New Log Likelihood Estimation Function

Louangrath, P.I. ★

About the author

★Louangrath, P.I. is an Assistant Professor in Business Administration at Bangkok University, Bangkok, Thailand. He could be reached by email at: Lecturepedia@gmail.com

ABSTRACT

This paper provides a New Log-Likelihood Estimator (NLLE) function as a tool for value approximation. We improved the accuracy of the log MLE in two steps (i) determine the log likelihood of a random variable *X*, and (ii) adjust the estimate by a factor of $(1/n-1)(\ln L(X))$. In-Sample testing was accomplished by using daily SET100 indices over a period of 60 days. Out-of-sample data were used for confirmatory verification; out-of-sample data came from 5 major stock markets: NASDAQ, DOW, SP500, DAX, and CAC40. Relevant tests used to compare the results of the proposed NLLE include Cramer-Rao Lower Bound (CRLB), Likelihood Ratio Test, Wald statistic, and Lagrange Multiplier (Score Statistic). It was found that NLLE is more efficient than the conventional MLE. It gives practitioners a better tool for value estimation in many fields of natural and social sciences.

Keywords

Cramer-Rao Lower Bound (CRLB), maximum likelihood estimator (MLE), Monte Carlo, Lagrange multiplier (Score Statistic), likelihood ratio test, log likelihood estimator (LLE), and Wald statistic.

JEL CODE: C10, C13, C14, C46, E27. G11, G17

CITATION:

Louangrath, P.I. (2015). "New Log Likelihood Estimation Function." *Inter. J. Res. Methodol. Soc. Sci.*, Vol., 1, No. 2: pp. 25-35. (Apr. – Jun. 2018).

1. INTRODUCTION

The research question presented in this paper is: "whether the current log-likelihood function is adequately accurate for value estimation?" Researchers always seek better tool for stock price forecast or value estimation. The likelihood function is a common tool for value estimation (Myung, 2003). The current likelihood function lacks precision (Efron, 1981). This paper addresses this weakness in two steps: (i) determine the log likelihood of $(X_i : x_1,...,x_n)$, and (ii) adjust the estimate by a factor of $(n-1)(\ln L(X))$. The data used for this research consists of 5 major stock markets: NASDAQ, DOW, SP500, CAC40, FTSE. over a period of 60 days. The minimum sample size requirement was met by using the *n*-omega method (Louangrath, 2014).

There is a gap in the literature concerning the improvement of likelihood function. For example, there is the need for an accurate forecasting tool to estimate stock price in face of price fluctuation. Under price fluctuation, the current MLE method cannot provide accurate estimation of the expected value. A more accurate estimation tool would have practical application in investment risk management. This paper attempts to fill that gap. In this paper, $\{X_i\}$ represents stock price. Two general cases classified by data types introduce the subject matter of likelihood function and the log-likelihood function as tools for estimating the expected value for random variable $\{X_i\}$.

This paper is presented in five sections. Section 1 introduces the likelihood function and its corresponding log-likelihood as the motivation for stock price analysis in light of price fluctuation. Section 2 explains the data used for the Monte Carlo simulations. Section 3 presents the proposed correction to the log-likelihood estimator and introduces a predictive function to track and forecast stock price under volatile condition. The findings and discussions are covered in section 4. Section 5 concludes the paper by urging practitioners to employ the new LLE as a new tool for stock price estimation and risk management.

2. DATA

The data used for the simulation and testing in this research came from two sources. The first set of data is comprised of the daily CLOSE price of the 100 companies comprising SET100 index for the Stock Exchange of Thailand. The daily prices of these 100 over a period of 30 training sessions were used as a sample. A second set of data consists of the daily indices of 10 major stock markets: (i) NASDAQ; (ii) DOW; (iii) SP500; (iv) CAC40; and (v) FTSE. This second set of data is used for out-of-sample confirmatory verification so that the proposed LLE could be generalized. The generalizability of the proposed LLE is based on the result of consistency in Monte Carlo simulation of 217 iterations using out-of-sample data.

3. METHODOLOGY

The validity and reliability of the old and new log-likelihood functions are compared. The in-sample employed the data from the Stock Exchange of Thailand. Out-of-sample tests were achieved through the use of stock indices taken from 10 major stock markets over a period of at least 60 trading sessions. Comparative results of the following tests were observed: Cramer-Rao Lower Bound (CRLB), Likelihood Ratio Test, Wald statistic, and Lagrange Multiplier (Score Statistic).

The log-likelihood function in the current literature may be written linearly as:

$$\ln L(\mathrm{X}) = \frac{1}{n}\sum\left(\mathrm{X}_i \ln L(F(Z) + (1-\mathrm{X}_i)\ln L(1-F(Z))\right) \tag{1}$$

where X is the observed values $(\mathrm{X}_i : \mathrm{x}_1,...,\mathrm{x}_n)$, Z is the standard score given by $Z = (X_i - \bar{X})/S$ and $F(Z)$ is the percentage probability read from the Z-table at a given critical Z value. In this study, X_i is the CLOSE price of the stock in the SET100 index. For the out of sample test, X_i represents the daily market indices of 5 major stock markets. The result of (1) is called the expected value of series $(\mathrm{X}_i : \mathrm{x}_1,...,\mathrm{x}_n)$ or $E[\mathrm{X}]$. The result of the estimate is compared to the arithmetic means. The arithmetic mean is given by: $\bar{X} = (1/n)\sum X_i$.

The arithmetic mean is a biased estimator if the difference between the arithmetic mean and the expected value is non-zero: $\bar{X} - E[X] \neq 0$. In the learning data of 15 days of trading, the arithmetic mean of the PTT per share price is $\bar{X} = 237.00$, but the estimated value under the log likelihood method is 549.39 Baht per share. The over-shoot is twice the arithmetic mean. The current MLE method does not produce adequately precise value for small sample.

This paper proposes to modify the log likelihood function in order to reduce the bias or inaccuracy between the arithmetic mean and the expected value. The proposed function consists of two steps: (i) determine the log likelihood of $(\mathrm{X}_i : \mathrm{x}_1,...,\mathrm{x}_n)$ under the new LLE, and (ii) adjust the estimate the LLE in (i) by subtracting $\left(\ln L(X)\right)^{-n-1}$. The proposed log-likelihood in step 1 is given by:

$$\ln L(X) = \sum\left[\left(\frac{X_i \ln F(Z)}{n-1}\right)+\left(\frac{(1-X_i)(1-F(Z))}{n^2-n-1}\right)\right] \tag{2}$$

With the estimate obtained in (2), the expected value for $(\mathrm{X}_i : \mathrm{x}_1,...,\mathrm{x}_n)$ is obtained through:

$$E[X] = \frac{1}{n}\sum\left[\left(X_i - \ln(X_i)\right) - \left|\frac{\sum \ln L(X)}{n-1}\right|\right] \tag{3}$$

The result of (3) is approximates equal to $\bar{X} \cong E[X]$ and, thus, the bias for the estimator is minimized. Under this proposed log-likelihood estimator function, the arithmetic mean approximately equals to the expected value. Note that if the denominator of the correcting factor is n such that:

$$\frac{1}{n}\sum\left[\left(X_i - \ln(X_i)\right) - \left|\frac{\sum \ln L(X)}{n}\right|\right] \tag{4}$$

The expected value is observationally equivalent to the arithmetic mean of the series $(\mathrm{X}_i : \mathrm{x}_1,...,\mathrm{x}_n)$, this breaks the first property of *identity* where the condition $\hat{\theta} \to \theta$, but $\hat{\theta} \neq \theta$ must be maintained. Therefore, in (3) following the Bessel correction, the degree of freedom is used. Under the new LLE method, the Fisher information $(I(\theta_0))$ is calculated by $I(\theta_0) = 1/\mathrm{var}[X_i - \ln L(X_i) - |\ln L(X_i)/n-1|]$.

3.1 Monte Carlo Simulation

Monte Carlo is a class of algorithm used in repeated measurement as a simulation tool to approximate the true value (Kroese et al., 2014). One function served by the Monte Carlo simulation is to iterate the measurement until asymptotic normality is achieved or series of estimated values is stabilized. The issue of Fisher information as a means for explaining asymptotic normality is a common approach in the literature. Although theoretically sound, this approach may not be practicable in a case of smaller sample size and data volatility. In the context of stock price movement, investors are expected to see price volatility within a short period. For purposes of risk management, decisions must be made shorter time frame. Therefore, the Monte Carlo iteration size must be adjusted accordingly.

Define that $\hat{\theta} = T$ and $\theta = Z$ then the asymptotic normality condition $\sqrt{n}\left(\hat{\theta}-\theta\right) \to N(0,\sigma^2)$ becomes $\sqrt{n}\left(T-Z\right) \to N(0,\sigma^2)$. Since the sample and population distribution under normality is almost always equal, their relationship may be written as $\sigma^2 = S^2$. Therefore, the argument $\sqrt{n}\left(T-Z\right) \to N(0,\sigma^2)$ may also be written as

$\sqrt{n}\left(S^2 - \sigma^2\right) \to N(0,\sigma^2)$. Under this approach, the variance of the two distributions is used as the basis for the analysis. In order to test that this proposed method to assure that it is reliable, we suggest the use of repeated measurement through Monte Carlo simulation. Monte Carlo is given by:

$$\lim_{n \to R} \Pr\left[\frac{1}{N}\sum_{i=1}^{N} \xi - \mu\right] \le \frac{3\sigma}{\sqrt{n}} = 99.80\% \quad (5)$$

Given a series of random variable $\{X_i\}$, there exists the maximum and minimum values, the error for Monte Carlo is given by:

$$\xi^* = \left(\frac{\max - \min}{2}\right) / 50 \quad (6)$$

The number of simulation needed in Monte Carlo is obtained through:

$$N = \left(\frac{3\sigma_{xi}}{\varepsilon^*}\right)^2 \quad (7)$$

where σ_{xi} is the estimated standard deviation of three values: $X_1 = \max$, $X_1 = \min$, and $X_3 = (\max + \min)/2$. Similarly, μ in (24) $\mu = \bar{X} - \left(T\left(S/\sqrt{n}\right)\right)$ The learning data of PTT price for 15 days produces the following N iteration for the series. In the initial example of 15 days of PTT price, the maxima is 250 and the minima is 209. Thus, $X_1 = 250$; $X_1 = 209$; and $X_3 = (250 + 209)/2 = 459/2 = 229.5$. Use these three X's to determine μ. The standard deviation of the 3 X's is $S = 20.15$ with the corresponding estimated standard deviation of $\sigma_{xi} = 16.74$. For this information, under (7), the expected error is $\xi^* = [(250 - 209)/2]/50 = 0.41$. The Monte Carlo iterations size under (7) is simply: $N = ([3(16.74)]/0.41)^2 = 122.49^2 = 15{,}003$. This means that with a sample of 15 days trading, the Monte Carlo requires 15,003 iterations in order to obtain an acceptable estimated value of stock price.

In the case of 5 major stock market indices out of sample test, the required Monte Carlo iteration is 22,500 runs. The required Monte Carlo iteration for SET100 components is 217 using adjusted result from the new LLE. For practical purpose, this conventional approach may not be practical. In stock trading where decisions are required within shorter span of time and the available of sample size is also smaller, demands for efficiency requires a faster method.

The new LLE method approximates the value of the set as $E[X]$ and compared it to the arithmetic mean of the sample ($\bar{X}$). These two values are used as the maximum and minimum values to obtain the Monte Carlo iteration. The required number of iteration is reduced considerably. This reduction evidences the efficiency of the new LLE function.

4. FINDINGS & DISCUSSION

The proposed new log-likelihood estimation (LLE) method provides a better estimation for the expected value of a given series of random variable $(X_i : x_1, ..., x_n)$. The improvement is evidenced through the result of the following tests: Cramer-Rao Lower Bound (CRLB), Likelihood Ratio Test, Wald statistic, and Lagrange Multiplier (Score Statistic). This improvement was achieved without sacrificing the general requirements of MLE: (i) consistency, (ii) asymptotic normality, and (iii) efficiency.

For the SET100 data set, the general finding is that the proposed new log-likelihood function produces more efficient estimation while retaining acceptable level of Fisher information. When subjected to hypothesis tests, the new LLE performed well.

4.1 Cramer-Rao Lower Bound Test

The Cramer-Rao Low Bound test is used to verify the efficiency of the proposed new LLE method. Efficiency is defined as the optimality of the estimator, i.e. experimental design (Everitt, 2002) or hypothesis testing procedure (Nikulin, 2001). More efficient procedure needs less observations, i.e. if the model is efficient, the required sample size is smaller. The efficiency of an unbiased estimator, T, for parameter θ is defined as:

$$e(T) = \frac{1/I(\theta)}{\mathrm{var}(T)} \tag{8}$$

where $I(\theta)$ is the Fisher information of the sample and $e(T)$ is the minimum possible variance of an unbiased estimator divide by its actual variance (Fisher, 1921). The Cramer-Rao bound is used to prove that $e(T) \le 1$. Efficiency is achieved at $e(T) = 1$. This is proved by the Cramer-Rao inequality for θ. The Cramer-Rao bound is given by:

$$\mathrm{var}(\hat{\theta}) \ge \frac{1}{\mathrm{I}(\theta)} = \frac{1}{-E\left[\frac{\partial^2}{\partial\theta^2}\log f(X \mid \theta)\right]} \tag{9}$$

Table 1. Comparison of Results from Old and New LLE

Market	$\overline{X}$	Old LLE	Diff*	New	Diff**
Dow	17,614.72	13,548.25	-4,066.47	17,708.73	94.01
SP500	2,080.28	1,382.22	-698.06	2,095.27	14.99
NASDAQ	5,036.54	4,473.78	-562.76	5,037.91	1.37
DAX	4,927.64	3,245.74	-1,681.90	4,896.49	- 31.15
CAC40	11,117.15	12,470.79	1,353.64	11,103.28	- 13.87
TWSE	8,476.10	9,847.43	1,371.33	8,562.26	86.16
Heng Seng	24,097.87	27,551.49	3,453.62	24,303.04	205.17
Shanghai	4,032.44	5,389.37	1,356.93	3,693.45	- 338.99
KOSPI	2,025.44	2,063.35	37.91	2,020.38	- 5.06
NIKKEI	20,240.28	- 1,371.32	-21,611.60	20,403.06	162.78

* $\overline{X}$ - Old LLE = Off from the arithmetic mean. ** $\overline{X}$ - New LLE = Off from the arithmetic mean.

The result of the Cramer-Rao test is shown in Table 2. The efficiency achieved under the new log-likelihood function is at or near 1.00. This is consistent with the findings in Table 1 where the accuracy of the old LLE is about 80% and the accuracy of the new LLE is 99.82% which also meets the requirement of the Monte Carlo simulation for $3\sigma/\sqrt{n} = 99.80\%$.

Table 2. Result of the Cramer-Rao Test under Conventional MLE and New LLE

Markets	$Var(\hat{\theta})$	$I(\theta)$	$eT = \frac{1/I(\theta)}{\mathrm{var}(T)}$	$Var(\hat{\theta}) \ge \frac{1}{I(\theta)}$
Dow	308,045.00	0.00000	1.00	Yes
SP500	4,505.00	0.00022	1.00	Yes
NASDAQ	11,268.00	0.00009	1.00	Yes
DAX	51,553.00	0.00002	1.00	Yes
CAC40	15,654.00	0.00006	1.00	Yes

4.2 Likelihood Ratio Test

The likelihood ratio test is based on chi square distribution with degree of freedom of $df = df_2 - df_1$ (Huelbeck, 1997). The ratio calculation is the likelihood of the null divided by the likelihood of the proposed model. The test statistic was given by as $\Lambda(x)$ by Wilk (1938) as:

$$\Lambda(x) = \frac{L(\theta_0 \mid X)}{L(\theta_1 \mid X)} \tag{10}$$

or equivalently:

$$\Lambda(x) = \frac{L(\theta_0 \mid X)}{\sup\{L(\theta \mid X) : \theta \in \{\theta_0 m \theta_1\}\}} \tag{11}$$

where $L(\theta \mid X)$ is likelihood function, *sup* is the supremium function. The decision rule is governed by if $\Lambda > c$ do not reject the null hypothesis and if $\Lambda < c$ then reject the null hypothesis. The rejection point is the probability $\Lambda = c$. The variable c and q are selected at specified alpha (error) level whose relationship may be summarized as: $qP(\Lambda = c \mid H_0) + P(\Lambda < c \mid H_0) = \alpha$. The likelihood ratio test is a tool against Type I error. Type I error occurs when the null hypothesis is wrongly rejected. In the seminal literature, the likelihood ratio test has been classified as a power test (Neyman & Pearson, 1933). Casella and Berger (2011) wrote (10) and (11) as:

$$\Lambda(x) = \frac{\sup\{L(\theta \mid x) : \theta \in \theta_0\}}{\sup\{L(\theta \mid x) : \theta \in \theta\}} \tag{12}$$

Equations (10), (11) and (12) yield the same result.

The calculation for the likelihood ration follows a chi square hypothesis testing. With 60 counts for each market, the ratio is 1 or near one for all markets. This near 1 result shows that the estimation is close to the actual observed value or arithmetic mean. The critical value against which the ratio is test is 79.10. The null hypothesis that the two groups are not significantly different cannot be rejected.

Table 3. Result of the Likelihood Ratio Test under Conventional MLE and New LLE

Market	$L(\theta_0 \mid X)$	$L(\theta_1 \mid X)$	$\Lambda(x)$	$\chi^2(60) = 79.10$
Dow	17,614.72	17,708.73	0.99	Not significant
SP500	2,080.28	2,095.27	0.99	Not significant
NASDAQ	5,036.54	5,037.91	1.00	Not significant
DAX	4,927.64	4,896.49	1.01	Not significant
CAC40	11,117.15	11,103.28	1.00	Not significant

4.3 Wald Statistic

The third test to assess the likelihood function is the Wald statistic. For a single-parameter scenario, the Wald statistic is given by:

$$W = \frac{\left(\hat{\theta} - \theta_0\right)^2}{\operatorname{var}\left(\hat{\theta}\right)} \tag{13}$$

This test is compared to the chi square in case where the data distribution is not normal. In case where the data is normally distributed, the Wald test is given by:

$$W_N = \frac{\hat{\theta} - \theta_0}{se(\hat{\theta})} \quad (14)$$

where se is the standard error of the MLE estimate which is given by:

$$se = \frac{1}{\sqrt{I_n(MLE)}} \quad (15)$$

where I_n is the Fisher information (Harell, 2001, Fears *et. al.*, 1996, Engle, 1983, and Agresti, 2002). The finding of the Wald test in Table 4 shows that there is no significant difference between the arithmetic mean and the estimated mean. The practical implication for stock price analysis is that the new LLE can provide a more accurate estimation.

Table 4. Result of the Wald Test under Conventional MLE and New LLE

Market	θ_0	θ_1	$W = \frac{(\hat{\theta} - \theta_0)^2}{\text{var}(\hat{\theta})}$	$\chi^2(60) = 79.10$
Dow	17,614.72	17,708.73	0.00	Not significant
SP500	2,080.28	2,095.27	0.00	Not significant
NASDAQ	5,036.54	5,037.91	0.00	Not significant
DAX	4,927.64	4,896.49	0.00	Not significant
CAC40	11,117.15	11,103.28	0.00	Not significant

4.4 Lagrange Multiplier (Score Statistic)

The Langrange multiplier test is also called the score test. The score test had been explained by several authors, such as Bera (2001), Lehman and Casella (1998), Engle (1983), and Cook and Demets (2007). The score test is more appropriate where the deviation between $\hat{\theta}$ and θ is small; this is the case of the adjusted log likelihood proposed by this paper. The score test is given by:

$$U(\theta) = \frac{\partial \log L(\theta \mid X)}{\partial \theta} \quad (16)$$

The null hypothesis is $\theta = \theta_0$. If the null hypothesis cannot be rejected, the data is treated as chi square distribution. The test statistic is given by:

$$S(\theta_0) = \frac{U(\theta_0)^2}{I(\theta_0)} \quad (17)$$

where $I(\theta_0)$ is the Fisher information or $I(\theta_0) = -E\left[\frac{\partial^2}{\partial \theta^2} \log L(X \mid \theta) \mid \theta\right]$. For normally distributed data, the score test is given by:

$$S^*(\theta) = \sqrt{S(\theta)} \quad (18)$$

The result of the score statistic shows that the null hypothesis $\theta = \theta_0$ is true for all markets under 95% confidence interval. Only Shanghai index shows a significant difference between the actual and estimated value. Under this test, the new LLE could show 10 our of 10 cases in accuracy.

Table 5. Result of the Score Statistic Test under Conventional MLE and New LLE

Market	θ_0	θ_1	$\theta = \theta_0$	$\chi^2(60) = 79.10$
Dow	17,614.72	17,708.73	-0.5%	Not significant
SP500	2,080.28	2,095.27	-0.7%	Not significant
NASDAQ	5,036.54	5,037.91	0.0%	Not significant
DAX	4,927.64	4,896.49	0.6%	Not significant
CAC40	11,117.15	11,103.28	0.1%	Not significant

4.5 A Proposed Test for New LLE under Chi Square

This paper proposed a test statistic for the new LLE method by using chi square statistic in form:

$$\chi^2 = \left. \frac{(n-1)S^2_{\ln L(X)}}{\sigma^2_{E[X]}} \right|_{df = n-1} \tag{19}$$

where n is the sample size; $S^2_{\ln L(X)}$ is the variance of the new $\ln L(X)$ under equation (18), and $\sigma^2_{E[X]}$ is the estimated variance of the $E[X]$ in (19). The critical value for the null hypothesis is read from the chi square table at degree of freedom $n-1$. Under this proposed test statistic, we are able to achieve more consistent result than the score statistic in Table 5.

Table 6. Result of the LLE Ratio Test under New LLE

Market	$S^2_{\ln L(X)}$	$\sigma^2_{E[X]}$	$\frac{(n-1)S^2_{\ln L(X)}}{\sigma^2_{E[X]}}$	$\chi^2(60) = 79.10$
Dow	308,045.00	1,834,128,509.00	0.01	Not significant
SP500	4,505.00	20,287,376.00	0.01	Not significant
NASDAQ	11,268.00	94,746,147.00	0.01	Not significant
DAX	51,553.00	606,504,707.00	0.01	Not significant
CAC40	15,654.00	143,094,150.00	0.01	Not significant

4.6 Out-of-Sample Test

Findings made from in-sample testing may not be reproduced when an out-of-sample test is conducted (Inuoe & Lutz, 2002). An out-of-sample test is the re-testing of the claimed made by a prior empirical whose conclusion was reach by using in-sample testing. In some cases, the out-of-sample data comes from the original sample where it has been split (Hansen & Timmermann, 2012). In the present case, the out-of-sample test employs a set of data different from the sample use for the hypothesis testing. The out-of-sample test consists of the daily indices of 10 major stock markets: (i) NASDAQ; (ii) DOW; (iii) SP500; and (iv) CAC40. The data was taken from a period of 60 days between June and August 2015.

Table 7. Out-of-Sample Test for Estimated Value under Old and New LLE Methods

Market	Max X_1	Min X_2	* X_3	Mean $\bar{X}$	S	μ
Dow	18,144.07	17,515.42	17,829.75	17,829.75	314.32	17,299.84
SP500	2,124.20	2,046.68	2,085.44	2,085.44	38.76	2,020.10
NASDAQ	5,160.09	4,909.76	5,034.93	5,034.93	125.16	4,823.91
DAX	11,542.54	10,676.78	11,109.66	11,109.66	432.88	10,379.88
CAC40	5,059.17	4,604.64	4,831.91	4,831.91	227.26	4,448.77

*((X1+X2)/2) $= X_3$.

The result of the Monte Carlo for the out-of-sample test follows the following decision rule: $H(0): \Pr\left[\frac{\bar{X}-\mu}{S/\sqrt{n}}\right] \leq \Phi\left(\frac{3\sigma}{\sqrt{n}}\right)$, otherwise H(A) where S is the pool sample standard deviation. The result of the test is summarized in Table 7 below.

Table 8. Monte Carlo Simulation for 5 Major Stock Markets in Out-of-Sample Test

Market	$\bar{X}$	μ	$H(A): \Phi\left(\frac{\bar{X}-\mu}{S/\sqrt{n}}\right)$	$H(0): \Phi\left(\frac{3\sigma}{\sqrt{n}}\right)$
DOW	17,829.75	17,299.84	0.841	0.998
SP500	2,085.44	2,020.10	0.841	0.998
NASDAQ	5,034.93	4,823.91	0.841	0.998
DAX	11,109.66	10,379.88	0.841	0.998
CAC40	4,831.91	4,448.77	0.841	0.998

*Monte Carlo iteration: $n = 217$.

The result of the test shows that the null hypothesis cannot be rejected. The Monte Carlo simulation of 271 iterations produces estimated value within the range of confidence interval at 99.8%. In terms of stock price analysis, this out-of-sample test confirms that the proposed estimating method is generalizable, i.e. apply n situation outside of the empirical data. Recall that the empirical data used in this paper consists of stock prices of 100 companies in the SET100 index. The out-of-sample tests uses data from a different source: indices from 5 major stock markets. The practical implication of this confirmation is that the new log likelihood estimator is generalizable.

5. CONCLUSION

The New Log-Likelihood Estimation (NLLE) function presented in this paper is a novel discovery. It may serve as a better tool for risk management and value estimator. This innovation is a contribution to the field because they fill the gap in the literature and have practical utility. Beyond stock price analysis, NLLE also has general applications in other fields in natural and social sciences.

REFERENCES

DATA: The SET100 data used for this research may be obtained at the share drive at the following page link:
https://drive.google.com/file/d/0B772RxSNNsoVaURhMjMtOEFzN2M/view?usp=sharing

Agresti, Alan (2002). Categorical Data Analysis, (2nd ed.). Wiley. p. 232. ISBN 0471360937.

Bera, Anil K.; Bilias, Yannis (2001). "Rao's score, Neyman's C(α) and Silvey's LM tests: An essay on historical developments and some new results." Journal of Statistical Planning and Inference **97**: 9–44. doi:10.1016/S0378-3758(00)00343-8.

Casella, George; Berger, Roger L. (2001). Statistical Inference (Second ed.). ISBN 0-534-24312-6; p. 375.

Cook, T. D.; DeMets, D. L., eds. (2007). Introduction to Statistical Methods for Clinical Trials. Chapman and Hall. pp. 296–297. ISBN 1-58488-027-9.

Cramér, Harald (1946). Mathematical Methods of Statistics. Princeton, NJ: Princeton Univ. Press. ISBN 0-691-08004-6. OCLC 185436716.

Efron, Bradley (1981). "Maximum Likelihood and Decision Theory." *Technical Report No. 70* (Sep 1981). Stanford University: p. 1.

Engle, Robert F. (1983). "Wald, Likelihood Ratio, and Lagrange Multiplier Tests in Econometrics." In Intriligator, M. D.; and Griliches, Z. Handbook of Econometrics **II**. Elsevier. pp. 796–801. ISBN 978-0-444-86185-6.

Engle, Robert F (1984) . "Wald, Likelihood Ratio and Lagrange Multiplier tests in Econometrics," in *Handbook of Econometrics*, Volume II, Edited by Z. Griliches and M.D. Intriligator. Elsevier Science Publishers BV.

Everitt, Brian S. (2002). The Cambridge Dictionary of Statistics. Cambridge University Press. ISBN 0-521-81099-X.

Fears, Thomas R.; Benichou, Jacques; Gail, Mitchell H. (1996). "A reminder of the fallibility of the Wald statistic." The American Statistician **50** (3): 226–227. doi:10.1080/00031305.1996.10474384.

Fisher, R (1921). "On the Mathematical Foundations of Theoretical Statistics." Philosophical Transactions of the Royal Society of London. Series A **222**: 309–368.

Fisher, R.A., Tippett, L.H.C., (1928). "Limiting forms of the frequency distribution of the largest and smallest member of a sample." *Proc. Cambridge Philosophical Society* **24**:180–190.

Hansen, Peter Reinhard and Timmermann, Allan (2012). "Choice of Sample Split in Out-of-Sample Forecast Evaluation."
http://rady.ucsd.edu/docs/faculty/timmerman/samplesplitmining2012_feb07.pdf
Accessed; January 13, 2016.

Harrell, Frank E., Jr. (2001). "Sections 9.2, 9.3.1., 9.3.3 & 10.5" Regression modeling strategies. New York: Springer-Verlag. ISBN 0387952322.

Huelsenbeck, J. P.; Crandall, K. A. (1997). "Phylogeny Estimation and Hypothesis Testing Using Maximum Likelihood." Annual Review of Ecology and Systematics 28: 437–466. doi:10.1146/annurev.ecolsys.28.1.437.

Inoue, Atsushi and Kilian, Lutz, In-Sample or Out-of-Sample Tests of Predictability: Which One Should We Use? (December 2002). CEPR Discussion Paper No. 3671. Available at SSRN: https://ssrn.com/abstract=369560

Kroese, D. P.; Brereton, T.; Taimre, T.; Botev, Z. I. (2014). "Why the Monte Carlo method is so important today." *WIREs Comput. Stat.* **6**: 386–392. doi:10.1002/wics.1314.

Laplace, Pierre-Simon (1814). *Essai philosophique sur les probabilités*. Paris: Courcier. file:///C:/Documents%20and%20Settings/Administrator/Desktop/NEW%20BOOK/Essai%20philosophique%20sur%20les%20probabilités%20-%20Pierre%20Simon%20marquis%20de%20Laplace%20-%20Google%20Books.html

Lehmann, Erich L.; Casella, George (1998). Theory of Point Estimation, 2nd ed. Springer. ISBN 0-387-98502-6; p. 115.

Myung, In Jae (2003). "Tutorial on Maximum Likelihood Estimation." *Journal of mathematical Psychology*, **47**: 90-100.
Online: http://times.cs.uiuc.edu/course/410/note/mle.pdf

Neyman, Jerzy; Pearson, Egon S. (1933). "On the Problem of the Most Efficient Tests of Statistical Hypotheses." *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **23** (694–706): 289–337.
Bibcode:1933RSPTA.231.289N. doi:10.1098/rsta.1933.0009. JSTOR 91247.

Nikulin, M.S. (2001). "Efficiency of a statistical procedure," in Hazewinkel, Michiel, Encyclopedia of Mathematics, Springer, ISBN 978-1-55608-010-4.