Adv. & Appl. in Stat. **5**(3) (2005), 353-370

A NEW LARGE SAMPLE TEST OF UNIVARIATE SYMMETRY: A COMPARATIVE SIZE-POWER STUDY

J. ANNAERT[*, †], G. BRYS[†] and M. DE CEUSTER[†]

[*]Faculty of Economics and Business Administration
Ghent University (UGent)
W. Wilsonplein 5D, B-9000 Ghent, Belgium
e-mail: jan.annaert@ugent.be

[†]Faculty of Applied Economics
University of Antwerp (UA)
Prinsstraat 13, B-2000 Antwerp, Belgium
e-mail: guy.brys@ua.ac.be; marc.deceuster@ua.ac.be

Abstract

In this paper we will present a comparative size-power study of several known tests of symmetry. We discuss both the tests with a known and an unknown symmetry point. We propose a new tail symmetry based test, that appears to often outperform the other tests under study, especially at large sample size. We illustrate the use of the tests on stock index returns.

1. Introduction

The question whether data exhibit skewness has been an issue of general interest. In finance, the role of higher order moments has become increasingly important since traditional risk measures fail to capture the true down-side risk (such as variance) under skewed asset distributions. Harvey and Siddique [11] is only one example of an asset pricing model

2000 Mathematics Subject Classification: 62F03, 62G10.

Key words and phrases: skewness, stock index returns, tail test.

Received June 16, 2005

that explicitly incorporates skewness. Still in finance, the option pricing literature [7, 8] and the theory of portfolio selection [15, 5] have suggested new models to take potential skewness into account.

The hypothesis of symmetry is often an assumption in application of parametric or non-parametric procedures. For instance, in adaptive estimation of regression coefficients the conditional density function of the error terms should be symmetrical around zero [2]. Moreover, the Wilcoxon signed-rank test asks that the parent population is symmetric [10]. Nevertheless, despite its importance, the empirical testing of (unconditional) symmetry has not received much attention. Notable exceptions to this, is a series of papers by Peiró [18, 19, 20] and Kim and White [13]. Using a parametric Kolmogorov-Smirnov based symmetry test Peiró concludes that symmetry often cannot be rejected and that hence a multitude of skewness based models is superfluous and of little use. Kim and White argue that S&P index returns are better described as a mixture containing a predominant component that is nearly symmetric with mild kurtosis and a relatively rare component that generates highly extreme behaviour. Crucial to the underlying economic discussion are the statistical properties of the tests used since it is not clear whether the non-rejection of the null of symmetry is due to the true symmetric nature of the data or due to a lack of power in the tests used. Kim and White examine the performance of robust skewness tests. Also in this paper we will study size and power properties of several symmetry tests.

It is possible to classify tests of symmetry in those with symmetry point θ known and those with θ unknown. Most tests belong to the first group (see, e.g., [6, 17, 23] and many others). As Lehmann [16] pointed out, the problem of testing symmetry with θ unknown is more difficult. A possible solution is given by estimating θ first, but this may lead to the loss of interesting asymptotic properties (see, e.g., [12] with regard to the sign test). Being aware of the incompleteness of this empirical study, we will restrict ourselves to some symmetry tests as it is cumbersome to handle all available tests. To compare both types of symmetry tests we will assume $\theta = 0$ in our simulations.

Let $X = \{x_1, x_2, ..., x_n\}$ be a random sample drawn from a continuous distribution function F. Then we will discuss tests of symmetry about the symmetry point θ, i.e., tests of

$$H_0 : F(x - \theta) = 1 - F(-x + \theta) \text{ for all } x \in \mathbb{R}$$

against

$$H_a : F(x - \theta) \neq 1 - F(-x + \theta) \text{ for some } x \in \mathbb{R}$$

in which the need of knowing θ depends on the test.

In Section 2 we give an overview of the different symmetry tests discussed in the paper. We also contribute to the literature by suggesting new tests which we label the tail symmetry test (TS_k) and the medcouple symmetry test (MS_F). Section 3 presents the results of a simulation study in which the mis-specification of several tests is detected. The power of different tests is assessed vis-à-vis a wide range of alternative distributions. Section 4 applies the tests on real data. Finally, Section 5 concludes.

2. Tests of Symmetry

Here we will present the different tests of symmetry compared in this paper. Firstly, we discuss those with symmetry point θ known and secondly those with θ unknown. An obvious drawback of the first is that they cannot always be applied as in empirical applications θ is not always known. Note that the Matlab code of all tests is available at

http://www.agoras.ua.ac.be/.

2.1. Symmetry tests with θ known

2.1.1. Butler-Smirnov test (BS)

The *Butler-Smirnov* test (BS) is designed to decide whether the cumulative distribution function is symmetric about zero. In case the symmetry point θ differs from zero, we will subtract it from all observations. As described in Chatterjee and Sen [4] the asymptotic p-value is given by

$$e^{n \ln((1+d)^{(-0.5-d/2)}(1-d)^{(-0.5+d/2)})}$$

with

$$d = \sup_{x \geq 0} | F_X(x) + F_{-X}(x) - 1 |$$

and with $F_X(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{(x_i \leq x)}$ and $F_{-X}(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{(-x_i \leq x)}$. Hereby $\mathbf{1}_{(\cdot)}$ stands for the indicator function. Intuitively, this test searches for the largest deviation d between the empirical distribution function of the upper half and that of the lower half of the sample. If this value is too large for certain x, then the null hypothesis of symmetry, i.e., $F_X(x) = 1 - F_{-X}(x)$ for all $x \geq 0$, is rejected.

2.1.2. Kolmogorov-Smirnov symmetry test (KS)

The second test we study is the *Kolmogorov-Smirnov* symmetry test (KS) which is based on the goodness-of-fit test of Kolmogorov-Smirnov [22]. Consider $(X - \theta)^-$ and $(X - \theta)^+$, the negative and positive halves of the sample after subtracting the symmetry point θ, with $X^- = \{-x \,|\, x < 0, x \in X\} = \{x_1^-, \ldots, x_{n_1}^-\}$ and $X^+ = \{x \,|\, x > 0, x \in X\} = \{x_1^+, \ldots, x_{n_2}^+\}$. If X is symmetric around θ, then these two random variables have the same distribution. The Kolmogorov-Smirnov goodness-of-fit test based on $(X - \theta)^-$ and $(X - \theta)^+$ gives us the p-value of the test.

2.1.3. Kozubowski test (KO)

The *Kozubowski* symmetry test (KO) [14] was designed under a skew Laplace model, but the same paper also gives the extension towards a general skew model. They propose to use

$$\tau = \frac{(1 + \kappa^2)^2}{1 + \kappa^4}$$

with

$$\kappa = \sqrt[4]{\frac{\sum_{j=1}^{n_1} x_j^-}{\sum_{j=1}^{n_2} x_j^+}}$$

as absolute measure of skewness (i.e., it leads to the same skewness value when applied on the mirror image of the distribution). Note that κ varies between zero and infinity and τ between one and two. In case of a symmetric distribution $\kappa = 1$ and $\tau = 2$. When κ (resp. τ) tends to zero or infinity (resp. one), the distribution becomes skewed. When testing for symmetry we intrinsically assume that $\kappa = 1$ as the observations at both sides of the symmetry point are equal in absolute value, and then the test statistic is given by

$$\frac{2(2-\tau)}{\gamma}$$

which is asymptotically χ_1^2 distributed with

$$\gamma = \frac{\sqrt{\dfrac{\sum_{j=1}^{n_1}(x_j^-)^2 \sum_{j=1}^{n_2}(x_j^+)^2}{\left(\sum_{j=1}^{n_1} x_j^-\right)^2 \left(\sum_{j=1}^{n_2} x_j^+\right)^2}}}{\sqrt{\dfrac{\sum_{j=1}^{n_1}(x_j^-)^2 \sum_{j=1}^{n_2} x_j^+}{\sum_{j=1}^{n_1} x_j^- \sum_{j=1}^{n_2}(x_j^+)^2}} \sqrt{\dfrac{\sum_{j=1}^{n_1} x_j^- \sum_{j=1}^{n_2}(x_j^+)^2}{\sum_{j=1}^{n_1}(x_j^-)^2 \sum_{j=1}^{n_2} x_j^+}}}.$$

2.2. Symmetry tests with θ unknown

2.2.1. Triples test (TR)

An asymptotically distribution-free test of symmetry with θ unknown is given by Randles et al. [21]. To apply the *Triples* test (TR) each triple $\{x_i, x_j, x_k\}(i < j < k)$ is labeled as right, neutral or left in the following manner:

$$\frac{x_i + x_j + x_k}{3} > med(x_i, x_j, x_k) \Rightarrow \text{right triple},$$

$$\frac{x_i + x_j + x_k}{3} = med(x_i, x_j, x_k) \Rightarrow \text{neutral triple},$$

$$\frac{x_i + x_j + x_k}{3} < med(x_i, x_j, x_k) \Rightarrow \text{left triple},$$

in which $med(\cdot)$ stands for the median. Let T be the number of right triples minus the number of left triples, B_i be the same difference between the right and left triples but only counting the triples involving x_i, and analogously B_{ij} but only counting those triples involving x_i and x_j. Then, the test statistic is given by T/σ_T, where

$$\sigma_T^2 = \frac{(n-3)(n-4)}{(n-1)(n-2)}\sum_{i=1}^{n} B_i^2 + \frac{n-3}{n-4}\sum_{i \le j < k \le n} B_{jk}^2 + \frac{n(n-1)(n-2)}{6}$$

$$-\left(1 - \frac{(n-3)(n-4)(n-5)}{n(n-1)(n-2)}\right)T^2.$$

It is shown in [21] that this test statistic is asymptotically standard normal distributed. A serious drawback of the TR test is its computational complexity as it needs to handle $\frac{n(n-1)(n-2)}{6}$ possible triples.

2.2.2. Kolmogorov-Smirnov-based symmetry tests (PM, HS and $\mathbf{TS}_k$)

Based on the Kolmogorov-Smirnov symmetry test of Subsection 2.1.2. some other tests can be constructed in case θ is unknown. Peiró [18] proposes to estimate θ by the sample mean $\overline{X}$, and than applying the Kolmogorov-Smirnov goodness-of-fit test as has been done in Subsection 2.1.2. We label this test as the *Peiró-Mean* symmetry test (PM). Alternatively, but not considered by Peiró, the median (Q_{50}) can also be used as symmetry point estimator. In this case, we label the test as *Half Symmetry* test (HS).

Pushing the idea of the HS test further, we propose to use the *Tail Symmetry* test (TS_k) with θ unknown. Here we only consider the tails of the distribution in order to determine the (a)symmetric character of the distribution. In order to let each tail start at the same numerical value, we subtract $\frac{Q_k + Q_{100-k}}{2}$ from the observations, before applying the two-sample Kolmogorov-Smirnov test in which Q_k represents the k^{th}

percentile of sample X. Hence the TS test applies the Kolmogorov-Smirnov test on Y^- and Y^+ given by

$$Y^- = \left\{\frac{Q_k + Q_{100-k}}{2} - x \mid x < Q_k,\ x \in X\right\},$$

$$Y^+ = \left\{x - \frac{Q_k + Q_{100-k}}{2} \mid x > Q_{100-k},\ x \in X\right\}.$$

Obviously the choice of k is arbitrarily. A trade-off must be made depending on the number of observations available and the power of the Kolmogorov-Smirnov test. We use $k \in \{15, 20, 25\}$. It is clear that the TS_k test encompasses the HS test for $k = 50$.

2.2.3. Medcouple symmetry test $(\mathbf{MS}_F)$

Another approach is by noting that symmetry implies that skewness is equal to zero. Therefore the asymptotic distribution of any skewness measure under the null hypothesis of a symmetric distribution can be used to test for the presence of skewness. Unfortunately, the asymptotic variance of a skewness measure differs across different symmetric distributions. The latter property is remarkably less pronounced for robust skewness measures as these measures are less influenced by outlying values. Therefore, we propose to use the medcouple, which is a robust skewness measure proposed by Brys et al. [3], and which is defined by

$$MC = med_{x_1 \le Q_{50} \le x_2} h(x_1, x_2)$$

with $x_1 \in X$, $x_2 \in X$ and with the kernel function h given by

$$h(x_i, x_j) = \frac{(x_j - Q_{50}) - (Q_{50} - x_i)}{x_j - x_i}.$$

Note that MC can be computed in $O(n \log(n))$ time due to the fast algorithm given in Brys et al. [3]. Depending on the underlying distribution, we have an asymptotic variance σ^2_{MC} of 1.246 (Normal, MS_N), 1.331 (Uniform, MS_U) and 2.029 (Cauchy, MS_C), and thus the test statistic of the *Medcouple Symmetry* test MS_F becomes

$$\frac{MC}{\sigma_{MC}}$$

which is according to [3] asymptotically standard normal distributed.

3. Simulation Results

In this section we present the simulation results of the given tests for a variety of distributions. The results of our simulation studies are summarised by looking at tables of size and power, and through the use of size-size and size-power curves proposed by Wilk and Gnanadesikan [25] and recently reviewed by Davidson and MacKinnon [9].

3.1. Families of distributions considered

We investigate the given tests by generating samples from the Normal Inverse Gaussian (NIG) family [1]. The *NIG*-distribution emerges as the marginal distribution of X in (X, Z), where

$$X \mid Z \sim N(\mu + \beta Z, Z)$$

with

$$Z \sim IG(\delta, \sqrt{\alpha^2 - \beta^2})$$

in which N stands for the normal distribution and IG for the inverse Gaussian distribution. Shortly, a *NIG*-distributed variable is uniquely determined by giving its four parameters $(\alpha, \beta, \mu, \delta)$, and so we will denote it as $NIG(\alpha, \beta, \mu, \delta)$. The domain of the *NIG*-family is the so-called *NIG*-shape triangle $0 \leq |\chi| < \xi < 1$, where $\xi = (1 + \delta\sqrt{\alpha^2 - \beta^2})^{-1/2}$ and $\chi = \beta\xi/\alpha$. For $\chi < 0$ we get negatively skewed distributions, for $\chi = 0$ we get symmetric distributions and for $\chi > 0$ we get positively skewed distributions. Also, tail thickness enlarges with increasing ξ and $\xi = 0$ yield normal tails.

To get an idea about the shape of the studied distributions, Table 1 shows for the studied parameter sets the theoretical mean, variance, skewness and kurtosis, together with the average median based on 1000 samples of size 10000 (between brackets the p-value of the Student t-test is given). Venter and De Jongh [24] provide the formulas to compute the theoretical values. In all blocks of Table 1 the kurtosis of the first two distributions is more or less the same, but skewness differs (i.e., the first distribution is symmetric while the second one is skewed to the right). The third distribution always shows more kurtosis (and skewness) compared to the second one.

Table 1. Theoretical mean, variance, skewness and kurtosis at a variety of *NIG*-distributions and the empirical averaged median based on 1000 samples of size 10000. Between brackets the p-value of the Student t-test is given

	mean	variance	skewness	kurtosis	Q_{50}
NIG(30, 0, 0, 30)	0	1	0	0.003	0.000 (0.888)
NIG(30, 25, – 22.613, 15)	0	2.96	0.159	0.046	– 0.030 (0.000)
NIG(30, 29, – 4.531, 1.2)	0	2.383	0.955	1.542	– 0.058 (0.000)
NIG(1.9, 0, 0, 1.9)	0	1	0	0.831	0.001 (0.023)
NIG(2.37, 0.8, – 0.85, 2.37)	0	1.199	0.440	0.826	– 0.009 (0.000)
NIG(2, 1.2, – 1.125, 1.5)	0	1.465	1.162	3.050	– 0.050 (0.000)
NIG(1, 0, 0, 1)	0	1	0	3	0.000 (0.273)
NIG(1.14, 0.2, – 0.178, 1)	0	0.919	0.497	3.002	– 0.001 (0.000)
NIG(1, 0.5, – 0.577, 1)	0	1.540	1.612	6.928	– 0.040 (0.000)

3.2. Tables of size and power

In our simulation study, we generated $m = 10000$ samples of size n from the *NIG*-family. Table 2 gives the percentage of p-values of the TS_k and the MF_F tests, respectively, in the simulated m samples smaller than the significance level. In case the underlying distribution is truly symmetric, we expect this percentage, the (empirical) *size* of the test, to be equal to the significance level. Then, the *test* is said to be well specified. The *power* of a test is given by the number of appropriate rejections of the null hypothesis. Alternatively, the power equals the percentage of p-values smaller than the chosen significance level.

Statistical tests with power values close to one are preferred. In Table 2 we used a significance level of 5% and a sample size of $n = 2500$.

From the table it becomes clear that all the selected tests show a percentage nearly equal to 5% in case of symmetry. They also tend towards one in case of asymmetry, although it is straightforward to see that the TS_k tests have the highest power values. Indeed, the MS_F tests hardly detect small positive skewness, which is due to the fact that the medcouple suffers the same problem [3]. Nevertheless, with regard to robustness issues, these tests remain interesting. In the remainder of this paper we will focus upon TS_{20} (abbreviated to TS) and MS_U (abbreviated to MS) as they showed in this table good results regarding size and power.

Table 2. Percentage of rejections of the null hypothesis of symmetry at the significance level of 5% (i.e., size or power) at a variety of *NIG*-distributions, simulated by $m = 10000$ samples of size $n = 2500$

	TS_{15}	TS_{20}	TS_{25}	MS_N	MS_U	MS_C
NIG(30, 0, 0, 30)	0.055	0.054	0.059	0.052	0.044	0.012
NIG(30, 25, – 22.613, 15)	0.832	0.856	0.861	0.402	0.376	0.220
NIG(30, 29, – 4.531, 1.2)	1.000	1.000	1.000	1.000	1.000	1.000
NIG(1.9, 0, 0, 1.9)	0.053	0.053	0.054	0.053	0.046	0.013
NIG(2.37, 0.8, – 0.85, 2.37)	0.999	1.000	1.000	0.824	0.807	0.657
NIG(2, 1.2, – 1.125, 1.5)	1.000	1.000	1.000	1.000	1.000	1.000
NIG(1, 0, 0, 1)	0.049	0.049	0.052	0.057	0.049	0.015
NIG(1.14, 0.2, – 0.178, 1)	0.947	0.946	0.945	0.558	0.530	0.350
NIG(1, 0.5, – 0.577, 1)	1.000	1.000	1.000	1.000	1.000	1.000

A similar table can be made of BS, KS, KO, TR, PM and HS. We omitted it here because such a table actually does not give much information as it is based on a single sample size (n) and on a single significance level ($\alpha = 5\%$). Therefore, we will use further on graphical tools to visualize size and power of a statistical test.

3.3. Size-size and size-power curves

A drawback of Table 2 is its restriction to the chosen significance level, and therefore we summarize more comprehensive results in size-size and size-power curves. A size-size curve represents the empirical distribution function $F(p)$ of the p-values obtained by the simulation of a symmetry test at the null distribution. Since significance values are uniformly distributed, we expect this line to be as close to the 45 degree line as possible for a well specified test. A 95% confidence bound (shaded area) is plotted across this bissectrice to take into account of sampling errors:

$$\left[p - 1.96\sqrt{\frac{p(1-p)}{m}};\ p + 1.96\sqrt{\frac{p(1-p)}{m}}\right]$$

in which m stands for the number of simulations performed in the study (here $m = 10000$) and p stands for nominal size.

In the Figures 1-5 the left, middle and right panel is based on a sample size of respectively $n = 100$, $n = 500$ and $n = 2500$. The size-size plots of Figure 1 show that the KS and the KO tests are overall well specified. Also the TR test follows closely the 45 degree line, but due to its computational complexity, it is only simulated with $n = 100$. The size-size curve of the TS test is stepwise, a conclusion which could also be made with HS and BS, and mainly caused by the discreteness of underlying statistical tests. Nevertheless, when n increases (i.e., $n = 2500$) their size-size curves become almost continuous, and we may conclude the TS test to be well specified. Sometimes the MS test deviates slightly from the confidence bound but overall it could be said that it has correct size values. With the other given symmetry tests actual size and true nominal size differ significantly. The actual size of the PM and the BS symmetry test constantly underestimates the true nominal size, while on the contrary the actual size of the HS test always lies above the true nominal size.

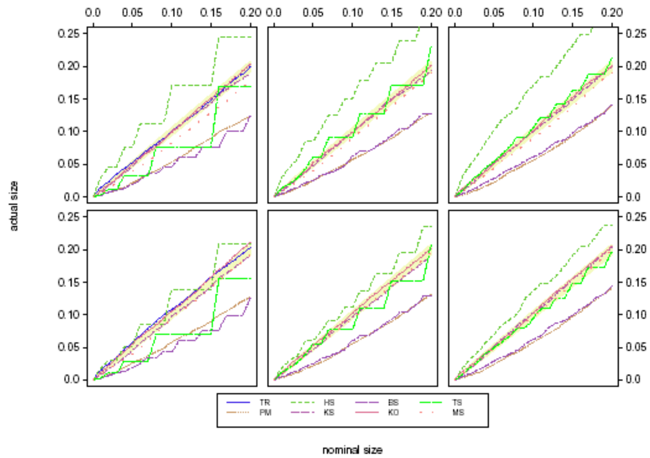

Figure 1. Size-size plots with $NIG(30, 0, 0, 30)$ (upper panel) and $NIG(1, 0, 0, 1)$ (lower panel) as null distribution.

A size-power curve plots the empirical distribution function of the p-values at the null distribution against its counterpart at an alternative distribution. In this way we are able to compare tests with different empirical size values. A powerful test will have a size-power curve converging very rapidly towards one.

The size-power curves of Figures 2-4 clearly show that overall the TS test is superior. Nevertheless, for small samples $(n = 100)$ the TR test clearly outperforms the TS test, but again due to its computational complexity, it is prohibitive to simulate this test if $n = 500$ or $n = 2500$. The well specified symmetry tests (KO, KS and MS) always show worse power values. The other tests are hard to interpret as they are biased at symmetric distributions.

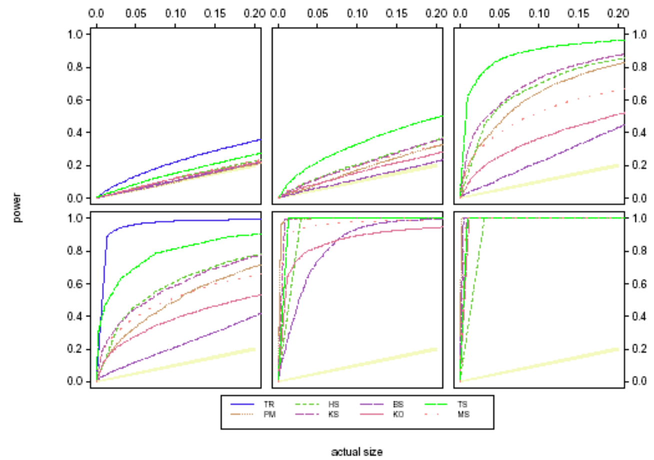

Figure 2. Size-power curves with $NIG(30, 0, 0, 30)$ and $NIG(30, 25, -22.613, 15)$ (upper panel) and $NIG(30, 0, 0, 30)$ and $NIG(30, 29, -4.531, 1.2)$ (lower panel) as respectively null and alternative distributions.

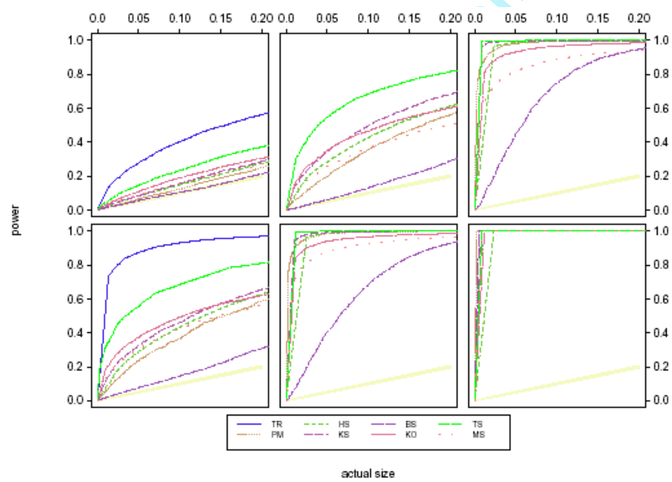

Figure 3. Size-power curves with $NIG(1.9, 0, 0, 1.9)$ and $NIG(2.37, 0.8, -0.85, 2.37)$ (upper panel) and $NIG(1.9, 0, 0, 1.9)$ and $NIG(2, 1.2, -1.125, 1.5)$ (lower panel) as respectively null and alternative distributions.

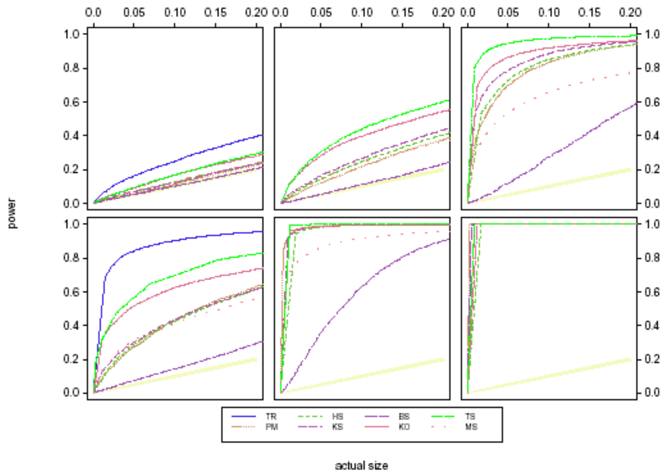

Figure 4. Size-power curves with $NIG(1, 0, 0, 1)$ and $NIG(1, 14, 0.2, -0.178, 1)$ (upper panel) and $NIG(1, 0, 0, 1)$ and $NIG(1, 0.5, -0.577, 1)$ (lower panel) as respectively null and alternative distributions.

To examine the influence of outlying values of the studied tests, we take a look at Figure 5. Here, we generated samples of size $n(1-\varepsilon)$ of $NIG(30, 0, 0, 30)$ or of $NIG(1, 0, 0, 1)$ and added a contaminated sample of size $n\varepsilon$ with $\varepsilon = 1\%$. The latter sample is respectively a normal sample with mean 1.96 or 2.06 and variance 0.1. In this way, we have added a cloud of (wrong) data points in the right tail of the (regular) distribution. A *robust symmetry test* is said to base its results on the majority of the data points, and thus it should conclude that these samples are symmetric. Therefore, Figure 5 plots the size-size curves of these two situations. Clearly, of the well specified tests, MS performs best here as it often remains in the confidence region. Of course, this property is due to the insensitivity of the medcouple at small positive skewness. Moreover, in financial applications it often holds that any observation is important and so the use of robust tests is only sometimes desirable. In that case, it is important to note that of the well specified tests the TS and KO test are well able to detect the asymmetry in these contaminated samples.

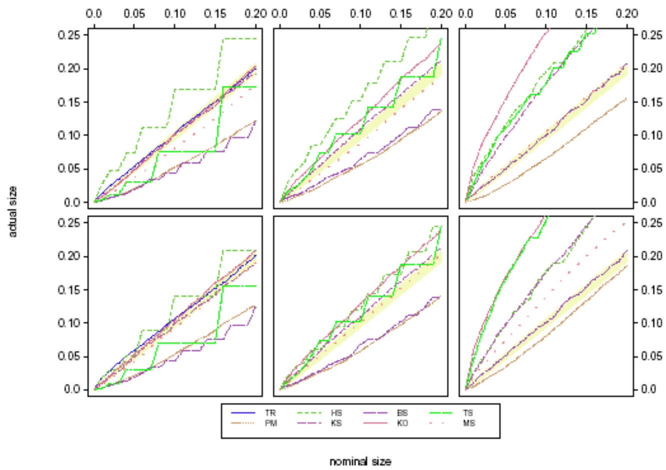

Figure 5. Size-size curves of contaminated situations with $\varepsilon = 1\%$, so with 99% $NIG(30, 0, 0, 30)$ and 1% $N(1.96, 0.1)$ (upper panel) and 99% $NIG(1, 0, 0, 1)$ and 1% $N(2.06, 0.1)$ (lower panel) as null distribution.

4. Applications on Stock Index Returns

From Datastream we downloaded the daily closing prices P_t of the Standard and Poor's (S&P), the FTSE 100 (FTSE) and the Nikkei (NIKK) indices, starting respectively from 20 October 1982, from 2 April 1984 and from 4 January 1984 and all ending on 20 August 2004. Then daily returns were obtained by logarithmic differences as $R_t = \log\left(\frac{P_t}{P_t - 1}\right)$. For the three stock index returns the sample size is above 5000. Table 3 shows the number of observations, the sample mean, variance, skewness and kurtosis, together with the median of the three time series.

Table 3. The mean, variance, skewness, kurtosis and the median (Q_{50}) of the discussed data sets

	mean	variance	skewness	kurtosis	Q_{50}
S&P	0.000	0.011	1.921	44.963	0.000
FTSE	0.000	0.011	0.537	10.818	– 0.001
NIKK	0.000	0.014	0.109	10.572	0.000

In Table 4 we list the significance values of the eight tests considered. Here, the TR test was also performed, although it took some time. We consider this test here as reference point as it was superior at small samples, a conclusion which can probably also be made at large samples due to the asymptotical properties of the test statistic. With a significance level of 5% only the HS test succeeds to copy the results of the TR test. But as this test is biased, we prefer the TS and the MS tests who lead to the same conclusions as the TR test at a significance level of 6%. Also the PM test makes the same rejections of the null hypothesis of symmetry, but already at a significance level of 23%. All other tests detect symmetry in the S&P time series before they detect symmetry in the other data sets.

Table 4. The significance values for the different symmetry tests at the discussed data sets

	TR	PM	HS	KS	BS	KO	TS	MS
S&P	0.214	0.675	0.661	0.278	0.000	0.009	0.167	0.890
FTSE	0.002	0.222	0.032	0.388	0.001	0.069	0.003	0.055
NIKK	0.006	0.140	0.032	0.128	0.134	0.926	0.056	0.050

Also Peiró [20] finds symmetry in stock price indices, although, as we have seen on the size-size and size-power curves, his result is merely based on (good) luck. We made this conclusion statistically more exact by using well specified tests as the TS and TR tests. Finding symmetry in stock (index) returns, Peiró questions the validity of models that explicitly incorporate skewness. The lack of evidence against unconditional skewness, however, does not preclude conditional skewness. As such, the proposed skewness based asset and option pricing models remain useful, if properly applied.

5. Conclusion

In this paper we examined several tests of symmetry, on simulated as well as on real data. It became clear that the existing tests often suffer lack of power. Also the proposed tail symmetry test (TS) is not very powerful, but it nevertheless performs better than the other tests, especially at large sample size. Moreover, the TS test can be regarded as a good alternative to the triples test (TR) which is hard to perform at large data sets $(n \geq 500)$. In small data sets $(n < 500)$ we propose to use the TR test.

References

[1] O. Barndorff-Nielsen, Normal inverse Gaussian distributions and stochastic volatility modeling, Scand. J. Statist. 24 (1997), 1-13.

[2] P. Bickel, On adaptive estimation, Ann. Statist. 10 (1982), 647-671.

[3] G. Brys, M. Hubert and A. Struyf, A robust measure of skewness, J. Comput. Graph. Statist. 13(4) (2004), 996-1017.

[4] S. K. Chatterjee and P. K. Sen, On Kolmogorov-Smirnov-type tests for symmetry, Ann. Inst. Statist. Math. 25 (1973), 287-300.

[5] P. Chunhachinda, K. Dandapani, S. Hamid and A. J. Prakash, Portfolio selection and skewness: evidence from international stock markets, J. Banking and Finance 21 (1997), 143-167.

[6] J. P. Cohen and S. S. Menjoge, One-sample runs test of symmetry, J. Statist. Plann. Inference 18 (1988), 93-100.

[7] C. J. Corrado and T. Su, Skewness and kurtosis in the S&P 500 index returns implied by option prices, J. Financial Research 19 (1996), 175-192.

[8] C. J. Corrado and T. Su, Implied volatility skews and stock return skewness and kurtosis implied by stock option prices, European J. Finance 3 (1997), 73-85.

[9] R. Davidson and J. G. MacKinnon, Graphical methods for investigating the size and power of test statistics, The Manchester School 66 (1998), 1-26.

[10] J. D. Gibbons, Nonparametric Statistical Inference, McGraw-Hill, New York, 1971.

[11] C. R. Harvey and A. Siddique, Conditional skewness in asset pricing tests, J. Finance 55 (2000), 1263-1295.

[12] T. Hettmansperger, Statistical Inference Based on Ranks, John Wiley, New York, 1984.

[13] T.-H. Kim and H. White, On more robust estimation of skewness and kurtosis: simulation and application to the S&P 500 index, Finance Research Lett. 1 (2004), 56-70.

[14] T. J. Kozubowski and A. K. Panorska, Testing symmetry under a skew Laplace model, J. Statist. Plann. Inference 120 (2004), 41-63.

[15] T. Y. Lai, Portfolio selection with skewness: a multiple-objective approach, Review of Quantitative Finance and Accounting 1(3) (1991), 293-305.

[16] E. L. Lehmann, Testing Statistical Hypotheses, John Wiley, New York, 1986.

[17] R. Modarres and J. L. Gastwirth, A modified runs test for symmetry, Statist. Probab. Lett. 31 (1996), 107-112.

[18] A. Peiró, Skewness in financial returns, J. Banking and Finance 23 (1999), 847-862.

[19] A. Peiró, Skewness in individual stocks at different investment horizons, Quant. Finance 2 (2002), 139-146.

[20] A. Peiró, Asymmetries and tails in stock index returns: are their distributions really asymmetric?, Quant. Finance 4 (2004), 37-44.

[21] R. H. Randles, M. A. Fligner, G. E. Policello and D. A. Wolfe, An asymptotically distribution-free test for symmetry versus asymmetry, J. Amer. Statist. Assoc. 75 (1980), 168-172.

[22] N. V. Smirnov, Table for estimating the goodness of fit of empirical distributions, Ann. Math. Statist. 19 (1948), 279-281.

[23] I. H. Tajuddin, Distribution-free test for symmetry based on Wilcoxon two-sample test, J. Appl. Stat. 21(5) (1994), 409-415.

[24] J. H. Venter and P. J. De Jongh, Risk estimation using the normal inverse Gaussian distribution, J. Risk 4(2) (2002), 1-23.

[25] M. B. Wilk and R. Gnanadesikan, Probability plotting methods for the analysis of data, Biometrika 33(1) (1968), 1-17.