Stochastic differential equation (SDE) model of opening gold share price of bursa saham malaysia

F N Hussin, H A Rahman and A Bahar

Department of Mathematical Sciences, Faculty of Science, Universiti Teknologi Malaysia, 81310, Skudai, Malaysia

E-mail: halizarahman@utm.my

Abstract: Black and Scholes option pricing model is one of the most recognized stochastic differential equation model in mathematical finance. Two parameter estimation methods have been utilized for the Geometric Brownian model (GBM); historical and discrete method. The historical method is a statistical method which uses the property of independence and normality logarithmic return, giving out the simplest parameter estimation. Meanwhile, discrete method considers the function of density of transition from the process of diffusion normal log which has been derived from maximum likelihood method. These two methods are used to find the parameter estimates samples of Malaysians Gold Share Price data such as: Financial Times and Stock Exchange (FTSE) Bursa Malaysia Emas, and Financial Times and Stock Exchange (FTSE) Bursa Malaysia Emas Shariah. Modelling of gold share price is essential since fluctuation of gold affects worldwide economy nowadays, including Malaysia. It is found that discrete method gives the best parameter estimates than historical method due to the smallest Root Mean Square Error (RMSE) value.

1. Introduction

Gold share price plays an important role in the economy. The fluctuation of gold can affect the economy anytime, not only in Malaysia, but also worldwide. This is because the supply and demand of the domestic currency will be effected when central banks purchase gold, which may also result in inflation. The changes occurred in other financial assets does not affect the gold demand while gold demands kept increasing despite the financial crisis that occurred [1]. If the change of price of gold has certain randomness to them, then one may use Brownian motion (BM) to describe these price changes [2].

Black-Scholes follows a stochastic process called Geometric Brownian motion, under which there will be including a random draw from the standardized normal distribution [3].

Geometric Brownian motion process assumes only positive value, just as what will be discovered in reality. Furthermore, the expected returns of GBM are independent of the value of the process, which also agrees with reality and it shows same kind of ‘roughness’ in its paths as can be seen in real stock prices [4].
2. Methodology

2.1 Distribution fitting of the observed data
Prior to model the observed data with Black-Scholes model given by,

\[ dX_t = \mu + \sigma X_t dB_t \]  \hspace{1cm} (1)

where \( \mu \) is the average drift term, \( \sigma \) is the diffusion term and \( dB_t \) is the Brownian noise. A standard Brownian motion (SBM) \( \{B(t) : t > 0\} \) is a stochastic process having;

a) Continuous paths, that is \( B(0) = 0 \)

b) Stationary and independent increments

c) \( B(t) \sim N(0,t) \) for all \( t \geq 0 \).

The stock price must satisfy the assumptions of having a log-normal random walk also known as the geometric random walk. Thus, distribution fitting procedure via Kolmogorov-Smirnov, Anderson-Darling and Chi-squared testings will be performed for lognormal distribution using Easyfit Software.

2.2 Parameter estimation methods
The parameter estimation methods are historical methods and discrete methods based from Khaled (2010) [5].

(i) The historical method
Historical method is a statistical method which uses the property of independence and normality logarithmic returns. Volatility of a stock is defined to be the standard deviation of log-returns of the stock. The log-returns are the logarithms of the ratio of successive prices. The Black-Scholes model assumes a constant volatility, and one way to estimate this is to use historical volatility as an estimator. If we have price data from \( n + 1 \) periods (in our case days), then the estimate for historical volatility is given by

\[ \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} (r_i - \hat{r})^2 = s^2 \]

One has \( R_j = log \left( \frac{x_j}{x_{j-1}} \right), j \geq 1 \). A very important constraint of the Black-Scholes model is that the returns \( R_j \) are independent and have the same Gaussian law. Note that,

\[ m = \mu - \frac{1}{2} \sigma^2 \]

If there are \( n \) observations \( (r_1, ..., r_n) \) of the returns \( R_i \), the density of this sample, attachment likelihood is given by the product of the Gaussians densities. The estimates of \( m \) and \( s \) are given by:

\[ \hat{m} = \hat{\mu} - \frac{1}{2} \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} r_i = \hat{r} \]

and therefore the estimated parameters are:

\[ \hat{\mu} = r + \frac{1}{2} \hat{\sigma}^2 \]

\[ \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} (r_i - \hat{m})^2 = s^2 \]
(ii) **Discrete method**

\( X(t) \) is a process of diffusion log normal characterized by the function of density of following transition, for \( s < t, X(s) = y, \)

\[
 f\left(x, \frac{t}{x_0}, t_0\right) = \frac{1}{\sigma \sqrt{2\pi}} \exp\left\{ \frac{-1}{2\sigma^2(t-t_0)} \left[ \log x - \log x_0 - \left( \mu - \frac{1}{2}\sigma^2 \right)(t-t_0) \right] \right\}.
\]

(2)

Considering \( b \) independent paths of the process \( X, (i = 1, ..., b) \) and by applying the method of maximum likelihood from the density, the parameter estimated can be obtain as follows. The likelihood function of (2) is given by,

\[
 L\left(x, \frac{t}{x_0}, t_0\right) = \prod_{i=1}^{b} \frac{1}{\sigma \sqrt{2\pi}} \left( \frac{1}{x_i} \right) \exp\left\{ -\frac{1}{2\sigma^2(t_i-t_0)} \left[ \sum_{n} \left( \log x_i - \log x_0 - \left( \mu - \frac{1}{2}\sigma^2 \right)(t_i-t_0) \right) \right] \right\}.
\]

Differentiating the log-likelihood function with respect to the wanted parameter (for this case \( \mu \)), resulting in

\[
 \frac{dl}{d\mu}\left(x, \frac{t}{x_0}, t_0\right) = -\frac{2\sum \log \left( \frac{x_i}{x_0} \right)(t_i-t_0) + 2\left( \mu - \frac{\sigma^2}{2} \right)(t_i-t_0)^2}{[2\sigma^2(t_i-t_0)]^2}.
\]

Solving the maximum of the likelihood function,

\[
 \frac{dl}{d\mu}\left(x, \frac{t}{x_0}, t_0\right) = 0
\]

thus,

\[
 \hat{\mu} = \frac{\log \left( \frac{x_n}{x_0} \right)}{t_n-t_0} + \frac{1}{2} \sigma^2
\]

and with the same procedure,

\[
 \hat{\sigma}^2 = \frac{\sum_{j=1}^{k} \left( \frac{x_j}{x_{j-1}} \right) \left( \frac{x_{j-1}}{x_j} \right) - \frac{\log \left( x_{j-1} \right)}{t_j-t_{j-1}} - \frac{\log \left( x_n \right)}{t_n-t_0} \right)^2}{k}.
\]

### 2.3 Goodness of fit test of the SDE model

If a stochastic differential model been developed for certain stochastic processes the goodness of fit of a statistical model describes how well it fits a set of observations. Measures of goodness of fit typically summarize the discrepancy between observed values and the values expected under the model in question. Suppose that stochastic process is observed at times \( t_0, t_1, ..., t_{n-1} \) where \( t_i = i\Delta t \) for constant \( \Delta t > 0 \). Let denote the \( n \) observations of the process.

Based from Allen [6] whereby in this procedure, \( M \) simulations of (2) are calculated from time \( t_{i-1} \) until time \( t_i \) starting at \( x_{i-1} \). Let \( X_{j,m}^{(n)} = X_{j,m}^{(n)} \) be the \( m \)th simulated value at \( m = 1, 2, ..., M \) and for \( i=1,2, ..., N-1 \). Now define
\( s_{i}^{(m)} = \begin{cases} 1, & \text{if } x_i \geq X_i^{(m)} \\ 0, & \text{if } x_i < X_i^{(m)} \end{cases} \)

and let \( r_i = 1 + \sum_{m=1}^{M} s_{i}^{(m)} \) for \( i = 1, 2, \ldots, n-1 \).

Then, \( r_i \) is the rank of value \( x_i \) as compared with the \( M \) simulated values, \( X_i^{(m)}, 1 \leq m \leq M \), for \( i = 1, 2, \ldots, n-1 \). The null hypothesis is that the model (2) describes the stochastic process. Under the null hypothesis, the ranks \( r_i \) have equally likely value between 1 and \( M+1 \). A \( \chi^2 \) goodness-of-fit test is used to test this hypothesis. To perform this test, the observed and expected frequencies are needed. Let,

\[
I_{i,q} = \begin{cases} 1, & \text{if } r_i = q \\ 0, & \text{if } r_i \neq q \end{cases}
\]

for \( i = 1, 2, \ldots, n-1 \) and let \( \Omega(q) = \sum_{i=1}^{n-1} I_{i,q} \) for \( q = 1, 2, \ldots, M+1 \). The expected frequency under the null hypothesis is \( \frac{n-1}{M+1} \). The test statistics is

\[
Q_m = \sum_{q=1}^{M+1} \left( \frac{\Omega(q) - \frac{n-1}{M+1}}{\frac{n-1}{M+1}} \right)^2
\]

under the null hypothesis is approximately distributed as a chi-square random variables with \( M \) degrees of freedom. If \( P(\chi^2(M) \geq Q_m) \) is smaller than a preset level of significance, then the null hypothesis is rejected indicating a lack-of-fit of the stochastic differential equation with the data.

3. Result and analysis

In order to model with Black-Scholes option pricing model, the gold stock prices were assumed to follow lognormal distribution; hence this study aimed to test whether the data of gold share price indeed follows the distribution, or the opposite. The distribution of FBM Emas is as follows:

![Figure 1. The graph of distribution of FTSE Bursa Malaysia Emas.](image)

Depicted by figure 1, it is found that the data of FTSE Bursa Malaysia Emas is nearly followed the lognormal distribution, as drawn in the graph as blue line. The graph of distribution of FTSE Bursa Malaysia Emas Shariah is:
The graph above shows the distribution of FTSE Bursa Malaysia Emas Syariah in accordance to lognormal distribution in the yellow line. For further investigation on whether to assume the data follows lognormal distribution or not, the goodness of fit test was run on the next section.

Hypothesis:
- \( H_0 \): The data follows lognormal distribution
- \( H_1 \): The data does not follow lognormal distribution

| Goodness of fit test       | Significance level | Critical value | Test statistic | Conclusion |
|----------------------------|--------------------|----------------|---------------|------------|
| Kolmogorov-Smirnov         | 0.05               | 0.11163        | 0.0451        | Accept \( H_0 \) |
| Anderson-Darling           | 0.05               | 2.5018         | 0.3120        | Accept \( H_0 \) |
| Chi-Squared                | 0.05               | 14.067         | 2.7831        | Accept \( H_0 \) |

From table 1 and table 2 it can be concluded that both data follows lognormal distribution as stated in previous study. It is crucial since it has been mentioned that the data of gold stock can be modeled by Geometric Brownian Motion, which lognormally distributed.

3.1 Parameter estimation
For this method, the parameter estimates for for both data are also not far from each other. However, the main aim for the study was to compare which method gives the best parameter estimate therefore, the table of all estimated parameters can simplified as shown below:
Table 3. Summary of the estimated parameters using historical and discrete method for both data.

| Estimated parameters | Methods       | FTSE Bursa Malaysia Emas | FTSE Bursa Malaysia Emas Syariah |
|----------------------|---------------|--------------------------|----------------------------------|
| $\hat{\mu}$          | Historical    | 0.000206192              | 0.000300741                      |
|                      | Discrete      | 0.000207572              | 0.00022082                       |
| $\hat{\sigma}^2$     | Historical    | 0.0000645157             | 0.000067691                      |
|                      | Discrete      | 0.00006451290            | 0.0000676895                     |

To determine which method gives the better estimates to both FTSE Bursa Malaysia Emas and FTSE Bursa Malaysia Emas Malaysia, root mean square errors (RMSE) was applied and the lowest RMSE between two methods will be chosen to be the best method. The formulae of calculating RMSE is

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{148} (x_1(i) - x_2(i))^2}$$

where $x_1$ is the observed value and $x_2$ is the predicted value.

Table 4. Root Mean Square Error (RMSE) for both methods.

| Type of stock                   | RMSE          | Historical Method | Discrete Method |
|---------------------------------|---------------|-------------------|-----------------|
| FTSE Bursa Emas                 | 231.2073685   | 230.556283        |
| FTSE Bursa Emas Syariah         | 2576.306      | 253.1843989       |

From table 4, it can be seen that Discrete method gives the lowest RMSE. Therefore, it can be concluded that Discrete method gives the better estimate of the share data.

4. Goodness of fit test of the SDE model

Performing 28 simulations, and applying goodness of fit test to FTSE Bursa Malaysia Emas data, $Q_{test}$ is calculated as $Q_{test} = 42.1905$. Since the parameters estimated are two, hence $M - 2 = 26$. The probability of having $\chi^2_{26}$ with degrees of freedom 26, are as listed in the table below.

Table 5. Significance levels and decision.

| $\alpha$ | 0.10  | 0.05  | 0.025 | 0.010 | 0.005 |
|----------|-------|-------|-------|-------|-------|
| $\chi^2_{26}$ | 35.563 | 38.885 | 41.923 | 45.642 | 48.290 |
| **Decision** | Reject | Reject | Accept | Accept | Accept |

Figure 3. Graph of comparison between actual value and expected value for RMSE Bursa Malaysia Emas (Simulated value).
5. Conclusion
After running the test on two samples of gold share price, FBM Emas and FBM Emas Syariah the results show that discrete method gives the best estimate. Discrete method gave the lowest RMSE compared to historical method for both data. Therefore, it can be concluded that discrete method is more accurate in estimating the parameters of the Black-Scholes model.

The Black-Scholes model does an adequate job of generally fitting the FBM Emas data. The model shows the same kind of “roughness” in its paths as can be seen in the real gold stock price. The goodness of fit test also proved that at certain confidence intervals, the model is significant to the actual data.

However, for FBM Emas Syariah, the model does not fit for all values. This means the model is not adequate for the data. At this state, there are many possible factors that needed to be considered such as for this set of data, there is a possibility of the volatility changes over time but in GBM volatility is assumed to constant or possibly the returns are not normally distributed, which means there is a higher chance of large price changes.

References
[1] Shahriar S and Erkan T 2010 An overview of global gold market and gold price forecasting Resources Policy 35(3) 178-189
[2] Radomski P 2013 http://beforeitnews.com/gold-and-precious-metals/2013/10/gold-silver-and-mining-stocks-decline-as-stocks-leads-the-way-20-2542550.html
[3] Brewer K D, Feng Y and Kwan C C 2012 Geometric brownian motion, option pricing, and simulations Some Spreadsheet-Based Exercises in Financial Modeling 38
[4] Zhong L 2010 Betting on Volatility: A Delta Hedging Approach 53
[5] Khaldi K M S 2010 Black-Scholes Model Share Price of Gold Estimation of the Parameters of the Stochastic Differential Equations 4
[6] Allen E 2007 Modelling With Ito Stochastic Differential Equations (The Netherlands: Springer)