International Journal of Mathematical Archive-7(6), 2016, 1-7
Available online through www.ijma.info ISSN 2229 – 5046

CHECKING THE NORMALITY ASSUMPTION OF STOCK RETURNS IN SELECTED ASIAN COUNTRIES

PARIMAL KUMAR SEN
Goenka College of Commerce and Business Administration
210, Bepin Behari Ganguly Street, Kolkata, West Bengal 700012, India.

DEBOJYOTI DAS*
UGC Junior Research Fellow
C-29/18 Ananda Nagar, DakshinBehala Road, Kolkata-700061, India.

(Received On: 24-05-16; Revised & Accepted On: 10-06-16)

ABSTRACT

The study endeavors to check whether the log of index stock returns follows the theoretical connotation of normal distribution in selected countries in Asia. The study briefly outlines various important properties of normal distribution. The normality assumption of stock return data of four Asian developing countries (India, Pakistan, Sri Lanka and China) is tested using the Shapiro-Wilk test. The study finds that all of the sample countries appear to fail the normality test.

Keywords: *Normality Assumptions, Shapiro-Wilk Test, Stock Returns.*

2000 Mathematics Subject Classification: *62P05.*

1. INTRODUCTION

In the theory of finance, the assumption that stock returns are normally distributed is expressed far and wide implicitly or explicitly. The existence of this assumption stems out of random walk theory of stock prices, which follows the property that if stock prices follow a random walk, then stock returns should be independently and identically distributed (i.i.d). Further, if the returns are collected the central limit theorem implies that the limiting distribution of these returns should be normal (Aparicio and Estrada, 1997).

In reality, the normality assumption of stock returns is perhaps questionable, if the information does not linearly arrive to the market or even it does, if the investors do not react promptly to its arrival a leptokurtic distribution of stock return is expected to occur. Alternatively, if the information arrives in frequent clusters, the investors might be forced to behave in the similar fashion. In other words, if the distribution of information is leptokurtic, so should be the distribution of stock returns (Peters, 1991). Studies in developing countries in this respect are limited, where the level of information asymmetry is assumed to be higher.

Empirical evidence against the normality assumption, on the other hand, has been mounting since the pioneering articles by Mandelbrot (1963), Fama (1965), and Clark (1973). In the present endeavor it is attempted to substantiate *whether stock returns follows a normal distribution as underlined theoretically or does it contradict to the normality assumption.* The subsequent section will provide a brief outline of Normal Distribution followed by a section dedicated to normality check and goodness of fit test of stock returns.

2. The Normal Distribution

X is normally distributed with parameters μ and σ^2 if the density of X is given by:

$$f(x|\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, -\infty < x < \infty \qquad (2.1)$$

The density function of the normal distribution is a bell shaped curve that is symmetric around its mean.

Corresponding Author: Debojyoti Das, UGC Junior Research Fellow*
C-29/18 Ananda Nagar, DakshinBehala Road, Kolkata-700061, India.

If X is assumed to be normally distributed with parameters μ and σ^2 then $Y = \alpha X + \beta$ is normally distributed with parameters $\alpha\mu+\beta$ and $\alpha^2\sigma^2$. To prove this, suppose first that $\alpha > 0$ and note that $F_Y(\cdot)$* the cumulative distribution function of the random variable Y is given by:

$$\begin{aligned} F_y(a) &= P\{Y \le a\} \\ &= P\{\alpha X + \beta \le a\} \\ &= P\{X \le \tfrac{a-\beta}{\alpha}\} \\ &= F_X\left(\tfrac{a-\beta}{\alpha}\right) \\ &= \int_{-\infty}^{\frac{a-\beta}{\alpha}} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\ &= \int_{-\infty}^{a} \frac{1}{\sqrt{2\pi}\,\alpha\sigma} \exp\left\{\frac{-(v-(\alpha\mu+\beta))^2}{2\alpha^2\sigma^2}\right\} dv \qquad (2.2) \end{aligned}$$

Where the last equality is obtained by the change in variables $v = \alpha x + \beta$. However, since $F_Y(a)= \int_{-\infty}^{a} f_Y(v)dv$, it follows from the equation above that the probability density function $f_Y(\cdot)$ is given by

$$f_Y(v) = \frac{1}{\sqrt{2\pi}\,\alpha\sigma} \exp\left\{\frac{-(v-(\alpha\mu+\beta))^2}{2\alpha^2\sigma^2}\right\}, -\infty < x < \infty \qquad (2.3)$$

2.1. Moment Generating Function (The Normal Distribution with parameters μ and σ^2)

The moment generating function of a standard normal random variable Z is obtained as follows:

$$\begin{aligned} E(e^{tZ}) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{tx} e^{-x^2/2} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2-2tx}{2}} dx \\ &= e^{\frac{t^2}{2}} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(x-t)^2/2}\, dx \\ &= e^{t^2/2} \quad \textbf{[2.4.]} \end{aligned}$$

If Z is a standard normal, then $X = \sigma Z + \mu$ is normal with parameters μ and σ^2; therefore,

$$\varphi(t) = E[e^{tX}] = E\left[e^{t(\sigma Z+\mu)}\right] = e^{t\mu} E[e^{t\sigma Z}] = \exp\left\{\frac{\sigma^2 t^2}{2} + \mu t\right\}$$

By differentiating we obtain,

$$\varphi'(t) = (\mu + t\sigma^2) \exp\left\{\frac{\sigma^2 t^2}{2} + \mu t\right\}$$

$$\varphi''(t) = (\mu + t\sigma^2)^2 \exp\left\{\frac{\sigma^2 t^2}{2} + \mu t\right\} + \sigma^2 \exp\left\{\frac{\sigma^2 t^2}{2} + \mu t\right\}$$

And so,

$$E[X] = \varphi'(0) = \mu$$
$$E[X^2] = \varphi''(0) = \mu^2 + \sigma^2$$

Implying that,

$$\begin{aligned} Var(X) &= E[X^2] - E([X])^2 \\ &= \sigma^2 \qquad (2.5) \end{aligned}$$

2.2. Fisher Information Matrix

Let x be a random variable. Consider the parametric distribution of x with parameter$\boldsymbol{\theta}$, $p(x|\boldsymbol{\theta})$. The continuous random variable ($x \in \boldsymbol{R}$) can be modeled by normal distribution:

$$\begin{aligned} p(x|\boldsymbol{\theta}) &= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{\frac{-(x-\mu)^2}{2\sigma^2}\right\} \\ &= N(x|\mu, \sigma^2) \end{aligned}$$

Where,

$$\boldsymbol{\theta} = (\boldsymbol{\mu\sigma^2})^{\boldsymbol{T}}$$

The Fisher Score is determined as follows:

$$g(\theta, x) = \nabla_\theta \ln p(x|\boldsymbol{\theta})$$

The Fisher Information matrix is defined as follows:

$$\boldsymbol{F} = E_x[g(\theta, x) g(\theta, x)^T]$$

Now, let us calculate the Fisher matrix for the univariate normal distribution. First, we need to take the logarithm:

$$lnN(x|\mu,\sigma^2) = -\frac{1}{2}ln2\pi - \frac{1}{2}ln\sigma^2 - \frac{1}{2\sigma^2}(x-\mu)^2$$

Second, we need to calculate the partial derivatives:

$$\frac{\partial}{\partial\mu}N(x|\mu,\sigma^2) = \frac{1}{\sigma^2}(x-\mu)$$
$$\frac{\partial}{\partial\sigma^2}N(x|\mu,\sigma^2) = -\frac{1}{2\sigma^2} + \frac{1}{2\sigma^4}(x-\mu)^2$$

Hence, we get the following Fisher score for Normal Distribution

$$g(\theta,x) = \begin{cases} \frac{\partial}{\partial\mu}N(x|\mu,\sigma^2) \\ \frac{\partial}{\partial\sigma^2}N(x|\mu,\sigma^2) \end{cases}$$
$$= \begin{cases} \frac{1}{\sigma^2}(x-\mu) \\ -\frac{1}{2\sigma^2} + \frac{1}{2\sigma^4}(x-\mu)^2 \end{cases}$$

Now let us calculate the Fisher score and its transposition,

$$\begin{bmatrix} \frac{1}{\sigma^2}(x-\mu) \\ -\frac{1}{2\sigma^2} + \frac{1}{2\sigma^4}(x-\mu)^2 \end{bmatrix} \begin{bmatrix} \frac{1}{\sigma^2}(x-\mu) & -\frac{1}{2\sigma^2} + \frac{1}{2\sigma^4}(x-\mu)^2 \end{bmatrix}$$
$$= \begin{bmatrix} \frac{1}{\sigma^4}(x-\mu)^2 & -\frac{1}{2\sigma^4}(x-\mu) + \frac{1}{2\sigma^6}(x-\mu)^3 \\ -\frac{1}{2\sigma^4}(x-\mu) + \frac{1}{2\sigma^6}(x-\mu)^3 & \frac{1}{4\sigma^4} - \frac{1}{2\sigma^6}(x-\mu)^2 + \frac{1}{4\sigma^8}(x-\mu)^4 \end{bmatrix} \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix}$$

Where $g_{12}= g_{21}$

In order to calculate the Fisher information matrix we need to calculate the expected values of all g_{ij}. Hence for g_{11}:

$$E_x[g_{11}] = E_x\left(\frac{1}{\sigma^4}(x-\mu)^2\right)$$
$$= \frac{1}{\sigma^2}E_x[x^2] - 2\mu^2 + \mu^2$$
$$= \frac{1}{\sigma^2}(\mu^2 + \sigma^2 - 2\mu^2 + \mu^2 = \frac{1}{\sigma^2}$$

For g_{12}:

$$E_x[g_{12}] = E_x\left(-\frac{1}{2\sigma^4}(x-\mu) + \frac{1}{2\sigma^6}(x-\mu)^3\right)$$
$$= -\frac{1}{2\sigma^4}(E_x[x]-\mu) + E_x\left(\frac{1}{2\sigma^6}(x^3 - 3x^2\mu + 3x\mu^2 - \mu^3)\right)$$
$$= \frac{1}{2\sigma^6}((E_x[x^3] - 3\mu E_x[x^2] + 3\mu^2 E_x[x] - \mu^3))$$
$$= \frac{1}{2\sigma^6}(\mu^3 + 3\mu\sigma^2 + 3\mu(\mu^2+\sigma^2) + 3\mu^3 - \mu^3) = 0.$$

For g_{22}:

$$E_x[g_{22}] = E_x\left(\frac{1}{4\sigma^4} - \frac{1}{2\sigma^6}(x-\mu)^2 + \frac{1}{4\sigma^8}(x-\mu)^4\right)$$
$$= \frac{1}{4\sigma^4} - \frac{1}{2\sigma^6}E_x[x^2 - 2x\mu + \mu^2] + \frac{1}{4\sigma^8}E_x[x^4 - 4x^3\mu + 6x^2\mu^2 - 4x\mu^3 + \mu^4]$$
$$= \frac{1}{4\sigma^4} - \frac{1}{2\sigma^6}(E_x[x^2]) - 2E_x[x]\mu + \mu^2] + \frac{1}{4\sigma^8}E_x[x^4] - 4E_x[x^3]\mu + 6E_x[x^2]\mu^2 - 4E_x[x]\mu^3 + \mu^4]$$
$$= \frac{1}{4\sigma^4} - \frac{1}{2\sigma^6}\sigma^2 + \frac{1}{4\sigma^8}3\sigma^4$$
$$= \frac{1}{2\sigma^4}$$

Finally, we get,

$$\boldsymbol{F} = \begin{bmatrix} \frac{1}{\sigma^2} & 0 \\ 0 & \frac{1}{2\sigma^4} \end{bmatrix} \qquad (2.6)$$

3. DATASET AND ANALYSIS

The dataset consist of stock index data of four Asian countries i.e. India, Pakistan, Sri Lanka [South Asia] and China [East Asia] for a period from 01/12/2014 to 01/12/2015. The return has been calculated as $\ln(P_{t-1}/P_t)$. The key objective of the study is to check the normality assumption of the log index returns of various stock markets in Asia. Presented below is the stock index movement of the sample countries:

Figure-1: Market Index (NIFTY) [India]

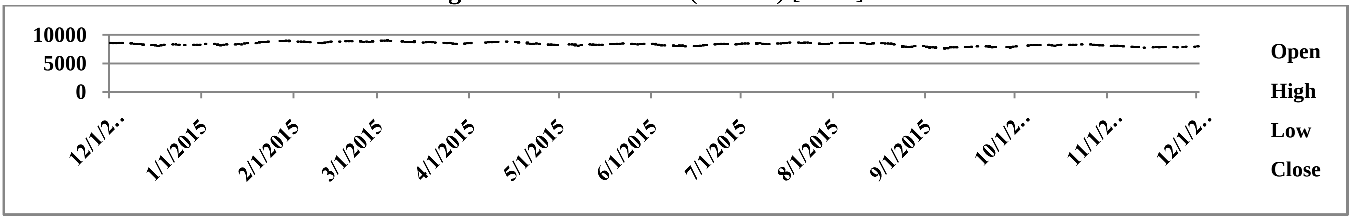


Figure-2: Market Index (Karachi Stock Exchange) [Pakistan]

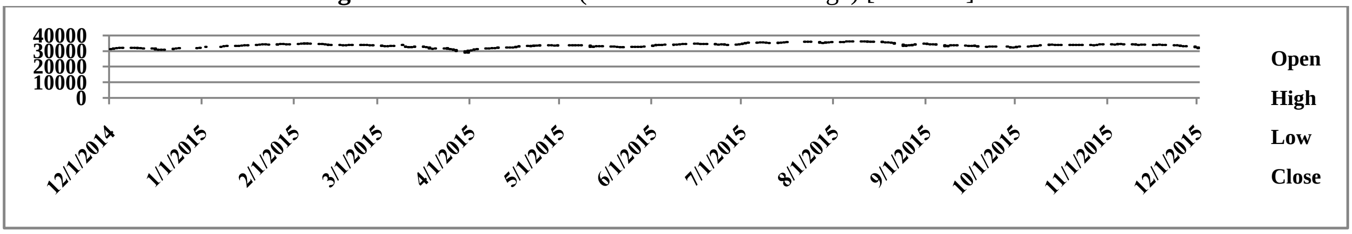


Figure-3: Market Index (Colombo Stock Exchange) [Sri Lanka]

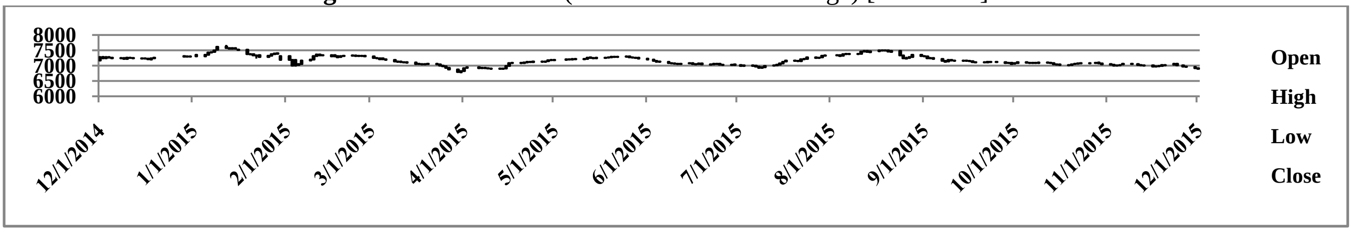
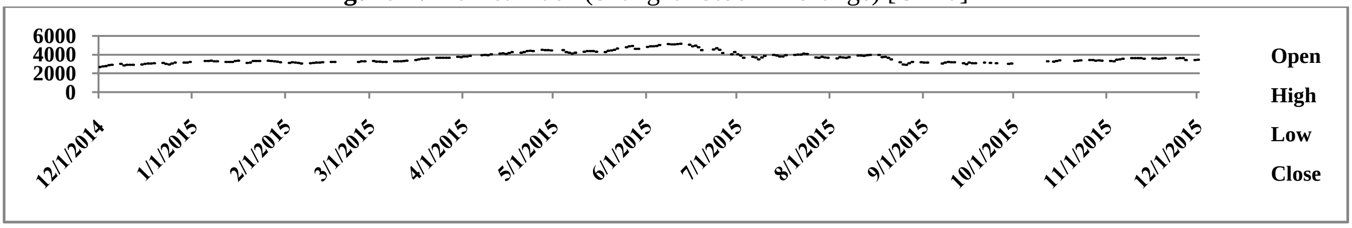
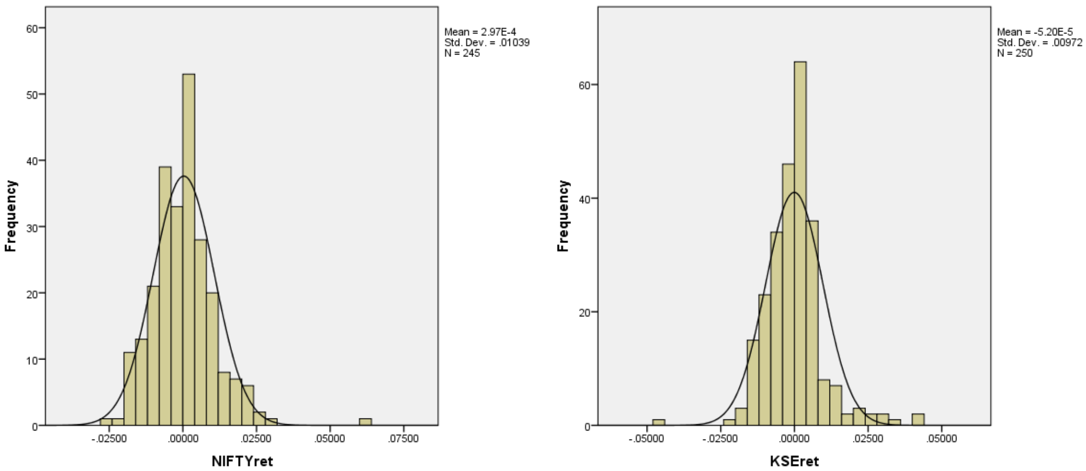


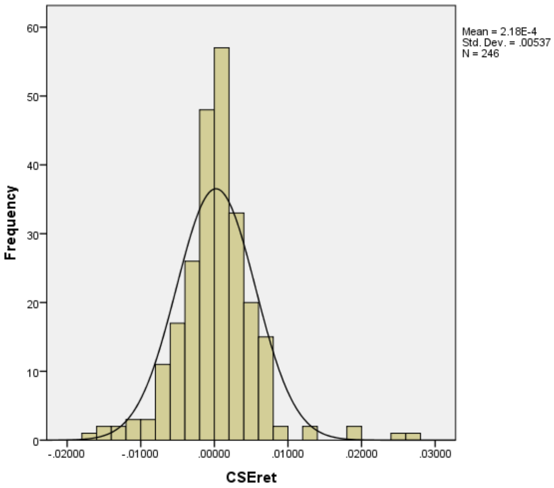
Figure-4: Market Index (Shanghai Stock Exchange) [China]



The visual inspection of the above figures depicts that the markets in India and Pakistan were relatively stable as compared to the remaining two samples over the period of study. However, comparing the magnitude of variability is a difficult task as the index construction parameters may be different. Now we look forward for studying the characteristics of the stock market returns.

Figure-7: Histogram and Normal Curve of Log Index Returns of Sample Countries [NIFTY, KSE, CSE, SSE]





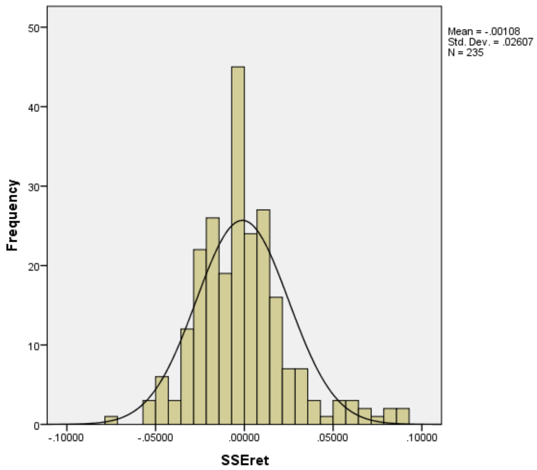


Table-1: Descriptive Statistics and Test of Normality

			Statistic	**Std. Error**				**Statistic**	**Std. Error**
NIFTYret	Mean		0.0001393	0.0006866	**KSEret**	Mean		-0.0000293	0.00063431
	95% Confidence Interval for Mean	Lower Bound	-0.0012134			95% Confidence Interval for Mean	Lower Bound	-0.001279	
		Upper Bound	0.001492				Upper Bound	0.0012204	
	5% Trimmed Mean		-0.0002712			5% Trimmed Mean		-0.0004709	
	Median		0.0000899			Median		0.00E+00	
	Variance		0			Variance		0	
	Std. Deviation		0.01052544			Std. Deviation		0.00972376	
	Minimum		-0.02583			Minimum		-0.04419	
	Maximum		0.06097			Maximum		0.04199	
	Range		0.0868			Range		0.08617	
	Interquartile Range		0.01126			Interquartile Range		0.0097	
	Skewness		1.048	0.159		Skewness		0.716	0.159
	Kurtosis		4.325	0.316		Kurtosis		4.743	0.316

			Statistic	Std. Error				Statistic	Std. Error
CSEret	Mean		0.000198	0.0003554	**SSEret**	Mean		-0.0010822	0.00170049
	95% Confidence Interval for Mean	Lower Bound	-0.0005022			95% Confidence Interval for Mean	Lower Bound	-0.0044325	
		Upper Bound	0.0008982				Upper Bound	0.002268	
	5% Trimmed Mean		0.0001001			5% Trimmed Mean		-0.0023358	
	Median		0.00E+00			Median		-0.0035845	
	Variance		0			Variance		0.001	
	Std. Deviation		0.00544824			Std. Deviation		0.02606798	
	Minimum		-0.01625			Minimum		-0.07412	
	Maximum		0.02699			Maximum		0.08873	
	Range		0.04324			Range		0.16285	
	Interquartile Range		0.00495			Interquartile Range		0.02944	
	Skewness		0.845	0.159		Skewness		0.822	0.159
	Kurtosis		5.215	0.316		Kurtosis		1.742	0.316

4. CHECK OF NORMALITY ASSUMPTIONS

In the present segment the normality assumption of stock returns is verified using descriptive statistics and Shapiro-Wilk test.

NIFTY Returns: The Skewness is found to be 1.013 with S.E. of .156 i.e. the z-value is 6.493 and the kurtosis is 4.342 with S.E. of .310 i.e. the z-value turns out to be 14.006. Thus, both the values are over the range -1.96 to +1.96. Hence the data distribution is kurtotic and skewed.

KSE Returns: The Skewness is found to be .716 with S.E. of .159 i.e. the z-value is 4.503 and the kurtosis is 4.743 with S.E. of .316 i.e. the z-value turns out to be 15.009. Thus, both the values are over the range -1.96 to +1.96. Hence the data distribution is kurtotic and skewed.

CSE Returns: The Skewness is found to be 0.845 with S.E. of .159 i.e. the z-value is 5.314 and the kurtosis is 5.215 with S.E. of .316 i.e. the z-value turns out to be 16.503. Thus both the values are over the range -1.96 to +1.96. Hence, the data distribution is kurtotic and skewed.

SSE Returns: The Skewness is found to be 0.822 with S.E. of .159 i.e. the z-value is 5.16 and the kurtosis is 1.742 with S.E. of .316 i.e. the z-value turns out to be 5.512. Thus both the values are over the range -1.96 to +1.96. Hence, the data distribution is kurtotic and skewed. (Cramer, 1998; Cramer and Howitt, 2004; Doane and Seward, 2011)

The null hypothesis set for the study is that the data are normally distributed against the alternative hypothesis.

The Shapiro-Wilk test found ($p<0.05$) for all the sample countries and hence in this case we fail to accept the null hypothesis. (Shapiro and Wilk, 1965)

Thus, we found that over the period of study the data distribution is not normally distributed.

5. CONCLUDING THOUGHTS AND FUTURE DIRECTION OF RESEARCH

The evidence against the normality assumption of stock returns is mounting over the years of studies. Most of the studies have been conducted in respect of the U.S. data and some on European markets. The present study considers data from Asian markets for testing the hypothesis of normal distribution of stock returns. The study started with describing the data characteristics and an intention to test the hypothesis of normality. Not surprisingly, the data under consideration was found to be with fat tails and high peaks and also skewed in different directions. This result is consistent with previous literature in the same context.

A distribution that fits to the daily returns of the Asian stock market is yet to be explored. A previous study (Aparicio and Estrada, 1997) concluded that scaled-t distribution properly describes the daily return in Scandinavian countries. Similarly, whether this distribution fits in the Asian market or there is a room for further exploration of any other distribution may become a subject matter of further research.

REFERENCES

[1] Aparicio, F., & Estrada, J. (1997). Empirical distributions of stock returns: european securities markets, 1990-95.
[2] Cramer, D., & Howitt, D. L. (2004). *The Sage dictionary of statistics: a practical resource for students in the social sciences*. Sage.
[3] Fama, E. F. (1965). The behavior of stock-market prices. *Journal of business*, 34-105.
[4] Razali, N. M., & Wah, Y. B. (2011). Power comparisons of shapiro-wilk, kolmogorov-smirnov, lilliefors and anderson-darling tests. *Journal of Statistical Modeling and Analytics*, *2*(1), 21-33.
[5] Shapiro, S. S., &Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, 591-611.
[6] Frennberg, P., and B. Hansson (1993). "Testing the Random Walk Hypothesis on Swedish Stock Prices: 1919-1990." Journal of Banking and Finance, 17, 175-191.
[7] Gray, B. and D. French (1990). "Empirical Comparisons of Distributional Models for Stock Index Returns." Journal of Business, Finance & Accounting, 17, 451-459.
[8] Hsu, D. (1982). "A Bayesian Robust Detection of Shift in the Risk Structure of Stock Market Returns." Journal of the American Statistical Association, March 1982, 29-39.
[9] Kon, S. (1984). "Models of Stock Returns - A Comparison." Journal of Finance, 39, 147-65.
[10] Mandelbrot, B. (1963). "The Variation of Certain Speculative Prices." Journal of Business, 36, 394-419.
[11] Peiró, A. (1994). "The Distribution of Stock Returns: International Evidence." Applied Financial Economics, 4, 431-439.

[12] Peters, E. (1991). Chaos and Order in the Capital Markets. A New View of Cycles, Prices, and Market Volatility. John Wiley and Sons.

Source of support: Nil, Conflict of interest: None Declared