A Distribution-Free Test for Symmetry with an Application to S&P Index Returns

Manabu Asai*
Faculty of Economics, Soka University

Ulziijargal Dashzeveg
Graduate School of Social Science, Tokyo Metropolitan University

February 2006

Abstract

We propose a distribution-free test of symmetry. Monte Carlo results show the new test usually outperforms the non-normality robust version of the Jarque-Bera test. Empirical results indicate that the tail of the distribution is too heavy to apply the latter test, while the former is always valid.

JEL Classification: C14; G12

Key Words: Symmetry; Skewness; Stock returns

**Address for correspondence*: Faculty of Economics, Soka University, 1-236 Tangi-cho, Hachioji-shi, Tokyo 192-8577, Japan. Email address: m-asai@soka.ac.jp

1 Introduction

It has long been recognized that daily asset returns are leptokurtic. While the high kurtosis of the returns is a well-established fact, the situation is more unclear with respect to the symmetry of the distribution. The issue of skewness in financial returns is important for financial theories. For instance, as for option pricing theories, the widely used Black-Scholes option pricing model frequently misprices deep-in-the-money and deep-out-the-money options. Corrado and Su (1996, 1997) attribute this fact to the skewness and kurtosis of returns distribution. They show that when skewness and kurtosis-adjustment terms are added to the Black-Scholes formula, improved accuracy is obtained for pricing options.

The test proposed by Jarque and Bera (1981) is widely used to test for normality. The Jarque-Bera (JB) test checks whether the skewness and kurtosis of a sample accord with those of normal distribution, respectively. Since the JB test for skewness is only valid under the assumption of normal distribution, the test is inappropriate for a symmetric but non-normal distribution. The JB test may lead the rejection of true assumption of symmetry simply as a result of non-normality. For this reason, results of JB test for skewness are unreliable for heavy-tailed distributions which are observed for asset returns.

Some authors propose non-normality robust test for skewness. While Gofrey and Orme (1991) extend the JB test by adjusting the variance of sample third moment, Ahmad and Li (1997) and Hyndman and Yao (2002) develop nonparametric tests for symmetry, based on kernel estimation techniques.

The purpose of this article is to propose a distribution-free test for symmetry, based on the two-sample Kolmogorov-Smirnov test. We use Gofrey and Orme (1991) (GO) test as a benchmark. Section 2 explains the new and GO test statistics. Section 3 investigates their finite sample properties by Monte Carlo simulations. Section 4 presents the empirical results of S&P 500 returns.

It should be noted that the method of the current paper is different from distributional approach such as Harris and Kucukozmen (2001). They assumed the exponential generalized beta and the skewed generalized *t* distributions in order to describe asymmetry.

2 A Distribution-Free Test

In this section, we propose a distribution-free test for skewness based on the two-sample Kolmogorov-Smirnov test. In our test, the sample is divided into two groups, and examined whether the two groups are from the same distribution. The proposed test requires no distribution assumption except for the existence of the first moment.

Testing Procedure for the proposed test is as follows;

1. Given a sample, $y_1, \ldots, y_T$, divide it into two groups; One is the group which takes larger value than mean, $\bar{y} = T^{-1}\sum y_t$, while the other is the sample which takes smaller value than mean.
2. Transform y_t into $y_t - \bar{y}$ for the former group and into $\bar{y} - y_t$ for the latter, so that the transformed data takes values larger than zero.
3. Test the null hypothesis that the two datasets come form the same distribution, by using the two-sample Kolmogorov-Smirnov test.

If $y_1, \ldots, y_T$ comes from a symmetric distribution, then the two transformed datasets follow the same distribution. If the sample comes from an asymmetric distribution, then the empirical distribution of two groups may differ significantly. For example, we consider a uni-modal distribution. As for a symmetric distribution, the density function of the transformed datasets is monotonously decreasing function. This is not the case for an asymmetric distribution. Since the mode for the asymmetric distribution is different from the mean, one of the two groups has a mode for the density function. The two-sample Kolmogorov-Smirnov test can detect such a difference powerfully. It should be stressed that the proposed test can be applied to any multi-modal and/or non-modal distribution if its first moment exists.

It is also possible to use such test as the Wilcoxon rank-sum test instead of the Kolmogorov-Smirnov test. While the Kolmogorov-Smirnov test is sensitive to any difference in the distribution of the two samples, the Wilcoxon rank-sum test is especially appropriate to detect differences in location. For this reason we employed the Kolmogorov-Smirnov test.

Computation of the two sample Kolmogorov-Smirnov test is as follows; Let A and B be two (independent) samples. Given an ordinal variable which ranks subjects from 1 to *k*, for each sample separately compute the cumulative percentage of subjects by rank, such that the cumulative percentage of the *k*-th rank will be 100% for each sample. For each rank, the cumulative percentage in sample B is subtracted from the cumulative percentage in sample A. Let *D* equal the largest difference in cumulative percentages for any given ordinal rank. Although it is possible to obtain critical values for Kolmogorov-Smirnov *D*, we employ the large-sample approximation based on $KS^* = 4D^2(n_1n_2)/(n_1+n_2)$ where n_1 and n_2 are sample sizes for A and B, respectively. Under the null hypothesis that the two samples come from the same distribution, KS^* follows $\chi^2(2)$ approximately; See textbooks such as Stuart and Ord (1991).

In the next section we consider a regression model $y_i = x_i\beta + u_i,\quad u_i \sim iid(0,\sigma^2)$, and apply the proposed test for OLS residuals, $\hat{u}_i$, in order to investigate finite sample properties.

As a bench mark, we consider the test of Godfrey and Orme (1991) (GO), which is the non-normality robust test for symmetry. We briefly introduce their test statistic. As shown below, the test requires the existence of moments up to order six. Let $\hat{m}_3 = n^{-1}\sum \hat{u}_i^3$, then $n^{1/2}\hat{m}_3$ is asymptotically normally distributed with zero mean under the null of symmetry, i.e., $E(u_i^3) = 0$. Godfrey and Orme (1991) show that the variance of $n^{1/2}\hat{m}_3$ is estimated by $\hat{v} = n^{-1}\sum \hat{u}_i^6 + 9\hat{\sigma}^6 - 6\hat{\sigma}^2\left(n^{-1}\sum \hat{u}_i^4\right)$, where $\hat{\sigma}^2$ is the residual variance, $n^{-1}\sum \hat{u}_i^2$. Under the null hypothesis, $T = n^{-1}\left(\sum \hat{u}_i^3\right)/\hat{v}$ follows the $\chi^2(1)$ distribution.

It is important to note that the estimator of variance of $n^{1/2}\hat{m}_3$ assumes symmetry but not normality. If the disturbances, u_i, are normally distributed, then $n^{-1}\sum \hat{u}_i^4$

tends to $3\sigma^4$ and $n^{-1}\sum \hat{u}_i^6$ tends to $15\sigma^6$ and $\hat{v}$ reduces to $\hat{v} = 6\hat{\sigma}^6$. This is the case of JB test.

3 Monte Carlo Results

In this section, we investigate the finite sample properties of the new nonparametric test. While we use the GO test as a benchmark, we also report the results of JB test to warn the misuse of the JB test.

We consider a regression model $y_i = x_i\beta + u_i$ ($i = 1,\ldots,n$), where $\beta = (0,1)'$ and $x_i = (x_{1i}, x_{2i})$. We set $x_{1i} = 1$ and $x_{2i} \sim U(0,1)$, and consider six kinds of disturbances, u_i ; (i) $u_i \sim N(0,1)$, (ii) $u_i \sim t(3)$, (iii) $u_i \sim \chi^2(1) - 1$, (iv) $u_i \sim \chi^2(2) - 2$, (v) $u_i \sim \chi^2(50) - 50$, (vi) $u_i \sim Gamma(2,1) - 2$. The first two distributions are symmetric distributions, while the others are asymmetric. The fifth specification can be considered as an extreme case. When the degree of freedom of chi-squared distribution is large, the distribution can be approximated by normal distribution. The third specification is the case of non-modal distribution. We specified as $n = 100$ and conducted 5000 replications. The nominal size of the test is five percent, and thus critical value of the JB and GO tests are 3.84, while that of the new test is 5.99.

Table 1 reports the rejection frequencies of the JB, GO and the new tests. First, we examine the size of the tests. The rejection frequency of the JB test for *t*(3) indicates that the JB test reject the null of symmetry simply because the distribution is non-normal. For the sample size of 100, the GO test tends to under-reject the null when the null hypothesis is true, i.e., the cases of N(0,1) and *t*(3). It should be stressed that the assumption of the GO test is unsatisfied for *t*(3). The rejection frequencies of the new test are close to nominal size. Secondly, we investigate the power of the tests. We neglect the JB test for the reason stated above. The GO and new tests seem to have good power except for the case (v), i.e., $\chi^2(50)$. In the case (v), the rejection frequency of the GO is low, and that of the new test is close to 0.05. This result implies that, in the case that the distribution is well approximated by normal distribution, the new test may lose power, while the GO test has higher power than the proposed one as it uses information about moments. Except for the case (v), the new test has higher powers than those of the GO test. Summing up the above results, the new test has good size and power to detect the asymmetry of distributions.

4 Empirical Results

In this section, we test for symmetry of returns of Standard and Poor's 500 Composite Index (S&P). The sample period for S&P is 1/6/1986 to 12/4/2000, giving 3604 observations. The returns, R_t, are defined as $100\times\{\log P_t - \log P_{t-1}\}$ minus the sample mean, where P_t is the closing price on day t. Table 2 shows the results of JB, GO and the distribution-free tests for raw data, R_t, showing no evidence for asymmetric distribution. As noted before, JB test rejects the null simply because the distribution is non-normal.

We consider the asymmetric stochastic volatility (SV) model with standardized *t*-distribution as follows;

$$R_t = \sigma\xi_t \exp(h_t/2), \quad \xi_t = \varepsilon_t \Big/ \sqrt{\kappa_t/(\nu-2)}, \quad \kappa_t \sim \chi^2(\nu),$$
$$h_{t+1} = \phi h_t + \eta_t, \quad \begin{pmatrix} \varepsilon_t \\ \eta_t \end{pmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho\sigma_\eta \\ \rho\sigma_\eta & \sigma_\eta^2 \end{bmatrix} \right).$$

By the specification, ξ_t follows the standardized *t*-distribution with degree of freedom ν, and the correlation coefficient between ε_t and η_t is ρ. When $\nu \to \infty$, the model reduces to the asymmetric SV model considered by Harvey and Shephard (1996). We estimated the model by the approach proposed by Asai (2005), which is based on the Monte Carlo likelihood method of Durbin and Koopman (1997). The estimate and standard error of ρ are −0.5315 and 0.0683, respectively, while those of ν are 6.4031 and 0.7023. The results show the leptokurtic distribution for returns and negative correlation between innovation terms. We applied the tests for asymmetry to the standardize residual, $\hat{z}_t = R_t\hat{\sigma}_t^{-1}$, where $\hat{\sigma}_t^2$ is the estimates of $\sigma^2\exp(h_t)$. Table 2 indicates that the null hypothesis of symmetry is rejected by the GO test, while it is not rejected by the new test. The contradicted results may be caused by heavy-tailed distribution. Since the estimate of ν is close to six, there is a possibility that the true ν is smaller than six. If it is true, the GO test is invalid as the moment condition is not satisfied.

References

Ahmad, I.A. and Q. Li (1997), "Testing the Symmetry of an Unknown Density Function by the Kernel Method", *Journal of Nonparametric Statistics*, **7**, 279–293.

Asai, M. (2005), "Autoregressive Stochastic Volatility Models with Heavy-Tailed Distribution: A Comparison with Multi-Factor Volatility Models", forthcoming in *Journal of Empirical Finance*.

Corrado, C.J. and T. Su (1996), "Skewness and Kurtosis in S&P 500 Index Returns Implied by Option Prices", *Journal of Financial Research*, **19**, 175-192.

Corrado, C.J. and T. Su (1997), "Implied Volatility Skews and Stock Return Skewness and Kurtosis Implied by Stock Option Prices", *The European Journal of Finance*, **3**, 73-85.

Durbin, J. and S.J. Koopman (1997), "Monte Carlo Maximum Likelihood Estimation for Non-Gaussian State Space Models", *Biometrika*, **84**, 669--684.

Godfrey, L.G. and C.D. Orme (1991), "Testing for Skewness of Regression Disturbances", *Economics Letters*, **37**, 31-34.

Harris, R.D.F. and C.C. Kucukozmen (2001), "The Empirical Distribution of Stock Returns: Evidence from an Emerging European Market", *Applied Economics Letters*, **8**, 367-371.

Harvey, A.C. and N. Shephard (1996), "Estimation of an Asymmetric Stochastic Volatility Model for Asset Returns", *Journal of Business and Economic Statistics*, **14**, 429-434.

Hyndman, R.J. and Q. Yao (2002), "Nonparametric estimation and symmetry tests for conditional density functions", *Journal of Nonparametric Statistics*, **14**, 259-278.

Jarque, C.M., and A.K. Bera (1980), "Efficient Tests for Normality, Heteroskedasticity, and Serial Independence of Regression Residuals", *Economic Letters*, **6**, 255–259.

Stuart, A. and J.K. Ord (1991), *Kendall's Advanced Theory of Statistics*, **2**, 5th ed., Edward Arnold.

Table 1: Rejection Frequencies of Tests for Symmetry

Tests	N(0,1)	*t*(3)	$\chi^2(1)$	$\chi^2(2)$	$\chi^2(50)$	Gamma(2,1)
JB	0.0468	0.6164	1.0000	1.0000	0.3182	1.0000
GO	0.0408	0.0222	0.6592	0.7628	0.2340	0.7642
New Test	0.0467	0.0514	0.9884	0.9074	0.0514	0.9052

Note: Nominal size is five percent.

Table 2: Tests for Asymmetry of Returns of S&P

	JB	GO	New Test
Raw Data	6289.5 [0.0000]	1.2670 [0.2603]	4.6639 [0.0971]
Std. Res.	199.77 [0.0000]	6.8335 [0.0089]	3.9067 [0.1418]

Note: *P*-values are in square brackets. Standardized Residuals are calculated based on the asymmetric SV-*t* model.