Discussiones Mathematicae
Probability and Statistics 32 (2012) 47–58
doi:10.7151/dmps.1143

NORMALITY ASSUMPTION FOR THE LOG-RETURN OF THE STOCK PRICES

Pedro P. Mota[1]

Department of Mathematics
CMA of FCT/UNL

e-mail: pjpm@fct.unl.pt

Abstract

The normality of the log-returns for the price of the stocks is one of the most important assumptions in mathematical finance. Usually is assumed that the price dynamics of the stocks are driven by geometric Brownian motion and, in that case, the log-return of the prices are independent and normally distributed. For instance, for the Black-Scholes model and for the Black-Scholes pricing formula [4] this is one of the main assumptions. In this paper we will investigate if this assumption is verified in the real world, that is, for a large number of company stock prices we will test the normality assumption for the log-return of their prices. We will apply the Kolmogorov-Smirnov [10, 5], the Shapiro-Wilks [17, 16] and the Anderson-Darling [1, 2] tests for normality to a wide number of company prices from companies quoted in the Nasdaq composite index.

Keywords: Anderson-Darling, Black-Scholes, Geometric Brownian motion, Kolmogorov-Smirnov, Log-return, Normality test, Shapiro-Wilks.

2010 Mathematics Subject Classification: 62F03, 62F10, 62G30, 62Q05.

1. Introduction

The assumption that the observed data from a certain phenomenon belongs to a certain distribution is not new and a large number of goodness-of-fit tests can be applied under different conditions. There is an area where the assumption

[1]This work was partially supported by the Fundação para a Ciência e a Tecnologia (Portuguese Foundation for Science and Technology) through PEst-OE/MAT/UI0297/2011 (CMA)

of normality of the data is of extremely importance, the mathematical finance, where in many models the price of the stocks are assumed to follow a geometric Brownian motion process and the log-return of those prices will have normal distribution. This assumption is fundamental, for instance, in the definition of the Black-Scholes model and in the deduction of the Black-Scholes pricing formula for European options. Is easy to understand this importance if we refer that the Black-Scholes pricing formula is widely used in the pricing of derivative products in financial (derivative) markets and the value of those markets are of the order of trillions of dollars. We will test the assumption of normality of the log-returns of the historical stock prices for more than 1000 companies chosen from the Nasdaq composite index, this data is available from the site finance.yahoo.com. From the about 2500 companies that compose the index, we choose the ones with mean daily transaction volume bigger than 50.000 units in the considered year. We will test the normality for prices with different time interval between observations: daily prices, weekly prices and monthly prices. We will work with the data from individual years and from grouped years in a range from January of 2000 to December of 2011. We will test the data using the Kolmogorov-Smirnov, Shapiro-Wilk/Shapiro-Francia and the Anderson-Darling goodness-of-fit tests for normality.

2. Geometric Brownian motion and Black-Scholes model

In mathematical finance are widely used models where the stock prices dynamics are the solution of the stochastic differential equation

$$dX_t = \mu X_t dt + \sigma X_t dB_t, \tag{1}$$

where B_t is the Brownian motion process. For an overview in Brownian motion and stochastic differential equations, see [8] and [12]. The solution of the preview equation is the so called geometric Brownian motion

$$X_t = X_s e^{\left(\mu - \frac{1}{2}\sigma^2\right)(t-s) + \sigma(B_t - B_s)}, \ s < t, \tag{2}$$

and from this representation we can write

$$Ln\left(\frac{X_t}{X_s}\right) = \left(\mu - \frac{1}{2}\sigma^2\right)(t-s) + \sigma(B_t - B_s) \sim N\left(\left(\mu - \frac{1}{2}\sigma^2\right)(t-s); \sigma^2(t-s)\right) \tag{3}$$

because $B_t - B_s \sim N(0, t-s)$. One of the reasons for using this process, when modeling the stock prices, is because with this process one can introduce the Black-Scholes model, where a financial market, see [4] or [12], is composed by two assets, a risk free asset with price dynamics

$$dR_t = rR_t dt, \tag{4}$$

where r is the risk free rate of return, and a risky asset with price dynamics

$$dX_t = \mu X_t dt + \sigma X_t dB_t, \tag{5}$$

where μ and $\sigma > 0$ are real numbers. In this context a contingent claim with exercise date T is any random variable $\chi = \Phi(X_T)$, measurable with respect to the σ-field $\mathcal{F}_T = \sigma(\{X_t : 0 \leq t \leq T\})$.

An example of a contingent claim is the European call option with exercise price K and exercise date T, on the underlying asset with price process X. This claim is a contract that gives to the older of the option the right, but not the obligation, to buy one share of the asset X at price K from the underwriter of the option, at time T (only at the precise time T).

For a contingent claim $\Phi(X_T)$, the arbitrage free price, at time t, is given by the formula

$$F(t,x) = e^{-r(T-t)} \mathbb{E}^{\mathbb{Q}}_{t,x}[\Phi(X_T)] \tag{6}$$

with X verifying the equation

$$dX_u = rX_u du + \sigma X_u dB_u, \quad X_t = x, \tag{7}$$

under the (martingale) measure $\mathbb{Q}$. For a European call option with strike price K and maturity T this formula gives the well known Black-Scholes formula for the option price $\Pi(t) = F(t, X_t)$, where

$$F(t,x) = xN[d_1(t,x)] - e^{-r(T-t)} KN[d_2(t,x)]. \tag{8}$$

With $N[.]$ the $N(0,1)$ distribution function and with

$$d_1(t,x) = \frac{\ln(x/K) + \left(r + \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}}, \tag{9}$$

$$d_2(t,x) = d_1(t,x) - \sigma\sqrt{T-t}. \tag{10}$$

Being this formula widely used in the pricing of the European options.

3. Goodness-of-fit tests

We will apply several goodness-of-fit tests, both the Kolmogorov-Smirnov and the Anderson-Darling test use the cumulative distribution function and the empirical distribution function and are based in a measure of the discrepancy between those two functions and therefore are considered in the class of “distance testes”. Some advantages of this kind of tests is that they are easy to compute, they are more powerful than the χ^2 (Chi-Square) test, over a wide range of alternatives, and they provide consistent tests. The Shapiro-Wilks (or Shapiro-Francia) is a test based in the regression between the order statistics of the sample and the mean value of the order statistics from the tested distribution. This test for normality has higher power than the previous ones.

3.1. Kolmogorov-Smirnov test

The Kolmogorov-Smirnov statistic provides a mean of testing if a set of observations are from some completely specified continuous distribution, $F_0(X)$. The test statistic is given by

$$D = \max\left\{|F_0(X_{k:n}) - S_n(X_{k:n})|, |F_0(X_{k:n}) - S_n(X_{k-1:n})|\right\}, \tag{11}$$

where F_0 is the distribution function of the distribution being tested, S_n is the empirical distribution function and $X_{1:n}, X_{2:n}, \ldots, X_{n:n}$ are the order statistics. Tables with the critical values for this statistic can be found in many texts, for instance in [3] or [11]. However, when certain parameters of the distribution need to be estimated from the sample, like the mean or the standard deviation, then the commonly tabulated critical points are of no use. It seems that if the test is used with this critical values the results will be conservative, that is, the probability of rejecting the null hypothesis, when the hypothesis is true, will be smaller than it should be. One way to overcome this problem is to use the critical values presented by Lilliefors in [10] and for the ones that don't like to use extensive tables is possible to use a modification of the test statistic, see [18], to dispense the usual tables of percentage points. In our case, only when considering one year of monthly data we will work with samples of small dimension, for bigger dimension samples the table values are obtained from simple formulas depending on the sample dimension n and we do not need to use extensive tables. In our problem, when testing the normality of the log-returns of the prices we need to estimate the mean and the standard deviation of the data and because of that we will use the Lilliefors's critical values and if the observed value of the D statistic exceeds the critical value in the table, we will reject the hypothesis that the observations are from a normal population.

Table 1. Lilliefors's critical values for samples of dimension $n = 12$ and $n > 30$

	Percentage points				
n	$\alpha = .20$	$\alpha = .15$	$\alpha = .10$	$\alpha = .05$	$\alpha = .01$
$= 12$	.199	.212	.223	.242	.275
> 30	$.736/\sqrt{n}$	$.768/\sqrt{n}$	$.805/\sqrt{n}$	$.886/\sqrt{n}$	$1.031/\sqrt{n}$

3.2. Shapiro-Wilk test

In the Shapiro-Wilk's test presented in the paper [17] the test statistic W was constructed through the regression of the order sample statistics on the expected normal order statistics. The test statistic W is defined by

$$W = \frac{\left(\sum_{i=1}^{n} a_i X_{i:n}\right)^2}{\sum_{i=1}^{n}(X_{i:n} - \bar{X})^2}, \tag{12}$$

where

$$\mathbf{a}^T = (a_1, a_2, \ldots, a_n) = \frac{\mathbf{m}^T \mathbf{V}^{-1}}{\left(\mathbf{m}^T \mathbf{V}^{-1} \mathbf{V}^{-1} \mathbf{m}\right)^{\frac{1}{2}}}, \tag{13}$$

when

$$\mathbf{m}^T = (m_1, m_2, \ldots, m_n), \quad \mathbf{V} = [v_{ij}]_{n \times n} \tag{14}$$

represents the vector of expected values of standard normal order statistics and the corresponding covariance matrix, respectively. Tables for the expected values of the order statistics $m_i, i = 1, \ldots, n$ for sample sizes $n = 2(1)100(25)300(50)400$ can be found at [7]. The values of $\mathbf{a}$ and the percentage points of W are known up to sample sizes of $n = 50$ and can be found in the original paper [17]. For samples of larger dimension an extension of the Shapiro and Wilk's test can be found in [13] or in alternative, the Shapiro-Francia statistic W' (with simpler coefficients) introduced in [16] and defined by

$$W' = \frac{\left(\sum_{i=1}^{n} b_i X_{i:n}\right)^2}{\sum_{i=1}^{n}(X_{i:n} - \bar{X})^2}, \tag{15}$$

where

$$\mathbf{b}^T = (b_1, b_2, \ldots, b_n) = \frac{\mathbf{m}^T}{(\mathbf{m}^T \mathbf{m})^{\frac{1}{2}}}, \tag{16}$$

can be used. This test is proved to be consistent, see [15]. Values for $\mathbf{b}$ can be computed from the order statistics in [7] and percentage points for W' can be found in [16] for sample sizes $n = 35, 50, 51(2)99$. For samples of larger dimension, percentage points can be found in the paper [14] where an extension algorithm for W' is presented. Just for illustrative purposes we present a few percentage points from the W statistics.

Table 2. Shapiro-Wilk's critical values for samples of dimension $n = 12$ and $n = 30$

	Percentage points			
n	$\alpha = .01$	$\alpha = .02$	$\alpha = .05$	$\alpha = .10$
12	.805	.828	.859	.883
30	.900	.912	.927	.939

For this test, small values of W are significant, i.e. indicate non-normality.

3.3. Anderson-Darling test

The Anderson-Darling test is the third and the last goodness-of-fitness test that we will apply to test for normality. This test, presented in the papers [1] and [2], compares the observed cumulative distribution function to the expected cumulative distribution function as the Kolmogorov-Smirnov test. This test gives more weight to the tails of the distribution than the Kolmogorov-Smirnov test, that is, is more sensitive to deviations in the tails between the empirical and the theoretical distributions.

The test statistic A^2 is defined by

$$A^2 = -n - \frac{1}{n}\sum_{i=1}^{n}(2i-1)\left[\ln(F_0(X_{i:n})) + \ln(1 - F_0(X_{n-i+1:n}))\right], \tag{17}$$

where, as before F_0, represents the distribution function of the distribution to be tested. In the paper [2], asymptotic critical points for significance levels of $1\%, 5\%$ and 10% are presented, from Monte Carlo simulation, extensive tables of critical points can be found in [9] or from saddle point approximation to the distribution function, in [6]. For the situation where the distribution to be tested is normal or exponential and the parameters of the distribution must be estimated, we can find critical points in the papers [19, 20] or [21].

Table 3. Anderson-Darling's asymptotic critical values

Percentage points		
$\alpha = .01$	$\alpha = .05$	$\alpha = .10$
3.857	2.492	1.933

For this test, are the large values of A^2 that are significant, that is, indicate non-normality.

4. Data testing

In this section we will present the results of the normality tests. As we referred, we tested the assumption of normality of the log-returns of the historical stock prices for a large number of companies chosen from the Nasdaq composite index. This data is available from the site finance.yahoo.com and from, something like 2500 companies that compose the index, we choose the ones with daily transaction volume bigger (in mean) than 50.000 units in the considered data interval. We will test the normality for prices with different time interval between observations: daily prices, weekly prices and monthly prices. We will work with the data from individual years and from grouped years in a range from the beginning of 2000 to the end of 2011. We will test the data using the three goodness-of-fit tests for normality presented in the previous section.

First, we considered the daily prices from each one of the years of 2005 through 2011, as we referred, in each year we only test the data from the companies with daily transaction volume bigger than 50.000 unit in that year, this number of selected companies will be represented by N. After getting a large majority of rejected samples we decide to repeat the normality tests for weekly data. The results obtained will be presented in the following tables.

Using daily and weekly prices, the Kolmogorov-Smirnov test and a level of significance of 1% we get the results presented in Table 4.

Table 4. Total (and percentage) of samples rejected by the K.S. test for normality at 1% level of significance (daily and weekly prices)

		Number of rejected companies		% of rejected companies	
Year	N	Daily Prices	Weekly Prices	Daily Prices	Weekly Prices
2005	1029	730	149	70.94%	14.48%
2006	1162	835	150	71.86%	12.91%
2007	1322	1038	216	78.52%	16.34%
2008	1366	1163	288	85.14%	21.08%
2009	1385	1067	201	77.04%	14.51%
2010	1471	927	171	63.02%	11.62%
2011	1568	1128	211	71.94%	13.46%
2007-2011	1322	1321	1037	99.92%	78.44%

Using daily and weekly prices, the Shapiro-Wilk test and a level of significance of 1% we get the results presented in Table 5 and for the Anderson-Darling test the results are presented in Table 6.

Table 5. Total (and percentage) of samples rejected by the S.W. test for normality at 1% level of significance (daily and weekly prices)

		Number of rejected companies		% of rejected companies	
Year	N	Daily Prices	Weekly Prices	Daily Prices	Weekly Prices
2005	1029	917	310	89.12%	30.13%
2006	1162	1035	301	89.01%	25.90%
2007	1322	1240	422	93.80%	31.92%
2008	1366	1324	499	96.93%	36.53%
2009	1385	1278	387	92.27%	27.94%
2010	1471	1205	332	81.92%	22.57%
2011	1568	1458	389	92.98%	24.81%
2007-2011	1322	1322	1241	100%	93.87%

Table 6. Total (and percentage) of samples rejected by the A.D. test for normality at 1% level of significance (daily and weekly prices)

		Number of rejected companies		% of rejected companies	
Year	N	Daily Prices	Weekly Prices	Daily Prices	Weekly Prices
2005	1029	863	244	83.87%	23.71%
2006	1162	972	229	83.65%	19.71%
2007	1322	1196	354	90.47%	26.78%
2008	1366	1292	462	94.58%	33.82%
2009	1385	1261	300	91.05%	21.66%
2010	1471	1166	276	79.27%	18.76%
2011	1568	1387	315	88.46%	20.09%
2007-2011	1322	1322	1196	100%	90.47%

Looking at the results from all the normality tests, it seems reasonable to conclude that the daily data in all the considered years do not follow normal distributions. However, when we consider weekly prices, in the same years, the percentage of company prices where the normal assumption is rejected is much smaller. Then a question arises, the reason for the high/low percentage of rejections is because we consider daily/weekly prices or that difference is due to the difference in the samples dimension? When we consider daily data, we consider ± 252 observations but when we consider weekly prices the sample dimension is 52. To answer this question we repeat the tests using weekly prices from 2007 to 2011, that is, for samples with dimension $5 \times 52 = 260$ and in that situation high rejection percentages are obtained again, as can be verified in the last line in the previous Tables (4, 5 and 6). So it seems that weekly data can be assumed to be normal distributed only when we have small samples. Finally, putting the question on the other way around, that is, what will happen to small samples of daily data? Well, for samples with smaller dimension corresponding to daily prices for the different quarters of the different years of 2010 and 2011 (samples of dimension ± 63), we obtain the results that we present in the Table 7.

Table 7. Samples rejected by the three tests for normality, at 1% level of significance, for daily prices for each quarter of the years of 2010 and 2011

		Number of rejected companies			% of rejected companies		
Quarter	N	K.S.	S.W.	A.D.	K.S.	S.W.	A.D.
First 2010	1471	357	612	503	24.27%	41.60%	34.19%
Second 2010	1471	1456	1550	1539	92.86%	98.85%	98.15%
Third 2010	1471	1340	1455	1435	91.09%	98.91%	97.55%
Fourth 2010	1471	418	669	570	28.42%	45.48%	38.75%
First 2011	1568	384	648	532	24.49%	41.33%	33.93%
Second 2011	1568	1065	1407	1334	67.92%	89.73%	85.08%
Third 2011	1568	702	1081	969	44.77%	68.94%	61.80%
Fourth 2011	1568	228	444	329	14.54%	28.32%	20.98%

For small samples of daily prices we can observe radical changes from one quarter to the other. In some quarter we observe the same behavior as for the weekly prices in small samples, that is, small percentages of rejection, but in others we continue to observe high percentages of rejection of normality.

For a final confirmation of the normality failure, we also tested monthly prices. For monthly prices, we start with small samples, in this situation just considering the data from each year (12 observations) and then we grouped several years in order to build larger samples ($2007 - 2011 \Rightarrow 60$ observations; $2005 - 2011 \Rightarrow 84$ observations; $2000 - 2011 \Rightarrow 144$ observations). The results are presented in Table 8.

Once again, as the samples dimension increases also the percentage of normality rejection increases just as happened in the daily and weekly prices sets of data.

Observation 1. *The reason for the decreasing number of selected companies to test, when we go from the year of* 2011 *to the year of* 2000*, is because we start by considering the companies that belong to the Nasdaq composite index in the year of* 2011 *but not all those companies belong to the index in the previous years.*

Table 8. Samples rejected by the three tests for normality, at 1% level of significance, for monthly prices in different years

		Number of rejected companies			% of rejected companies		
Year(s)	N	K.S.	S.W.	A.D.	K.S.	S.W.	A.D.
2005	1029	24	40	20	2.33%	3.89%	1.94%
2006	1162	19	38	28	1.64%	3.27%	2.41%
2007	1322	22	57	35	1.66%	4.31%	2.65%
2008	1366	38	69	47	2.78%	5.05%	3.44%
2009	1385	38	76	50	2.74%	5.49%	3.61%
2010	1471	18	41	31	1.22%	2.79%	2.11%
2011	1568	25	64	42	1.59%	4.08%	2.68%
2007-2011	1322	188	416	296	14.22%	31.47%	22.39%
2005-2011	1029	224	428	337	21.77%	41.59%	32.75%
2000-2011	752	347	551	478	46.14%	73.27%	63.56%

5. Final remarks

The normality of the logarithm of the stock prices is widely accepted in the financial models. However, our study suggests that the assumption of normality will fail for an high percentage of companies prices from the Nasdaq composite index. This seems to be true even when we consider different observations interval, that is, when we consider weekly prices or even monthly prices. In fact, the number of observations, that is, the sample dimension is more important when testing the normality than the fact of considering daily, weekly or monthly prices. One can argue that to use the pricing Black-Scholes formula we can use the implied volatility (σ parameter) and in that approach we do not need to use the historical prices to estimate it. Even so, the Black-Scholes formula is deduced supposing that the stock prices follows the geometric Brownian motion dynamics and than the normality assumption must be verified.

References

[1] T.W. Anderson and D.A. Darling, *Asymptotic theory of certain goodness-of-fit criteria based on stochastic processes*, Ann. Math. Statist. **23** (1952) 193–212.

[2] T.W. Anderson and D.A. Darling, *A test of goodness of fit*, Journal of the American Statistical Association **49** (268) (1954) 765–769.

[3] Z.W. Birnbaum, *Numerical tabulation of the distribution of Kolmogorov's statistic for finite sample size*, Journal of the American Statistical Association **47** (259) (1952) 425–441.

[4] T. Bjork, Arbitrage Theory in Continuous Time (Oxford University Press, 1998).

[5] D.A. Darling, *The Kolmogorov-Smirnov, Cramer-von Mises tests*, Ann. Math. Statist. **28** (4) (1957) 823–838.

[6] D.E.A. Giles, *A saddlepoint approximation to the distribution function of the Anderson-Darling test statistic*, Communications in Statistics - Simulation and Computation **30** (4) (2001) 899–905.

[7] H.L. Harter, *Expected values of normal order statistics*, Biometrika **48** (1–2) (1961) 151–165.

[8] I. Karatzas and Shreve, Brownian Motion and Stochastic Calculus (Springer-Verlag, 2000).

[9] P.A. Lewis, *Distribution of the Anderson-Darling statistic*, Ann. Math. Statist. **32** (4) (1961) 1118–1124.

[10] H.W. Lilliefors, *On the Kolmogorov-Smirnov test for normality with mean and variance unknown*, Journal of the American Statistical Association **62** (318) (1967) 399–402.

[11] F.J. Massey, *The Kolmogorov-Smirnov test for goodness of fit*, Journal of the American Statistical Association **46** (253) (1951) 68–78.

[12] B. Oksendall, Stochastic Differential Equations (Springer-Verlag, 1998).

[13] J.P. Royston, *An extension of Shapiro and Wilk's W test for normality to large samples*, Journal of the Royal Statistical Society. Series C (Applied Statistics) **31** (1982) 115–124.

[14] J.P. Royston, *A simple method for evaluating the Shapiro-Francia W' test of non-normality*, Journal of the Royal Statistical Society. Series D (The Statistician) **32** (1983) 287–300.

[15] K. Sarkadi, *The consistency of the Shapiro-Francia test*, Biometrika **62** (2) (1975) 445–450.

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

[17] S.S. Shapiro and M. Wilk, *An analysis of variance test for normality (Complete Samples)*, Biometrika **52** (1965) 591–611.

[18] M.A. Stephens, *Use of the Kolmogorov-Smirnov, Cramer-Von Mises and related statistics without extensive tables*, Journal of the Royal Statistical Society. Series B (Methodological) **32** (1) (1970) 115–122.

[19] M.A. Stephens, *EDF statistics for goodness of fit and some comparisons*, Journal of the American Statistical Association **69** (347) (1974) 730–737.

[20] M.A. Stephens, *Asymptotic results for goodness-of-fit statistics with unknown parameters*, Annals of Statistics **4** (2) (1976) 357–369.

[21] M.A. Stephens, Goodness of Fit with Special Reference to Tests for Exponentiality, Technical Report No. 262 (Department of Statistics, Stanford University, Stanford, CA, 1977).

Received 18 May 2012