Identification of Jump Times of Large Jumps for the Nikkei 225 Stock Index from Daily Share Prices via a Stochastic Volatility Model

Saneyuki ISHIDA* and Shuya KANAGAWA**

* Graduate School of Systems Information Engineering, Tokyo City University, Tokyo, Japan
** Department of Mathematics, Tokyo City University, Tokyo, Japan

Abstract

We investigate daily share prices of the Nikkei 225 stock index to identify jump times of the stock index. Since such daily share prices are observed at discrete times, it is difficult to find real jump times. In this paper we consider how to separate jump times from the observed times.

MSC2010 Subject Classification. 91G70, 60J75, 65C30.

Key words: Stochastic Differential Equation, Compound Poisson Process, Lévy Process, Jump Diffusion, Black-Scholes Model.

1 Introduction

Bell and Torous (1983, 1985) ([1], [2]) investigated share prices to show that there exist some jumps in the behavior of share prices using the Merton model. After their works a lot of people have studied stochastic processes with jumps as a model of share prices. Recently the Gibbs sampling is used to estimate unknown parameters of the jump stochastic processes. (See e.g. Chan and Wong [3].) In general likelihood estimators are used to identify some parameters in jump stochastic models of share prices like jump diffusions. These methods treat small and frequent jumps rather than large jumps. (See e.g. Inoue and Ozaki [5] in the details.) Since such small and frequent jumps can not be separated easily from variations of the continuous part of the model, it is difficult to find jump times of share prices by these methods by estimating unknown parameters in jump stochastic models.

As a model of share prices some stochastic processes which involve jump terms and continuous terms are used, e.g. jump diffusion or the Lévy process are commonly used. Fig.1 shows daily share prices of the Nikkei 225 stock index in the period of February 27th ~
April 8th, 2015. Such real stock index data consist of only jumps without continuous parts. Therefore it is difficult to find the real jump times and distinguish precisely between jump parts and continuous parts. We investigate daily share prices of the Nikkei 225 stock index to identify the jump times using the following stochastic model

$$dX(t) = dS(\mu_t, s_t, t) + X(t) dZ(t), \quad 0 \leq t \leq T,$$

(1)

where $X(t)$ is the share price, $S(\mu_t, s_t, t)$ is a non-negative stochastic process with continuous paths and $Z(t)$ is a compound Poisson process with an intensity $\lambda > 0$ of Poisson distribution. A typical example of $S(\mu_t, s_t, t)$ is the Black-Scholes model defined by

$$dS(\mu_t, s_t, t) = \mu_t X(t) dt + \sigma_t X(t) dB(t), \quad 0 \leq t \leq T,$$

(2)

where $B(t)$ is a standard Brownian motion, $\mu_t$ is a trend parameter and $\sigma_t > 0$ is a stochastic volatility. In this paper we show an algorithm to separate the jumps of $Z(t)$ from daily share prices of the Nikkei 225 stock index without any correct model of the continuous process $S(\mu_t, \sigma_t, t)$.

As for estimation of unknown parameters of the jump diffusion (2) there are lots of papers (see e.g. [4], [8] and [9]), but they do not identify jump times. On the other hand Jorion [6], Kwakernaak [7] and Iino and Ozaki [5] et al. investigate large jumps as our results. For example Iino and Ozaki (1999) estimate unknown parameters of non-Gaussian time series for foreign exchange which involves large jumps with low frequency using AIC. However their results are strongly depend on the continuous part of their models. Since an error in the estimation of the continuous term has a huge effect on the jump term, it possibly leads to a significant error in the estimation of jumps. This is the reason why our model of share prices (1) does no: need a correct form of the continuous process $S(\mu_t, \sigma_t, t)$.

There are some stochastic models to investigate share prices, e.g. Lévy process, stable process, jump diffusion and so on. For example the Lévy process is defined by a characteristic
function (Lévy-Khinchin formula) and it can be separated into three terms in theory, i.e. a Brownian motion, a martingale with small jumps and a compound Poisson process with large jumps. Although we can use some likelihood estimators to estimate unknown parameters in the Lévy process, it is difficult to identify jump times from the real stock index data by the identification of the Lévy process.

In this paper we show that we can identify large jumps using only the compound Poisson process \( Z(t) \) and we do not need the proper model of the continuous process \( S(\mu_t, \sigma_t, t) \) for the identification. Therefore our model is robust as compared with other models. We have to notice that our algorithm in principle can not separate small size jumps with high frequency.

2 Compound Poisson Process

Let \( N(t) \) be a counting process which means a number of jumps until the time \( t > 0 \). The distribution of \( N(t) \) obeys the Poisson distribution such that

\[
P\{N(t) = k\} = e^{-\lambda t} \frac{(\lambda t)^k}{k!}, \quad k = 0, 1, 2, \ldots
\]

for some positive constant \( \lambda > 0 \) which is called "intensity". Let \( t_1, t_2, \ldots \) be jump times of \( N(t) \) and \( \tau_1, \tau_2, \ldots \) be differences of the jump times defined by

\[
\tau_k = t_k - t_{k-1}, \quad k = 1, 2, \ldots
\]

Define a compound Poisson process \( Z(t) \) by

\[
Z(t) = Y_1 + Y_2 + \cdots + Y_{N(t)},
\]

where \( Y_1, Y_2, \cdots \) are some i.i.d. random variables which mean jump sizes. It is well known that \( \tau_1, \tau_2, \cdots \) are i.i.d. random variables with the exponential distribution \( \text{Exp}(\lambda) \) such that

\[
P\{\tau_k \leq x\} = \int_0^x \lambda e^{-\lambda s} ds
\]

for each \( k = 1, 2, \cdots \). In the following sections we consider an algorithm to find the jump times \( t_1, t_2, \cdots \) by an estimator \( \chi_0^2 \) for the testing fitness to the Poisson distribution of samples of \( Z(1) \).

3 Estimation of \( \mu_t \) and \( \sigma_t > 0 \)

We estimate \( \mu_t \) and \( \sigma_t > 0 \) from the the Nikkei 225 stock index data. Let

\[
r(k) = \log \frac{S(k+1)}{S(k)}
\]
be the logarithmic return of the $k$th day and put
\[ s_t = \sqrt{\frac{1}{19} \sum_{k=1}^{20} (r(t-k) - \bar{r})^2}, \quad \bar{r} = \frac{1}{20} \sum_{k=1}^{20} r(t-k). \]
$s_t$ is the historical volatility estimating for $\sigma_t > 0$ using stock indexes of 20 days. Furthermore let
\[ \hat{\mu}(t) = \log \frac{S(t)}{S(t-20)} \]
be an estimator of $\mu_t$.

4 Estimation of the Intensity $\lambda$ of $Z(t)$

Since the counting process $N(t)$ obeys Poisson distribution
\[ P\{N(t) = k\} = e^{-\lambda t} \frac{(\lambda t)^k}{k!}, \quad k = 0, 1, 2, \ldots, \]
we estimate the intensity $\lambda$ by the following steps using the daily returns of the Nikkei 225 stock index for 20 years from 1994 to 2013. We define an unit period as one month (20 days). Therefore $N(1)$ means a number of jumps observed in a month. We first identify jump times from daily returns observed in each month in 20 years. Put a criterial parameter $\alpha > 0$. If the absolute value of a daily return $r(k)$ is larger than $\alpha r_{k-1}$, then we estimate that the return $r(k)$ jumps at $k$th day which is called the jump time.

1st step: Put a criterial parameter $\alpha > 0$ for jump size.

2nd step: If $|r(k)| \geq \alpha r_{k-1}$, then the $k$th day is a jump time in the $j$th month for $j = 1, 2, \ldots, 240$. Count a number of all jumps in the $j$th month. The number $y_k$ of jumps is a sample of $N(1)$ of the $j$th month, where $N(1)$ means a number of jumps in the $j$th month.

3rd step: For the 240 samples $m_1, m_2, \ldots, m_{240}$ of numbers of jumps in a month, put
\[ \chi^p_{\alpha} = \frac{1}{240} (m_1 + m_2 + \cdots + m_{240}), \]
which is an estimator of the intensity $\lambda$.

4th step: Let $n_k$ be a number of months of $k$ jumps in a month. From the definition of $n_k$
\[ n_0 + n_1 + \cdots + n_K = 240. \]
5th step: Calculate the following estimator $\chi^2_\alpha$ of the test of goodness of fit to Poisson distribution with the intensity $\lambda_\alpha^p$ for the samples $n_1, n_2, \cdots n_K$.

$$\chi^2_\alpha = \sum_{k=0}^{K} \frac{(n_k - 240p_k)^2}{240p_k},$$

where

$$p_k = e^{-\lambda_\alpha^p} \frac{(\lambda_\alpha^p)^k}{k!}, \quad k = 0, 1, 2, \cdots, K - 1 \quad p_K = 1 - \sum_{k=1}^{K} e^{-\lambda_\alpha^p} \frac{(\lambda_\alpha^p)^k}{k!}.$$

6th step: Comparing $\chi^2_\alpha$ for each $\alpha = 0.8, 1.0, 1.2, \cdots, 3.0$, find the applicable $\lambda_\alpha^p$.

5 Estimation of the Distribution of $Y_i$

From the 6 steps in the previous section, the applicable $\lambda_\alpha^p$ can be obtained. For the applicable $\alpha$, $r(k)$ is identified as the jump size at the $k$th day when $|r(k)| \geq \alpha s_{k-1}$. From all samples of returns $\{r(k)\}$ at jump times, we can estimate the distribution of $Y_i, i = 1, 2, \cdots$. Fig.2 shows a histogram of $Y_i, i = 1, 2, \cdots$ observed in 20 years.

![Figure 2: Histogram of $Y_1, Y_2, \cdots$](image)

6 Estimation of Jump Times

For each month (20 days) let $t_1^\alpha, t_2^\alpha, \cdots$ be jump times according to the applicable parameter $\alpha > 0$ estimated by 5 steps in Section 4. If the distribution of $N(1)$ obeys the Poisson distribution, then the intervals $\tau_k^\alpha, k = 1, 2, \cdots$ are i.i.d. random variables with the exponential distribution, i.e.

$$\tau_k^\alpha = \frac{t_k^\alpha - t_{k-1}^\alpha}{20}, k = 1, 2, \cdots, M \quad i.i.d. \sim Exp(\lambda_\alpha),$$
where

\[ M = n_1 + \cdots + n_K, \]

which means the number of jumps. Since \( \lambda_\alpha \) is the expectation of intervals of jump times \( \tau^\alpha_k, k = 1, 2, \cdots, \) we estimate \( \lambda_\alpha \) by the sample mean \( \lambda^e_\alpha \) of length of all intervals of jump times. From the definition of the compound Poisson process \( \int_0^t Z(s) ds \), it is well known that the parameter \( \lambda \) of the exponential distribution is the same as \( \lambda \) of the Poisson distribution. We can confirm the result from Tables 1 and 2.

Divide one month (20 days) into 5 periods. Each period contains 4 days. Let \( q_k \) be

\[ q_k = \int_{0.2(k-1)}^{0.2k} \lambda_\alpha e^{-\lambda_\alpha x} dx, \quad k = 1, \cdots, 5. \]

\( q_k \) is the theoretical probability of \( \tau^\alpha \) satisfying \( 0.2(k-1) \leq \tau^\alpha \leq 0.2k \).

Let \( b_k, k = 1, 2, \cdots, 5 \) be the number of \( \tau^\alpha \) which are contained in each interval \( 0.2(k-1) \leq \tau^\alpha \leq 0.2k \). Let \( \eta^2_\alpha \) be the test statistics of goodness of fit with respect to \( \{ \tau^\alpha_k \} \) defined by

\[ \eta^2_\alpha = \sum_{k=1}^{5} \frac{(b_k - Mq_k)^2}{Mq_k}. \]

We show \( \chi^2_\alpha, \eta^2_\alpha, \lambda^p_\alpha \) and \( \lambda^c_\alpha \) obtained from the share prices of 20 years in Tables 1 and 2.

**Table 1:** \( 1.0 \leq \alpha \leq 1.4, \chi^2_\alpha, \eta^2_\alpha, \lambda^p_\alpha, \lambda^c_\alpha \)

<table>
<thead>
<tr>
<th>( \alpha )</th>
<th>1.0</th>
<th>1.1</th>
<th>1.2</th>
<th>1.3</th>
<th>1.4</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \chi^2_\alpha )</td>
<td>13.923</td>
<td>19.185</td>
<td>14.046</td>
<td>18.162</td>
<td>4.4002</td>
</tr>
<tr>
<td>( \eta^2_\alpha )</td>
<td>52.738</td>
<td>32.119</td>
<td>23.306</td>
<td>21.942</td>
<td>0.22005</td>
</tr>
<tr>
<td>( \lambda^p_\alpha )</td>
<td>6.6333</td>
<td>5.7417</td>
<td>4.9308</td>
<td>4.2583</td>
<td>3.5708</td>
</tr>
<tr>
<td>( \lambda^c_\alpha )</td>
<td>6.4821</td>
<td>5.6050</td>
<td>4.7940</td>
<td>4.1603</td>
<td>3.5157</td>
</tr>
</tbody>
</table>

**Table 2:** \( 1.5 \leq \alpha \leq 2.0, \chi^2_\alpha, \eta^2_\alpha, \lambda^p_\alpha, \lambda^c_\alpha \)

<table>
<thead>
<tr>
<th>( \alpha )</th>
<th>1.5</th>
<th>1.6</th>
<th>1.7</th>
<th>1.8</th>
<th>1.9</th>
<th>2.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \chi^2_\alpha )</td>
<td>3.8461</td>
<td>5.7029</td>
<td>7.8972</td>
<td>4.2192</td>
<td>12.4926</td>
<td>6.1155</td>
</tr>
<tr>
<td>( \eta^2_\alpha )</td>
<td>0.6294</td>
<td>0.47035</td>
<td>0.4668</td>
<td>0.5406</td>
<td>0.7763</td>
<td>0.8343</td>
</tr>
<tr>
<td>( \lambda^p_\alpha )</td>
<td>3</td>
<td>2.5625</td>
<td>2.1875</td>
<td>1.8333</td>
<td>1.5917</td>
<td>1.3125</td>
</tr>
<tr>
<td>( \lambda^c_\alpha )</td>
<td>2.9382</td>
<td>2.5031</td>
<td>2.1424</td>
<td>1.8498</td>
<td>1.5620</td>
<td>1.2881</td>
</tr>
</tbody>
</table>
7 Conclusion

From Tables 1 and 2 $\chi^2_\alpha = 18.162$ at $\alpha = 1.3$ and $\chi^2_\alpha = 4.04$ at $\alpha = 1.4$, respectively. On the other hand $\eta^2_\alpha = 21.942$ at $\alpha = 1.3$ and $\eta^2_\alpha = 0.22605$ at $\alpha = 1.4$, respectively. $\eta^2_{1.4} = 0.22605$ takes the minimum value. Combining these results we conclude that we can find jumps whose absolute values of returns are more than $\alpha \sigma_t$ and their jump times of the Nikkei 225 stock index by choosing $\alpha \geq 1.4$. The fact $\chi^2_\alpha = 12.4929$ when $\alpha = 1.9$ is thought to be due to that the goodness of fit for the Poisson distribution of jump sizes rapidly decreases since there are very few jumps for $\alpha \geq 1.9$. Fig.3 shows large jumps more than $1.4\sigma_t$ of the Nikkei 225 stock index in the period of 2011/1/27 ~ 2011/3/3.

![Figure 3: Jump times for $\alpha = 1.4$ in 2011/1/27 - 2011/3/3](image)

Acknowledgements. The authors would like to express their thanks to referees for their valuable comments. The research of the second author is supported in part by JSPS KAKENHI Grant Number 25400211, Ministry of Education, Science and Culture, Japan.

REFERENCES


