3000	3.4561	4.4158	5.8202	8.4157	11.2803	15.0918	18.2819	26.6123
4000	3.3346	4.2730	5.7096	8.3806	11.1277	15.4032	18.7630	27.1330
5000	3.3931	4.3236	5.7419	8.3262	11.0594	14.9927	18.0884	27.5194
6000	3.3959	4.3217	5.7256	8.2645	10.7938	14.5022	17.4324	24.9720
7000	3.3096	4.2534	5.6234	8.2073	10.7264	14.2188	16.8695	23.5062
8000	3.3599	4.2755	5.6482	8.1808	10.8273	14.6689	17.4870	24.1832
9000	3.3612	4.2593	5.6926	8.0974	10.7713	14.4236	17.0406	23.4841
10000	3.3353	4.2301	5.5856	8.0521	10.7443	14.1910	16.9329	23.0113

Appendix C. Detailed Power Comparisons

In this section of the annex, we present the detailed results of the power tests, for samples of size 25, 50 and 100, at significance levels of 0.05 and 0.1. The results already presented in table 5 and table 6 are not repeated. Again, the subscripts indicate the significance of the differences in powers. They are obtained by the Cochran's Q test and the McNemar test, as indicated in the text. The most powerful tests for a given distribution at a given size and significance level are subscripted. The subscripts are omitted when we cannot reject that all tests have the same power.

Table 18. Power comparisons of normality tests based on 5000 samples at $\alpha = 0.05$ and $n = 25$, Symmetric Distributions

distribution	Δ_1	Δ_2	Δ_3	K^2	A^2	$W(W')$	W^*
Group I. Symmetric-Leptokurtic							
$\mathcal{N}(0,1)$	0.0370	0.0382	0.0278	0.0396	0.0492	0.0521	0.0588^1
t_{10}	0.1158^4	0.1154^4	0.0972	0.0796	0.1058^4	0.1022	0.1118^4
Logistic	0.1366^2	0.1380^2	0.1200	0.0956	0.1270	0.1096	0.1232
$SC(0.05,3)$	0.2488^2	0.2500^2	0.2288	0.2110	0.2042	0.1913	0.2154
$SC(0.1,3)$	0.3524^2	0.3552^2	0.3338	0.3124	0.3064	0.2938	0.3056
t_4	0.2930^2	0.2908^2	0.2726	0.2456	0.2678	0.2501	0.2470
$SC(0.05,5)$	0.4406^2	0.4408^2	0.4284	0.4150	0.4092	0.3742	0.4192
$SC(0.1,5)$	0.6232^2	0.6248^2	0.6032	0.6012	0.6072	0.5410	0.6040
t_2	0.5774	0.5774	0.5510	0.5406	0.6148^1	0.5549	0.5586
Group II. Symmetric-Platykurtic							
$U(0,1)$	0.0010	0.0010	0.0004	0.0002	0.2576^1	0.1571	0.1558
Beta(1.25,1.25)	0.0008	0.0008	0.0002	0	0.1694^1	0.0861	0.0852
Beta(1.5,1.5)	0.0016	0.0014	0.0006	0.0004	0.1100^1	0.0527	0.0518
Beta(2,2)	0.0012	0.0012	0.0006	0.0004	0.0708^1	0.0359	0.0368

Table 19. Power comparisons of normality tests based on 5000 samples at $\alpha = 0.05$ and $n = 25$, Asymmetric Distributions

distribution	Δ_1	Δ_2	Δ_3	K^2	A^2	$W(W')$	W^*
Group III. Asymmetric-Leptokurtic							
Weibull(2)	0.1416	0.1060	0.1030	0.0674	0.1680^1	0.1238	0.1234
$LC(0.10,2)$	0.0960	0.0774	0.0764	0.0462	0.1070^1	0.0871	0.0882
$LC(0.20,6)$	0.5744	0.5658	0.4254	0.4180	0.9926^1	0.6494	0.9876
χ^2_{10}	0.2562	0.2142	0.2074	0.1534	0.2738^1	0.2360	0.2294
$LC(0.10,6)$	0.8734	0.8680	0.8218	0.8036	0.9032	0.8921	0.9102^1
$LC(0.05,6)$	0.7218^1	0.7154	0.7066	0.6924	0.6916	0.6800	0.7088
χ^2_2	0.7108	0.6746	0.6020	0.5582	0.8872^1	0.8349	0.8670
χ^2_1	0.9114	0.8882	0.8242	0.8142	0.9942^1	0.9704	0.9930
Weibull(0.5)	0.9842	0.9756	0.9516	0.9540	1.0000^2	0.9878	1.0000^2
Lognormal(0,10)	0.8928	0.8732	0.8298	0.8092	0.9616^1	0.9302	0.9532
Group IV.Asymmetric-Platykurtic							
Beta(3,2)	0.0112	0.0178	0.0058	0.0052	0.0920^1	0.0453	0.0464
Beta(2,1)	0.0508	0.0684	0.0206	0.0200	0.3738^1	0.2630	0.2740
Beta(2.5,1)	0.1318	0.1706	0.0692	0.0680	0.4788^1	0.3951	0.3862

Table 20. Power comparisons of normality tests based on 5000 samples at $\alpha = 0.10$ and $n = 25$, Symmetric Distributions

distribution	Δ_1	Δ_2	Δ_3	K^2	A^2	$W(W')$	W^*
Group I. Symmetric-Leptokurtic							
$\mathcal{N}(0,1)$	0.0800	0.0852	0.0616	0.0806	0.1026^2	0.0998	0.1024^2
t_{10}	0.1826^3	0.1798^3	0.1640	0.1040	0.1752^3	0.1564	0.1680
Logistic	0.2112^3	0.2124^3	0.1906	0.1276	0.2036^3	0.1603	0.1896
$SC(0.05,3)$	0.3124^2	0.3136^2	0.2992	0.2384	0.2828	0.2591	0.2778
$SC(0.1,3)$	0.4300^2	0.4290^2	0.4144	0.3492	0.3886	0.3310	0.3798
t_4	0.3736^2	0.3786^2	0.3578	0.2818	0.3562	0.3058	0.3286
$SC(0.05,5)$	0.4910^2	0.4922^2	0.4776	0.4366	0.4716	0.3988	0.4670
$SC(0.1,5)$	0.6834^2	0.6852^2	0.6752	0.6304	0.6636	0.6174	0.6534
t_2	0.6550	0.6596	0.6392	0.5822	0.6862^1	0.6102	0.6210
Group II. Symmetric-Platykurtic							
$U(0,1)$	0.0068	0.0066	0.0020	0.0008	0.3998^1	0.2784	0.2896
Beta(1.25,1.25)	0.0024	0.0028	0.0008	0.0004	0.2842^1	0.1621	0.1700
Beta(1.5,1.5)	0.0068	0.0070	0.0028	0.0004	0.2122^1	0.1182	0.1202
Beta(2,2)	0.0056	0.0070	0.0026	0.0006	0.1450^1	0.0764	0.0838

Table 21. Power comparisons of normality tests based on 5000 samples at $\alpha = 0.10$ and $n = 25$, Asymmetric Distributions

distribution	Δ_1	Δ_2	Δ_3	K^2	A^2	$W(W')$	W^*
Group III. Asymmetric-Leptokurtic							
Weibull(2)	0.2362	0.2008	0.1746	0.0966	0.2628^1	0.2118	0.2040
$LC(0.10,2)$	0.1728^2	0.1446	0.1388	0.0706	0.1774^2	0.1404	0.1510
$LC(0.20,6)$	0.8318	0.7500	0.5806	0.5724	0.9948^1	0.9282	0.9930
χ^2_{10}	0.3720^2	0.3412	0.3052	0.1992	0.3830^2	0.3266	0.3234
$LC(0.10,6)$	0.9228^3	0.9148	0.8840	0.8672	0.9218^3	0.8944	0.9238^3
$LC(0.05,6)$	0.7452^1	0.7414	0.7368	0.7126	0.7300	0.6892	0.7398
χ^2_2	0.8442	0.8072	0.7210	0.6536	0.9388^1	0.9050	0.9252
χ^2_1	0.9732	0.9554	0.9042	0.8836	0.9970	0.9884	0.9978^2
Weibull(0.5)	0.9972	0.9924	0.9776	0.9782	1.0000^2	0.9986	1.0000^2
Lognormal(0,10)	0.9512	0.9344	0.8904	0.8582	0.9808^1	0.9650	0.9744
Group IV.Asymmetric-Platykurtic							
Beta(3,2)	0.0380	0.0506	0.0210	0.0094	0.1748^1	0.1108	0.1052
Beta(2,1)	0.1522	0.1902	0.0876	0.0418	0.5310^1	0.4361	0.4270
Beta(2.5,1)	0.2796	0.3452	0.1972	0.1174	0.6148^1	0.5374	0.5380

TABLE 22. Power comparisons of normality tests based on 5000 samples at $\alpha = 0.05$ and $n = 50$, Symmetric Distributions

distribution	Δ_1	Δ_2	Δ_3	K^2	A^2	$W(W')$	W^*
Group I. Symmetric-Leptokurtic							
$\mathcal{N}(0,1)$	0.0424	0.0432	0.0336	0.0414	0.0550[2]	0.0413	0.0522[2]
t_{10}	0.1672[2]	0.1688[2]	0.1534	0.1546	0.1214	0.0968	0.1310
Logistic	0.2044[2]	0.1982[2]	0.1866	0.1936	0.1540	0.1272	0.1538
$SC(0.05,3)$	0.4028[3]	0.4036[3]	0.3940	0.4014[3]	0.2860	0.3088	0.3406
$SC(0.1,3)$	0.5846	0.5902[2]	0.5700	0.5964[2]	0.4642	0.4679	0.5216
t_4	0.4472	0.4486	0.4308	0.4564[1]	0.4144	0.3563	0.3998
$SC(0.05,5)$	0.6750[4]	0.6730[4]	0.6722[4]	0.6762[4]	0.5960	0.6211	0.6412
$SC(0.1,5)$	0.8662	0.8664	0.8604	0.8742[1]	0.8230	0.8310	0.8496
t_2	0.8238	0.8228	0.8126	0.8490	0.8610[1]	0.8082	0.8290
Group II. Symmetric-Platykurtic							
$U(0,1)$	0.0011	0.0010	0.0004	0.0002	0.5840[2]	0.5356	0.5840[2]
Beta(1.25,1.25)	0.0008	0.0008	0.0002	0	0.3752[1]	0.3110	0.3190
Beta(1.5,1.5)	0.0016	0.0014	0.0006	0.0004	0.2484[1]	0.1784	0.1810
Beta(2,2)	0.0012	0.0012	0.0007	0.0004	0.1354[1]	0.0752	0.0748

TABLE 23. Power comparisons of normality tests based on 5000 samples at $\alpha = 0.05$ and $n = 50$, Asymmetric Distributions

distribution	Δ_1	Δ_2	Δ_3	K^2	A^2	$W(W')$	W^*
Group III. Asymmetric-Leptokurtic							
Weibull(2)	0.2980	0.2598	0.2154	0.1994	0.3182[1]	0.2706	0.3038
$LC(0.10,2)$	0.1694[1]	0.1412	0.1416	0.1132	0.1454	0.1134	0.1244
$LC(0.20,6)$	0.9592	0.8870	0.6560	0.9694	1.0000[2]	0.9976	1.0000[2]
χ^2_{10}	0.5164[1]	0.4746	0.4212	0.3984	0.4954	0.4382	0.4890
$LC(0.10,6)$	0.9950[4]	0.9938[4]	0.9872	0.9948[4]	0.9890	0.9866	0.9942[4]
$LC(0.05,6)$	0.9246[1]	0.9222	0.9226	0.9216	0.8836	0.8824	0.9148
χ^2_2	0.9796	0.9684	0.9138	0.9508	0.9978	0.9214	0.9992[1]
χ^2_1	0.9998	0.9992	0.9902	0.9994	1.0000	0.9996	1.0000
Weibull(0.5)	1.0000	1.0000	0.9998	1.0000	1.0000	1.0000	1.0000
Lognormal(0,10)	0.9976	0.9960	0.9880	0.9932	0.9996[2]	0.9894	1.0000[2]
Group IV.Asymmetric-Platykurtic							
Beta(3,2)	0.0160	0.0218	0.0042	0.0068	0.1750[1]	0.0987	0.1076
Beta(2,1)	0.1432	0.1936	0.0502	0.0920	0.7338[1]	0.7179	0.7254
Beta(2.5,1)	0.3570	0.4328	0.1882	0.2680	0.8312	0.8316	0.8454[1]

Table 24. Power comparisons of normality tests based on 5000 samples at $\alpha = 0.05$ and $n = 100$, Symmetric Distributions

distribution	Δ_1	Δ_2	Δ_3	K^2	A^2	$W(W')$	W^*
Group I. Symmetric-Leptokurtic							
$\mathcal{N}(0,1)$	0.0456	0.0458	0.0356	0.0492	0.0478	0.0501[1]	0.0454
t_{10}	0.2400	0.2336	0.2250	0.2596[1]	0.1508	0.1822	0.1764
Logistic	0.3090	0.3120	0.3030	0.3552[1]	0.2400	0.2620	0.2550
$SC(0.05,3)$	0.6006	0.5972	0.5958	0.6284[1]	0.4112	0.4746	0.5458
$SC(0.1,3)$	0.7966	0.8010	0.7930	0.8410[1]	0.6764	0.7882	0.7632
t_4	0.6928	0.6828	0.6816	0.7528[1]	0.6480	0.6440	0.6632
$SC(0.05,5)$	0.8806[2]	0.8760	0.8774	0.8834[2]	0.8060	0.8004	0.8604
$SC(0.1,5)$	0.9840[2]	0.9828	0.9828	0.9870[2]	0.9696	0.9556	0.9798
t_2	0.9742	0.9714	0.9672	0.9868[2]	0.9850[2]	0.9374	0.9810
Group II. Symmetric-Platykurtic							
$U(0,1)$	0.0024	0.0022	0	0.6832	0.9432	0.9321	0.9866[1]
Beta(1.25,1.25)	0.0008	0.0006	0	0.3406	0.7794	0.6174	0.8796[1]
Beta(1.5,1.5)	0.0010	0.0002	0	0.1474	0.5786	0.5133	0.6488[1]
Beta(2,2)	0.0002	0.0002	0.0002	0.0278	0.2980[1]	0.2044	0.2788

Table 25. Power comparisons of normality tests based on 5000 samples at $\alpha = 0.05$ and $n = 100$, Asymmetric Distributions

distribution	Δ_1	Δ_2	Δ_3	K^2	A^2	$W(W')$	W^*
Group III. Asymmetric-Leptokurtic							
Weibull(2)	0.6224	0.5716	0.4068	0.4878	0.5988	0.6194	0.6766[1]
$LC(0.10,2)$	0.2922[1]	0.2518	0.2458	0.2184	0.2228	0.1708	0.2024
$LC(0.20,6)$	1.0000[5]	0.9992	0.9108	1.0000[5]	1.0000[5]	1.0000[5]	1.0000[5]
χ^2_{10}	0.8572[1]	0.8300	0.7070	0.7754	0.7982	0.7870	0.8474
$LC(0.10,6)$	1.0000	1.0000	1.0000	1.0000	0.9998	0.9998	1.0000
$LC(0.05,6)$	0.9958[2]	0.9936	0.9952	0.9960[2]	0.9814	0.9872	0.9934
χ^2_2	1.0000[6]	0.9998[6]	0.9976	1.0000[6]	1.0000[6]	1.0000[6]	1.0000[6]
χ^2_1	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Weibull(0.5)	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Lognormal(0,10)	1.0000	1.0000	0.9998	1.0000	1.0000	1.0000	1.0000
Group IV.Asymmetric-Platykurtic							
Beta(3,2)	0.0444	0.0584	0.0078	0.0602	0.3874	0.3035	0.3666[1]
Beta(2,1)	0.5136	0.6724	0.2734	0.7778	0.9810	0.9682	0.9958[1]
Beta(2.5,1)	0.8348	0.9162	0.6532	0.9168	0.9942[2]	0.9856	0.9990[2]

Table 26. Power comparisons of normality tests based on 5000 samples at $\alpha = 0.10$ and $n = 100$, Symmetric Distributions

distribution	Δ_1	Δ_2	Δ_3	K^2	A^2	$W(W')$	W^*
Group I. Symmetric-Leptokurtic							
$\mathcal{N}(0,1)$	0.0998	0.1010	0.0778	0.0814	0.0988	0.0961	0.0960
t_{10}	0.3276^1	0.3184	0.3204	0.3082	0.2298	0.2562	0.2482
Logistic	0.4126^4	0.4086^4	0.4134^4	0.4090^4	0.3396	0.3368	0.3374
$SC(0.05,3)$	0.6606^4	0.6624^4	0.6618^4	0.6596^4	0.4976	0.5360	0.5972
$SC(0.1,3)$	0.8454	0.8478	0.8466	0.8626^1	0.7478	0.7762	0.8060
t_4	0.7634	0.7708	0.7708	0.7910^1	0.7346	0.7240	0.7234
$SC(0.05,5)$	0.8992^4	0.8964^4	0.8974^4	0.8950^4	0.8410	0.8460	0.8788
$SC(0.1,5)$	0.9870^4	0.9866^4	0.9880^4	0.9886^4	0.9760	0.9774	0.9840
t_2	0.9818	0.9828	0.9814	0.9900^2	0.9918^2	0.9752	0.9874
Group II. Symmetric-Platykurtic							
$U(0,1)$	0.0844	0.0818	0.0050	0.9482	0.9794	0.9746	0.9974^1
Beta(1.25,1.25)	0.0268	0.0364	0.0010	0.7912	0.8948	0.9454	0.9510^1
Beta(1.5,1.5)	0.0150	0.0156	0.0006	0.5536	0.7284	0.7992^2	0.8004^1
Beta(2,2)	0.0070	0.0068	0.0004	0.2208	0.4576^3	0.4510^3	0.4504^3

Table 27. Power comparisons of normality tests based on 5000 samples at $\alpha = 0.10$ and $n = 100$, Asymmetric Distributions

distribution	Δ_1	Δ_2	Δ_3	K^2	A^2	$W(W')$	W^*
Group III. Asymmetric-Leptokurtic							
Weibull(2)	0.7888^2	0.7570	0.5564	0.6658	0.7300	0.7650	0.7884^2
$LC(0.10,2)$	0.4114^1	0.3862	0.3664	0.2976	0.3206	0.2940	0.2846
$LC(0.20,6)$	1.0000^6	1.0000^6	0.9624	1.0000^6	1.0000^6	1.0000^6	1.0000^6
χ^2_{10}	0.9354^1	0.9220	0.8166	0.8760	0.8782	0.8572	0.9070
$LC(0.10,6)$	1.0000^6	1.0000^6	1.0000^6	1.0000^6	0.9998	1.0000^6	1.0000^6
$LC(0.05,6)$	0.9960^3	0.9950	0.9960^3	0.9960^3	0.9876	0.9924	0.9942
χ^2_2	1.0000^6	1.0000^6	0.9992	1.0000^6	1.0000^6	1.0000^6	1.0000^6
χ^2_1	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Weibull(0.5)	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Lognormal(0,10)	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Group IV.Asymmetric-Platykurtic							
Beta(3,2)	0.1626	0.2238	0.0812	0.2682	0.5434^1	0.5118	0.5256
Beta(2,1)	0.8188	0.9284	0.7410	0.9560	0.9950	0.9944	0.9986^1
Beta(2.5,1)	0.9662	0.9884	0.9408	0.9884	0.9982	0.9986	0.9998^1

References

Alles, A. and J. Kling (1994), "Regularities in the Variation of Skewness in Asset Returns," *Journal of Financial Reaserch*, 17: pp. 427–438.

Ané, T. and C. Labidi (2001), "Revisitng the Finite Mixture of Gaussian Distributions with Application to Futures Market," *Journal of Future Markets*, 21 (4): pp. 347–376.

Anderson, T. and D. Darling (1954), "A Test of Goodness of Fit," *Journal of the American Statistical Association*, 49 (268): pp. 765–769.

Berger, A. and R. Gold (1973), "Note on Cochran's Q-Test for the Comparison of Correlated Proportions," *Journal of the American Statistical Association*, 68 (344): pp. 989–993.

Blom, G. (1958), *Statistical Estimates and Transformed Beta-Variables*, Wiley, New York.

Bowman, K. and L. Shenton (1975), "Omnibus Test Contours for Departure from Normality Based on $\sqrt{b_1}$ and b_2," *Biometrika*, 62 (2): pp. 243–250.

Brown, B. and T. Hettmansperger (1996), "Normal Scores, Normal Plots and Tests for Normality," *Journal of the American Statistical Association*, 91 (436): pp. 1668–1675.

Campbell, J., A. Lo and C. MacKinlay (1997), *The Econometrics of Financial Markets*, Princeton University Press, Princeton.

Cochran, W. (1950), "The Comparison of Percentages in Matched Samples," *Biometrika*, 37 (3-4): pp. 256–266.

Cramer, H. (1946), *Mathematical Methods of Statistics*, Princeton University Press, Princeton.

D'Agostino, R. (1970), "Transformation to normality of the null distribution of g1," *Biometrika*, 57: pp. 679–681.

D'Agostino, R. (1971), "An Omnibus Test of Normality for Moderate and Large Size Samples," *Biometrika*, 58 (2): pp. 341–348.

D'Agostino, R. and E. Pearson (1973), "Tests for Departure of Normality. Empirical Results for the distributions of b_2 and $\sqrt{b_1}$," *Biometrika*, 60 (3): pp. 613–622.

D'Agostino, R. and M. Stephens (1986), *Goodness-of-Fit Techniques*, Marcel Dekker, New York.

Fama, E. (1963), "Mandelbrot and the stable Paretian hypothesis," *Journal of Business*, 36 (4): p. 420–429.

Fama, E. (1965), "The Behaviour of Stock-Market Prices," *Journal of Buisness*, 38 (1): pp. 34–105.

Fishburn, P. (1977), "Mean-Risk Analysis with Risk Associated with Below-Target Returns," *Amercian Economic Review*, 67 (2): pp. 116–126.

Fisher, R. (1929), "Moments and Product Moments of Sampling Distributions," *Proceedings of the London Mathematical Society*, 30: pp. 199–238.

Fisher, R. (1930), "The Moments of the Distribution for Normal Samples of Measures of Departure from Normality," *Proceedings of the Royal Society, A*, 130: pp. 16–28.

Gastwirth, J. and M. Owens (1977), "On Classical Tests of Normality," *Biometrika*, 64 (1): pp. 135–139.

Harter, H. (1961), "Expected Values of Normal Order Statistics," *Biometrika*, 48: pp. 151–165.

Johnson, N., S. Kotz and N. Balakrishnan (1994), *Continuous Univariate Distributions*, John Wiley & Sons, New York, 2nd edition.

Kahneman, D. and A. Tversky (1979), "Prospect Theory: an Analysis of Decision under Risk," *Econometrica*, 47 (2): pp. 263–292.

Kendall, M. and A. Stuart (1977), *The Advanced Theory of Statistics, Vol. 1: Distribution Theory*, MacMillan, New York, 4th edition.

Mandelbrot, B. (1963), "The Variation of Certain Speculative Prices," *Journal of Business*, 36 (4): pp. 394–419.

Mardia, K. (1980), "Tests of Univariate and Multivariate Normality," in P. Krishnaiah, ed., *Handbook of Statistics*, North-Holland, Amsterdam, pp. 279–320.

Mudholkar, G., G. Kollia, C. Lin and K. Patel (1991), "A Graphical Procedure for Comparing Goodness-of-Fit Tests," *Journal of the Royal Statistical Society. Series B*, 53 (1): pp. 221–232.

Oja, H. (1983), "New Tests for Normality," *Biometrika*, 70 (1): pp. 297–299.

Patel, K. and G. Mudholkar (1983), "Abstract of Profiles of some statistical distributions and some applications," *IMS Bull.*, 12 (266).

Pearson, E. (1963), "Some Problems Arising in Approximating to Probability Distributions, Using Moments," *Biometrika*, 50 (1-2): pp. 95–112.

Pearson, E., R. D'Agostino and K. Bowman (1977), "Tests for departure from normality : Comparison of powers," *Biometrika*, 64: pp. 231–246.

Pearson, E. and H. Hartley (1972), *Biometrika Tables for Statisticians*, volume 1, Cambridge University Press., Cambridge.

Pearson, K. (1900), "On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can reasonably be supposed to have risen from random sampling," *Philosophical Magazine*, 5 (50): pp. 157–175.

Peiró, A. (1999), "Skewness in Financial Returns," *Journal of Banking and Finance*, 23: pp. 847–862.

Peiró, A. (2002), "Skewness in Individual Stocks at Different Investment Horizons," Procedings of the Multimoment CAPM and Related topics Conference, Finance-sur-Seine, Paris.

Rao, C. (1973), *Linear Statistical Inference and its Applications*, John Wiley & Sons, New York, 2nd edition.

Royston, P. (1982), "An extension of Shapiro and Wilk's W test for normality to large samples," *Applied Statistics*, 31 (2): pp. 115–124.

Royston, P. (1993a), "Graphical Detection of Non-Normality by Using Michael's Statistic," *Applied Statistics*, 42 (1): pp. 153–158.

Royston, P. (1993b), "A toolkit for testing for non-normality in complete and censored samples," *The Statistician*, 42 (1): pp. 37–43.

Saniga, E. and J. Miles (1979), "Power of Some Standard Goodness-of-Fit Tests of Normality Against Asymmetric Stable Alternatives," *Journal of the American Statistical Association*, 74 (368): pp. 861–865.

Shapiro, S. and R. Francia (1972), "An approximate analysis of variance test for normality," *Journal of the American Statistical Association*, 67: pp. 215–216.

Shapiro, S. and M. Wilk (1965), "An analysis of variance test for normality (complete samples)," *Biometrika*, 52: pp. 591–611.

Shapiro, S., M. Wilk and H. Chen (1968), "A Comparative Study of Various Tests of Normality," *Journal of the American Statistical Association*, 63: pp. 1343–1372.

Siegel, S. and N. Castellan (1988), *Nonparametric Statistics for the Behavioral Sciences*, McGraw-Hill, New York, 2nd edition.

Stephens, M. (1974), "EDF Statistics for Goodness-of-Fit and Some Comparisons," *Journal of the American Statistical Association*, 69 (347): pp. 730–737.

Tibiletti, L. (2002), "Higher Order Moments and Beyond," Procedings of the Multimoment CAPM and Related topics Conference, Finance-sur-Seine, Paris.

Tversky, A. and D. Kahneman (1974), "Judgment under Uncertainty: Heuristic and Biases," *Science*, 185 (4157): pp. 1124–1131.

Weisberg, S. and C. Bingham (1975), "An approximate analysis of variance test for non-normality suitable for machine calculation," *Technometrics*, 17: pp. 133–134.

CEREG, Université de Paris Dauphine
E-mail address: f_desmoulins@yahoo.com