Advances and Applications in Statistics
Published Online: September 2015
http://dx.doi.org/10.17654/ADASJul2015_057_078
Volume 46, Number 1, 2015, Pages 57-78 **ISSN: 0972-3617**

APPLICATION OF BAYESIAN INFERENCE FOR THE ANALYSIS OF STOCK PRICES

Juliet Gratia D'Cunha[1], K. Aruna Rao[1] and T. Mallikarjunappa[2]

[1]Department of Statistics
Mangalore University
Mangalagangothri, Karnataka, India

[2]Department of Business Administration
Mangalore University
Mangalagangothri, Karnataka, India

Abstract

In the analysis of stock market data, it is important to predict the future price of various scripts. Generally, stock prices are analyzed using log transformation. This leads to the assumption that distribution of stock price is lognormal. In this paper we develop Bayes estimator and credible intervals for the mean of the lognormal distribution. The advantage of Bayesian inference is that future stock prices can be predicted using past information as well as the belief of the investor. The validity of proposed method is examined by analyzing the average daily prices (as reported by National Stock Exchange of India limited (NSE)) of 18 scripts listed in NSE for the period of 3 months from October 1st to December 31st, 2013. These scripts belong to 9 sectors namely automobile, bank, cement, finance, FMGC, IT software, pharmaceuticals, real estate, and telecommunication services. Out of these 18 scripts, 11 scripts belong to the list of Bombay Stock

Received: May 19, 2015; Revised: July 7, 2015; Accepted: July 15, 2015

2010 Mathematics Subject Classification: 62F15.

Keywords and phrases: Bayesian inference, stock prices, lognormal distribution, credible intervals, right invariant prior, left invariant prior.

Communicated by K. K. Azad

Exchange (BSE) sensex companies. The validity of the procedure is examined using training and validation data sets. The conclusion is that the data of previous 5 days is sufficient to accurately predict the stock price of the 6th day. A rule has also been suggested for making use of the forecasted value for taking decision by the market analyst and the stock broker and can also be incorporated in automated forecasting.

1. Introduction

Investors in the securities market are interested in predicting the prices for making investment decisions. What determines the price of a financial security is a complex and interesting issue. Investors and researchers have used fundamental and technical approaches to predict the prices. While the fundamental approach relies on the basic strength of a company that issues the underlying security, the technical analysis assumes that the historical price and volume can be used for predicting the future prices. A variety of techniques have been suggested as technical tools to understand the price and volume formations and to predict the future prices based on the past prices. This paper attempts to develop a Bayesian statistical tool to predict the future prices. We revisit the trusted method in Bayesian inference namely predicting the future observations. This approach is mentioned in standard textbooks on Bayesian inference and makes use of probability distribution of the observations under considerations; see for example Berger [4] and Ghosh et al. [13]. This approach is referred to as out of sample prediction in the literature. This approach finds its application in reliability theory also. For details see Kundu and Howlader [25] and Al-Hussaini [1] and the references cited there in. Forecasting stock prices and other economic indicators is well researched area in economics. These approaches use econometric models of time series for making the prediction. Section 2 provides review of literature in this area. This type of forecasting is referred to as economic forecasting (a term used in Litterman [27]) which differs from technical forecasting (Granger [15]). Technical forecasting is generally done using automated computer packages and is often employed by the market consultants. In this paper, we propose the Bayesian forecasting of future observation as a technical forecasting method. This method is based on scientific theory and inherits the features of technical forecasting. Unlike the fundamental forecasting, this forecasting approach does not require large amount of data for making the prediction and is based only on the past behaviour of the stock prices and does not involve any economic models. Most of the literatures on stock prices assume that log of the prices follow normal distribution, which in turn amounts to the assumption that distribution of the stock prices is lognormal. This is rigorously investigated in Antoniou et al. [3] who show that detrended stock prices follow lognormal distribution. Lognormal distribution has support on the positive part of the real line. It is right skewed and is widely applicable when normal distribution does not fit well to the data. It is used in the analysis of failure time data (Kalbfleisch and Prentice [22]; Lawless [26]), stock market data (Antoniou et al. [3]) and in the analysis of rainfall data (Ananthakrishnan and Soman [2]). The commonly used procedure in this area is to take log transformation and use the technique developed for normal models. The mean and variance of the lognormal distribution are not invariant under distributional transformation. Therefore it is necessary to develop estimators and confidence intervals for the mean and variance of the lognormal distribution.

In the analysis of stock prices, one of the problems of interest is to predict the future value of the stock prices. In Bayesian analysis one can predict the value of X_{n+1} given the observations $X_1, ..., X_n$. The estimated value of X_{n+1} is the expected value of θ, the mean of the distribution of θ, where the expectation is taken with respect to the posterior density $\pi(\theta | x)$, $x = x_1, ..., x_n$ (Berger [4]). In this paper we derive maximum likelihood and Bayes estimator of the mean of the lognormal distribution and derive the associated $100(1-\alpha)\%$ confidence/credible interval. The use of maximum likelihood estimator is undertaken as it is a more efficient estimator compared to the historical average. For such type of comparison see Campbell and Thompson [7] and the references cited therein. To check the accuracy of the proposed approach, we have carried out the analysis of stock price of the 18 scripts of the National Stock Exchange for the period of 3 months from 1st October to 31st December, 2013. We have used the short term data as required in technical forecasting, and therefore comparison with the well known approaches like autoregressive integrated moving average (ARIMA), Bayesian vector autoregression (BVAR) procedure or the artificial intelligence based methods cannot be made as they require long term data. The analysis shows that we can accurately predict the value of the stock price given the information for the last 5 days. The accuracy of the prediction does not improve if we change the past data to 10 or more days. The results also show that although Bayesian procedure outperforms the classical procedure, there are cases where maximum likelihood method performs well. Such a conclusion has also emerged in the analysis of macroeconomic indicators (see for example Campbell and Thompson [7]). The results also indicate that the proposed method is robust against the specification of the prior distribution. The same conclusion emerges if we use right invariant and left invariant objective priors. The organization of the paper is as follows:

In Section 2, we provide brief review of the various forecasting approaches and in Section 3, we derive Bayes estimator and the associated credible interval for the mean of the lognormal distribution. This procedure is used for predicting the future value of the stock prices using training set and the validation data set. The results are presented in Section 4. Section 5 gives concluding remarks and directions for future research.

2. Brief Review of Literature

The papers on forecasting can be classified into 3 categories namely (a) methods based on time series and econometric models, (b) Bayesian forecasting and (c) methods based on artificial intelligence. In the sequel we provide a brief review on the research in these 3 areas.

(a) *Time series and econometric models*

Brown and Warner [6] proposed a methodology to use of daily stock returns on predicting the excess return and event date. A combination of methods is used to estimate the excess return. They used regression models.

Forecasting stock market prices and their efficiency in the stock market was discussed by Virtanen and Yli-Olli [30]. They show that monthly and quarterly stock market prices can be adequately forecasted using either univariate time series analysis or multivariate econometric models. The analysis of Helsinki stock market data indicates that econometric models slightly outperformed the univariate ARIMA model. Granger [15] critically examines the various forecasting procedures till that date. They include the various versions of the autoregressive time series models as well as the technical or automated forecasting models of Brock et al. [5]. His recommendation to the forecasters was to break away from simple linear univariate ARIMA or multivariate transfer functions. Ding et al. [10] investigate a long memory property of stock market returns. They compare the variance of autoregressive conditional heteroskedasticity (ARCH) and generalized autoregressive conditional heteroskedasticity (GARCH) models. From the Monte Carlo study undertaken by them they conclude that both ARCH type models based on squared returns and those based on absolute return could produce this long memory property. Goyal and Welch [14] argue that the historical average excess stock return forecasts accurately the future excess stock returns compared to regression models based on excess stock returns. This claim was re-examined by Campbell and Thompson [7]. They show that restricted least square prediction outperforms the historical average in forecasting excess stock returns and dividend-price ratio.

(b) *Bayesian forecasting*

In this subsection, we review the research papers dealing with Bayesian forecasting of economic indicators. Mainly these are the papers where they use Bayesian procedures to estimate the parameters of time series and econometric models and to make out of sample prediction.

One of the earliest works on Bayesian forecasting using time series model is that of Harrison and Stevens [16]. They consider a growth model based on ARIMA, and Bayesian estimation is used for making short term prediction. They have not analyzed the real data and only use artificial data to check the prediction accuracy of their approach. In subsequent papers, Harrison and Stevens [17] extend their work for multi-process model. A rigorous treatment of Bayesian estimation and out of sample forecasting was taken by Thompson and Miller [29]. Although Bayesian estimation of the parameters of this model does not provide difficulties, the prediction of out of sample observation is computationally challenging. Using Monte Carlo integration and the importance sampling, they estimate the predictive density and provide the confidence interval for this density. Litterman [27] discuss the five year performance of Bayesian vector autoregression (BVAR) which was originally proposed by him. The basic model is the autoregressive model, and BVAR provides the prediction of out of sample observations. He considers the data on quarterly observations on the following seven variables: Annual growth rate of gross national product (GNP), annual inflation rates, the unemployment rate, logged levels of money supply, logged levels of gross private domestic investment, the rate on four to six month commercial paper, and the change in business inventories. They arise at the conclusion that BVAR is as accurate, on average, as provided by the best known commercial forecasting services. Harrison and West [18] discuss a dynamic variance component linear regression model for making forecast. Bayesian procedure is used for making the prediction. The procedure is general in nature and not necessarily developed for stock prices. This paper deals with application of model with reference to demand data in a major industrial sector and quarterly retail sales of a major confectionary product for a period over 11 years. Forecasting models based on large datasets is the topic discussed in Kapetanios et al. [24]. They compare 3 approaches namely factor analysis, Bayesian model averaging methods and the non Bayesian model averaging procedure based on decision theoretic approach. Their analysis shows that in some situations the information theoretic approach due to Akaike is better compared to the Bayesian model averaging scheme and to factor models. A comprehensive paper on Bayesian forecasting is due to Geweke and Whiteman [12]. This is a chapter in a book, and provides review of the Bayesian forecasting procedure till that date. The focus of this chapter is the Bayesian vector autoregression procedures and their variance. Never the less, other Bayesian forecasting approaches are also discussed and provide an extensive material to those who work on Bayesian forecasting. Wright [31] discussed the use of Bayesian model averaging procedure to predict out of sample exchange rate. A Monte Carlo investigation conducted by them reveals that the forecast yield results which are close to the results obtained by random walk procedure.

(c) *The procedure based on artificial intelligence*

The procedures discussed in the preceding paragraphs correspond to the model developed by economists and at times by statisticians. The procedure based on artificial intelligence is mainly developed by the computer scientist and is used to predict the macroeconomic indicators and stock market data. Oh and Kim [28] discuss four component model for analysis of stock market indices. The first phase of the core component of the model determines time-lag size in the input variables using chaotic analysis. The second phase detects successive change-point in the stock market data and the third phase forecasts the change-point group with back propagation neural networks (BPNs). The fourth or the final phase forecasts the output using BPN. Hidden Markov model (HMM) is used by Hassan and Nath [19] to forecast some of the airlines stock. The trained HMM is used to search for the variable of interest behavioural data pattern from the past dataset and to make future prediction by interpolating the neighbouring values. Garcia-Almanza and Tsang [11] discuss the decision tree based approach for making forecast of stock price. They use genetic algorithm for this purpose, which accommodates the chance discovery. A fusion model based on hidden Markov model (HMM), artificial neural network (ANN) and genetic algorithm (GA) was developed by Hassan et al. [20] to forecast financial market behaviour. ANN is used to develop a set of independent variables which become input to HMM. The initial parameters of the model were estimated using GA. The trained HMM is used to identify the price of the current day. The price of the next day of this identified day is the predicted value of the price of the stock under consideration. Weighted average of such forecast was taken as the forecast for the next day. Chang and Liu [8] developed Takagi-Sugeno-Kang (TSK) model to predict the Taiwan electronic shares. The forecasts were 97.6% accurate in Taiwan Stock Exchange (TSE) index and 98.08% in MediaTek. Chu et al. [9] use a dual factor modified fuzzy time series model, which take stock index and trading volume as forecasting factors to predict stock index. Wavelet transform and recurrent neural network was used to develop a forecasting model by Hsieh et al. [21]. Wavelet transformation is used to eliminate the systematic component namely, the seasonal effect. Using the transformed data the predictors for the model are then selected using stepwise regression. Artificial Bee colony (ABC) algorithm is then used for making the forecast. The authors use simulated data sets to test their approach.

3. Methodology

Let μ and σ denote the log location and scale parameter of the lognormal distribution. Given a random sample of size n, $x_1, \ldots, x_n$ from this distribution, let $Z_i = \log X_i$, $i = 1, \ldots, n$, where Z follows normal distribution with parameter μ and σ and maximum likelihood estimator of μ and σ^2 are $\bar{Z}$ and S_Z^2, respectively, where $\bar{Z} = \frac{\sum_{i=1}^{n} Z_i}{n}$ and $S_Z^2 = \frac{\sum_{i=1}^{n} (Z_i - \bar{Z})^2}{n}$. Using invariance property of maximum likelihood estimator (Kale [23]), the maximum likelihood estimate (MLE) of the mean of the lognormal distribution namely $\theta = e^{\mu + \frac{1}{2}\sigma^2}$ is given by $\hat{\theta} = e^{\hat{\mu} + \frac{1}{2}\hat{\sigma}^2} = e^{\bar{Z} + \frac{1}{2} S_Z^2}$. The asymptotic variance of $\hat{\theta}$ can be obtained using delta method and is given by

$$\operatorname{var}(\hat{\theta}) = e^{2\mu + \sigma^2} \frac{\sigma^2}{n} \left(1 + \frac{\sigma^2}{2}\right). \tag{1}$$

In the above expression, we have used $\frac{n-1}{n} \approx 1$. The $100(1-\alpha)\%$ asymptotic confidence interval for $\hat{\theta}$ is given by $\hat{\theta} \pm Z_{\alpha/2} S.E(\hat{\theta})$, where $Z_{\alpha/2}$ refers to upper $\alpha/2$th percentile value of the standard normal distribution (critical value) and $S.E(\hat{\theta})$ refers to estimated *standard error*$(\hat{\theta})$.

The estimate of μ and σ^2 is obtained by substituting the value of $\hat{\mu}$ and $\hat{\sigma}^2$ in the expression for variance of $\hat{\theta}$.

We have used two objective priors namely the right invariant prior $\pi(\mu, \sigma) = \frac{1}{\sigma}$ and left invariant Jeffrey's prior given by $\pi(\pi, \sigma) = \frac{1}{\sigma^2}$ (Ghosh et al. [13]; Berger [4]). The choice of the right invariant prior stems from the fact that Z follows normal distribution and the right invariant prior used in this paper is the one that is suggested for location scale family (Ghosh et al. [13]). The advantage of objective Bayesian analysis is that the prediction remains the same irrespective of the decision maker. Thus the procedure can be applied universally given the past data. The posterior density $\pi(\mu, \sigma | z_1, ..., z_n)$ for the right invariant and the left invariant Jeffry's priors are given by the following expression:

$$\pi(\mu, \sigma | z_1, ..., z_n)$$

$$= \frac{1}{\sqrt{2\pi}\,\frac{\sigma}{\sqrt{n}}} e^{-\frac{1}{2}\frac{(\bar{Z}-\mu)^2}{\sigma^2/n}} \frac{\left(\frac{(n-1)S_Z^2}{2}\right)^{\frac{n+2}{2}}}{\Gamma\left(\frac{n+2}{2}\right)} \left(\frac{1}{\sigma^2}\right)^{\left(\frac{n+2}{2}\right)-1} e^{-\frac{1}{2}\left(\frac{n-1}{\sigma^2}\right)S_Z^2}$$

(using right invariant prior), (2)

$$\pi(\mu, \sigma | z_1, ..., z_n)$$

$$= \frac{1}{\sqrt{2\pi}\,\frac{\sigma}{\sqrt{n}}} e^{-\frac{1}{2}\frac{(\bar{Z}-\mu)^2}{\sigma^2/n}} \frac{\left(\frac{(n-1)S_Z^2}{2}\right)^{\frac{n+3}{2}}}{\Gamma\left(\frac{n+3}{2}\right)} \left(\frac{1}{\sigma^2}\right)^{\left(\frac{n+3}{2}\right)-1} e^{-\frac{1}{2}\left(\frac{n-1}{\sigma^2}\right)S_Z^2}$$

(using left invariant prior). (3)

It may be noted that although we have used independent prior for μ and σ, the posterior density has a bivariate correlated distribution. The Bayes estimator of θ is $E(\theta | z_1, ..., z_n)$, where expectation is taken with respect to the posterior density of μ and σ. Closed form solution does not exist for this expectation and has to be evaluated using numerical integration or the Monte Carlo integration. In this paper we have used Monte Carlo integration to compute Bayes estimate. The samples of μ, σ^2 are generated using importance sampling approach. Since the marginal distribution of $\eta = \frac{1}{\sigma^2}$ follows gamma distribution for the right as well as left invariant priors, the observation for η is generated from the gamma distribution. The conditional posterior distribution of μ given σ^2 follows normal distribution with mean $\bar{Z}$ and variance $\frac{\sigma^2}{n}$. Using the previously generated value of σ^2, we generate an observation from the normal distribution. This constitutes a pair of observation μ, σ^2 from the posterior density. Then $E(e^{\mu + \frac{1}{2}\sigma^2} | z_1, ..., z_n) = \frac{1}{M}\sum_{i=1}^{M}(e^{\mu_i + \frac{1}{2}\sigma_i^2})$, where M denotes the number of paired samples generated from the posterior distribution. In this paper we have used $M = 10{,}000$. The generated samples were also used for obtaining equitailed credible intervals.

The forecasted price can be used to suggest a rule for the market analysts or the stock brokers. As in Brock et al. [5], form a band $B_t = (1 \pm 0.01)M_t$, where M_t is the average of previous prices, if the price is above the upper limit of the band do not sell the stock and if the price falls below the lower limit of the band to sell the stock.

4. Numerical Analysis

4.1. Bayes estimator

To evaluate the accuracy of the Bayes prediction we have carried out a numerical investigation using the approach of training and validation data sets. For this purpose, we have used the daily data on the average stock prices of 18 scripts belonging to NSE for the period of 3 months from October 1st to December 31st, 2013. We have identified 9 sectors namely automobile, bank, cement, finance, FMGC, IT software, pharmaceuticals, real estate, and telecommunication services. From each sector 2 scripts are selected. Out of these 18 scripts, 11 scripts belong to the list of BSE sensex companies. Some of the smaller companies are also included in the investigation to check the accuracy of the procedure. In each month the date selected for the prediction purpose are October 31st, November 25th and December 31st, 2013. To estimate the duration of the past data which gives accurate prediction, the average value of the absolute difference between the predicted and the average price (as reported by NSE) of the stock prices was computed. Tables 1(a) and 1(b) summarize the results.

Table 1(a). Average of the absolute difference between the estimated price and the average price as reported by NSE

Sr. no.	Stock	Length of the period	Average of the absolute difference	
			Right invariant prior	Left invariant prior
1	Tata Motors	5	2.53	2.99
		10	2.72	3.63
		15	2.35	2.74
		20	5.19	5.15
2	Maruti Suzuki	5	30.66	34.43
		10	46.66	51.18
		15	57.52	63.15
		20	74.75	80.58
3	HDFC Bank	5	3.27	3.88
		10	5.67	6.04
		15	11.15	10.38
		20	14.60	14.17

4	ICICI Bank	5	20.50	23.40
		10	31.35	34.55
		15	41.38	40.18
		20	46.66	45.80
5	ACC	5	45.26	17.16
		10	19.22	7.12
		15	23.96	12.62
		20	11.92	9.20
6	Ambuja	5	2.52	2.36
		10	3.84	3.99
		15	3.17	2.68
		20	3.99	3.39
7	Reliance Capital	5	4.74	4.82
		10	4.18	5.18
		15	7.75	7.09
		20	10.66	10.18
8	Muthoot Finance	5	1.90	1.94
		10	4.29	4.38
		15	4.65	4.62
		20	4.18	4.23
9	Hindustan Unilever	5	4.04	4.65
		10	4.23	4.54
		15	7.04	7.12
		20	6.88	6.91
10	Godrej Consumer	5	10.74	13.20
		10	15.58	18.02
		15	19.28	21.01
		20	15.54	17.29

11	Infosys	5	34.50	30.44
		10	22.77	18.94
		15	24.05	28.38
		20	50.82	56.72
12	Wipro	5	4.75	4.42
		10	11.83	11.59
		15	14.17	14.19
		20	16.52	16.74
13	Cipla	5	4.14	3.41
		10	8.83	8.44
		15	13.21	12.76
		20	15.09	14.63
14	Dr. Reddy's Lab	5	20.35	23.21
		10	34.38	37.22
		15	41.85	43.88
		20	44.95	46.45
15	India Bulls Real Estate Ltd	5	1.11	0.66
		10	0.83	0.97
		15	0.90	1.00
		20	1.16	1.46
16	Oberoi Realty Ltd.	5	2.12	2.20
		10	3.94	4.55
		15	6.42	7.02
		20	9.61	10.71
17	Idea	5	2.54	1.56
		10	2.97	2.76
		15	4.56	4.42
		20	4.50	4.33

18	Bharti Airtel	5	9.45	9.77
		10	9.09	9.27
		15	12.39	12.49
		20	13.54	13.83

Table 1(b). Frequency distribution of scripts for the duration at which the absolute difference is minimum under Bayes estimation

Length of the period	Frequency of count	
	Right invariant prior	Left invariant prior
5	12	13
10	4	4
15	1	1
20	1	0

From the table, it follows that average of the absolute difference is minimum for 5 days. It is 12 times out of 18 for the right invariant prior and 13 times out of 18 for the left invariant prior. The scripts which require larger data sets for accurate prediction are Tata Motors and ACC. These scripts belong to automobile and cement sector. Table 2 provides the Bayes estimator and the associated 95% credible interval using the data for the last consecutive 5 days for all the scripts along with the average price (as reported by NSE) and predicted price of the script for the date under consideration which is October 31st, 2013. From this table and the tables not reported in this paper, for the other two dates namely November 25th, 2013 and December 31st, 2013 it follows that in 50 cases out of 54 (18 scripts × 3 months) using right invariant prior, the actual price lies inside the credible interval. From the tables not reported here we notice that using left invariant prior in 53 cases out of 54, the length of the credible interval is shorter than the length of the confidence interval.

Table 2. Prediction of stock price for October 31st, 2013 (the 6th day) using Bayes estimation

Script name	Average Price as reported by NSE	Bayes estimator		Credible interval	
		Right invariant	Left invariant	Right invariant	Left invariant
Tata Motors	382.43	381.06	379.10	(364.02, 399.70)	(371.77, 386.61)
Maruti Suzuki	1631.2	1572.10	1562.30	(1483.90, 1669.40)	(1525.20, 1602.00)
HDFC Bank	679.37	674.74	673.84	(665.91, 684.11)	(670.11, 677.50)
ICICI Bank	1107.95	1052.10	1047.20	(1008.20, 1101.30)	(1028.50, 1066.10)
ACC	1133.82	1144.30	1140.60	(1110.30, 1179.70)	(1126.3, 1155.20)
Ambuja	188.4	193.81	192.99	(186.58, 201.87)	(189.92, 196.10)
Reliance Capital	366.33	357.25	356.04	(347.51, 367.6018)	(351.85, 360.29)
Muthoot Finance	106.87	102.30	101.94	(99.13, 105.75)	(100.59, 103.31)
Hindustan Unilever	609.31	606.81	603.77	(580.44, 635.75)	(593.05, 615.14)
Godrej Consumer	869.46	849.56	845.09	(809.43, 893.91)	(828.13, 862.44)
Infosys	3313.49	3328.30	3327.70	(3322.20, 3335.10)	(3325.00, 3330.20)
Wipro	481.04	481.89	481.30	(476.40, 487.76)	(479.00, 483.63)
Cipla	414.25	417.35	416.82	(412.45, 422.50)	(414.77, 418.89)
Dr. Reddy's Lab	2483.94	2429.60	2421.50	(2357.00, 2510.20)	(2390.60, 2452.60)
India Bulls Real Estate Ltd.	62.61	63.10	62.83	(60.61,65.71)	(61.78, 63.90)
Oberoi Realty Ltd.	182.43	182.36	181.80	(177.34,187.92)	(179.66, 183.95)
Idea	172.7	173.96	171.48	(154.48,196.12)	(163.11, 180.47)
Bharti Airtel	365.72	350.41	347.57	(325.46,378.48)	(336.93, 358.55)

4.2. Maximum likelihood estimator

Given a sequence of observations $Y_1, Y_2, \ldots, Y_n,$ in classical inference the mean of the random variable is taken as the predicted value of Y_{n+1}. In the case of lognormal distribution, the maximum likelihood estimate of the mean of the distribution is given by $e^{\bar{Z}+\frac{1}{2}S_Z^2}$. MLE is consistent and is asymptotically unbiased. The variance of the estimator asymptotically attains Cramer Rao lower bound for the variance of an unbiased estimator and is an efficient estimator. Therefore the predictive power of the Bayes estimator is compared with the MLE in terms of coverage probability and length of the confidence/credible interval. To compare MLE with the Bayes predicted value we have carried out a similar exercise as in the case of Bayes prediction. From the tables not reported here, it follows that even with maximum likelihood estimator, 5 days data is sufficient to predict the stock price of the 6th day (for 13 scripts out of 18). Table 3 presents the predicted and the average price (as reported by NSE) for the 6th day using maximum likelihood estimator. When we compare the average absolute difference in the case of Bayes and ML prediction, using previous consecutive 5 days data it is observed that out of 18 scripts, for 10 scripts under right invariant prior, 2 scripts under left invariant prior, 3 scripts under MLE have smaller average absolute difference compared to other two. The average absolute difference for the left invariant prior is at par with MLE in 3 scripts. Thus, 15 scripts out of 18 perform well in Bayes prediction compared to the MLE and thus this estimator is recommended for prediction purpose.

Table 3. Prediction of price for October 31st, 2013 (the 6th day) using maximum likelihood estimation

Script name	Average price as reported by NSE	MLE	Confidence interval
Tata Motors	382.43	379.01	(369.19, 389.08)
Maruti Suzuki	1631.2	1562.30	(1511.00, 1613.60)
HDFC Bank	679.37	673.85	(668.91, 678.78)
ICICI Bank	1107.95	1047.20	(1021.40, 1073.10)
ACC	1133.82	1140.50	(1121.30, 1160.00)
Ambuja	188.4	192.96	(188.85, 197.16)
Reliance Capital	366.33	356.05	(350.36, 361.73)
Muthoot Finance	106.87	101.95	(100.13, 103.76)

Hindustan Unilever	609.31	605.63	(589.10, 618.76)
Godrej Consumer	869.46	844.99	(822.42, 868.05)
Infosys	3313.49	34.50	(3324.10, 3331.20)
Wipro	481.04	481.30	(478.20, 484.39)
Cipla	414.25	416.82	(414.02, 419.61)
Dr. Reddy's Lab	2483.94	2421.50	(2379.10, 2464.00)
India Bulls Real Estate Ltd.	62.61	62.30	(61.41, 64.26)
Oberoi Realty Ltd.	182.43	181.80	(178.89, 184.71)
Idea	172.7	171.18	(160.19, 182.94)
Bharti Airtel	365.72	347.38	(332.95, 362.34)

Table 4 presents the range in the average price (as reported by NSE) of the 18 scripts for the month October 2013.

Table 4. Range in the average price (as reported by NSE) of the 18 scripts for the month October, 2013

Script Name	**Range**			
	5 days	**10 days**	**15 days**	**20 days**
Tata Motors	4.65	12.17	35.50	53.92
Maruti Suzuki	126.34	206.87	245.81	254.72
HDFC Bank	13.01	31.50	41.79	77.58
ICICI Bank	71.04	131.33	167.90	196.64
ACC	30.61	30.61	39.20	51.92
Ambuja	2.83	9.51	9.51	20.30
Reliance Capital	15.06	27.77	27.77	46.78
Muthoot Finance	5.10	10.57	10.57	14.27
Hindustan Unilever	16.22	22.30	22.30	22.30
Godrej Consumer	31.05	38.07	46.52	69.35

Infosys	8.61	48.09	283.70	329.66
Wipro	8.60	36.49	36.49	40.10
Cipla	8.50	8.50	21.83	28.22
Dr. Reddy's Lab	120.51	120.51	120.51	120.51
India Bulls Real Estate	3.81	7.54	9.44	13.41
Oberoi Realty Ltd.	7.06	14.75	14.75	17.66
Idea	7.79	15.61	16.96	16.96
Bharti Airtel	14.49	14.49	26.12	37.96

From Table 4, it follows that for all the 18 scripts the range of the prices is minimum for 5 days. A similar picture emerges for the month of November and December, 2013. Range is a measure of volatility in the stock prices. The volatility in the data affects the prediction accuracy and explains the reason why previous 5 days data is sufficient to accurately predict the average price for the 6th day.

5. Conclusion

In this paper, we have derived the Bayes estimator and the associated $100(1-\alpha)\%$ credible interval for the mean of the lognormal distribution. The advantage of Bayesian procedure is that the market uncertainties can be captured by treating the parameters as random and specifying prior distribution for the parameters. The procedure is illustrated by analyzing the future daily stock price of 18 scripts listed in National Stock Exchange of India limited. Out of these 18 scripts, 11 scripts are the one from the list of 31 BSE sensex companies. Some smaller companies were also included in the study so as to check the accuracy from the suggested procedures. It is recommended to the practitioners to use Bayes procedure to estimate the future price of the 6th day using the previous consecutive 5 days prices. This generalization is done so as to suggest an accurate procedure for the practitioners, although in 3 out of 18 scripts the estimate obtained from the MLE is slightly better than the Bayes estimator. Sometimes the historical averages perform well as noted in Campbell and Thompson [7]. In the case of lognormal distribution, the maximum likelihood estimator of the mean does not coincide with the historical average, and the conclusion arrived at is similar to the historical average.

If we increase the duration for the past data to 10, 15 or 20 days prior to the estimation, it does not improve the accuracy of the prediction. The result is quite encouraging and is in the expected line. If the duration of the past data is increased, it will lead to more volatility in the prices and thereby decreasing the prediction accuracy. The prediction accuracy can be obtained from the variance of the posterior density of the mean $\theta = e^{\mu + \frac{1}{2}\sigma^2}$ and can be routinely computed. In this paper we have only confined to $100(1-\alpha)\%$ equitailed credible interval for the mean of the lognormal distribution, as it is computationally simpler than the HPD (highest posterior density) credible interval.

We suggest that Bayes prediction be used for the decision making along with other financial indicators. This type of Bayesian forecasting is quite useful for inter day investors. For long term investors a suitable Bayesian model based on Bayesian autoregressive procedure can be used. We have not compared the performance of the rule proposed in this paper with the technical forecasting approach used by Brock et al. [5], as it requires a long historical data to estimate the returns, and Granger [15] does not advocate this type of comparison. Our procedure is based on short term data, and thus the accuracy of our approach cannot be compared with the other model based Bayesian procedures or the procedures based on artificial intelligence, as the latter procedure requires a long term data. For the analysis, a programme was written using the software MATLAB version 7.0, and can be obtained from the first author.

Acknowledgement

The authors are thankful to the anonymous referees whose critical comments have resulted in the substantial improvement in the quality of the paper. The first author would like to take the opportunity to thank the

Government of India, Ministry of Science and Technology, Department of Science and Technology, New Delhi, for sponsoring her with the award of INSPIRE fellowship, which enables her to carry out the research programme which she has undertaken. She is much honoured to be the recipient of this award.

References

[1] E. K. Al-Hussaini, Predicting observables from a general class of distributions, J. Stat. Plann. Infer. 79 (1999), 79-91.

[2] R. Ananthakrishnan and M. K. Soman, Statistical distribution of daily rainfall and its association with the coefficient of variation of rainfall series, Int. J. Climatology 9 (1989), 485-500.

[3] I. Antoniou, Vi. V. Ivanov, Va. V. Ivanov and P. V. Zrelov, On the log-normal distribution of stock market data, Physica A: Statistical Mechanics and its Applications 331 (2004), 617-638.

[4] J. O. Berger, Statistical Decision Theory and Bayesian Analysis, 2nd ed., Springer-Verlag, New York, 1985.

[5] W. Brock, J. Lakonishok and B. LeBaron, Simple technical trading rules and the stochastic properties of stock returns, Journal of Finance 47(5) (1992), 1731-1764.

[6] S. J. Brown and J. B. Warner, Using daily stock returns: the case of event studies, Journal of Financial Economics 14 (1985), 3-31.

[7] J. Y. Campbell and S. B. Thompson, Predicting excess stock returns out of sample: can anything beat the historical average?, The Review of Financial Studies 2(4) (2008), 1509-1531.

[8] P. Chang and C. Liu, A TSK type fuzzy rule based system for stock prediction, Expert System with Applications 34 (2008), 135-144.

[9] H. Chu, T. Chen, C. Cheng and C. Huang, Fuzzy dual-factor time series for stock index forecasting, Expert System with Applications 36 (2009), 165-171.

[10] Z. Ding, C. W. J. Grange and R. F. Engle, A long memory property of stock returns and a new model, Journal of Empirical Finance 1 (1993), 83-106.

[11] A. L. Garcia-Almanza and E. P. Tsang, Forecasting stock prices using genetic programming and chance discovery, 12th International Conference on Computing in Economics and Finance, 2006.

[12] J. Geweke and C. Whiteman, Bayesian Forecasting, Handbook of Economic Forecasting, 1, Elsevier, 2006.

[13] J. K. Ghosh, M. Delampady and T. Samanta, An Introduction to Bayesian Analysis: Theory and Methods, Springer Texts in Statistics, USA, 2006.

[14] A. Goyal and I. Welch, A comprehensive look at the empirical performance of equity premium prediction, The Review of Financial Studies 21(4) (2008), 1455-1508.

[15] C. W. J. Granger, Forecasting stock market prices: lessons for forecasters, International Journal of Forecasting 8 (1992), 3-13.

[16] P. J. Harrison and C. F. Stevens, A Bayesian approach to short term forecasting, Oper. Res. Soc. 22(4) (1971), 341-362.

[17] P. J. Harrison and C. F. Stevens, Bayesian forecasting, Journal of the Royal Statistical Society, Series B (Methodological) 38(3) (1976), 205-247.

[18] J. Harrison and M. West, Practical Bayesian forecasting, The Statisticians 36 (1987), 115-125.

[19] M. R. Hassan and B. Nath, Stock Market Forecasting using Hidden Markov Model: A New Approach, Published in the Proceedings of 5th International Conference on Intelligent System Design and Application (ISDA'05), 2005, pp. 192-196.

[20] M. R. Hassan, B. Nath and M. Kirley, A fusion model of HMM, ANN and GA for stock market forecasting, Expert System with Applications 33 (2007), 171-180.

[21] T. Hsieh, H. Hsiao and W. Yeh, Forecasting stock markets using wavelet transforms and recurrent neural networks: an integrated system based on artificial bee colony algorithm, Appl. Soft. Comput. 11 (2011), 2510-2525.

[22] J. D. Kalbfleisch and R. L. Prentice, The Statistical Analysis of Failure Time Data, 2nd ed., John Wiley & Sons, New York, 2002.

[23] B. K. Kale, A First Course on Parametric Inference, 2nd ed., Narosa, 1999.

[24] G. Kapetanios, V. Labhard and S. Price, Forecasting using Bayesian and information theoretic model averaging: an application to UK inflation, Journal of Business and Economics Statistics 26(1) (2008), 33-41.

[25] D. Kundu and H. Howlader, Bayesian inference and prediction of the inverse Weibull distribution for type-II censored data, Comput. Stat. Data Anal. 54 (2010), 1547-1558.

[26] J. F. Lawless, Statistical Models and Methods for Lifetime Data, 2nd ed., John Wiley & Sons, New York, 2003.

[27] R. B. Litterman, Forecasting with Bayesian vector autoregression: five years of experience, Journal of Business and Economic Statistics 4(1) (1986), 79-91.

[28] K. J. Oh and K. J. Kim, Analyzing stock market tick data using piecewise non-linear model, Expert Systems with Applications 22 (2002), 249-255.

[29] P. A. Thompson and R. B. Miller, Sampling the future: a Bayesian approach to forecasting from univariate time series models, American Statistical Association 4(4) (1986), 427-436.

[30] I. Virtanen and P. Yli-Olli, Forecasting stock market prices in a thin market, OMEGA Int. J. Mgmt. Sci. 15(2) (1987), 145-155.

[31] J. H. Wright, Bayesian model averaging and exchange rate forecasts, J. Econometrics 146 (2008), 329-341.