Proceedings of the 5th WSEAS Int. Conf. on SIMULATION, MODELING AND OPTIMIZATION, Corfu, Greece, August 17-19, 2005 (pp488-492)

Realized volatility estimation: new simulation approach and empirical study results

JULIA BONDARENKO, NIKOS MASTORAKIS
Institute of Applied System Analysis
National Technical University of Ukraine "Kyiv Politechnic Institute" (NTUU “KPI”)
Peremogy Prospect 37
03056 Kyiv
UKRAINE

World Scientific and Engineering Academy and Society (WSEAS)
Ag. Ioannou Theologou 17-23
15773 Zographou, Athens
GREECE
http://www.wseas.com/mastorakis/index.html

Abstract: - The paper empirically investigates several daily volatility estimators for the DAX index. Realized volatility is computed by means of standard, Parkinson, Garman-Klass estimators, which use the daily data samples, and also Andersen estimator based on the intraday information observed over time intervals of different sizes. A Monte Carlo simulation is conducted for two cases of underlying security fluctuation – the diffusion process and the process based on the “telegrapher” process; the theoretical results are compared with volatility values obtained from the studied estimators.

Key-Words: - realized volatility, Monte Carlo simulation, range estimators, geometric Brownian motion, “telegrapher” (Kac) process, estimator’s efficiency.

1 Introduction

In general, volatility is defined as fluctuations in the value of any financial security or in a portfolio of securities, and considered as a measure of market risk. The traditional method for modeling of underlying security prices assumes a “random walk” described by a geometric Brownian motion process [1]. Given the current price $S(0)$ of an underlying, the future underlying price $S(t)$ follows the stochastic differential equation:

$$d\,S(t) = \mu\,S(t)dt + \sigma\,S(t)dW(t), \qquad (1)$$

where $W(t)$ is a standard Brownian motion, μ represents expected rate of change or “drift rate” of the process, and σ represents volatility, $0 \le t \le T$. Using equation (1), the price of a European call $V(S,t)$ with a strike price K and maturity date T can be found by means the Black-Scholes formula [2] as follows

$$V(S,t) = S\;\Phi(d_1) - K\;e^{-r\,(T-t)}(d_2)\;, \qquad (2)$$

$$d_1 = \frac{\ln(S/K) + \left(r + \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}},\; d_2 = \frac{\ln(S/K) + \left(r - \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}},$$

where r is a risk free rate, $\Phi(x)$ is the cumulative normal probability for a standard normal random variable.

The value of σ can be technically expressed in several ways. In the valuation of options, the meanings of implied volatility and historical volatility are the most used. Implied volatility is estimated from traded option prices. Putting the current market price of an option $\tilde{V}$ into (2), we can compute the value of implied volatility as a unique solution of the equation $\tilde{V} = V(\sigma)$.

Historical (realized) volatility describes volatility observed in a security over a given period of time. Price movements in the security (historical data) are recorded at fixed time intervals over a given period. As before, it is assumed that prices are lognormally distributed. Using recent historical data provides better information on the current level of volatility. There is much discussion over the best method of calculating the historic volatility. The most usual and traditional measure is a standard deviation of the log-differenced close prices of asset. Other measures, such as Parkinson's extreme value estimator [3], Garman and Klass range-based estimator [4], Andersen's intraday estimator [5], improve the efficiency of realized volatility measures by using information embedded in daily high, low, open prices and high-frequency intra-day data.

The remainder of the paper is organized as follows. In Section 2, we start off with the features of data. Further, the section presents the realized volatility estimators, which will be employed, and also describes a Monte Carlo simulation technique for two different types of underlying process. Section 3 presents the simulation. Section 4 reports the empirical results. Conclusions are drawn and suggestions for future research offered in section 4.

2 Problem Formulation

2.1. Data

The data set we have used consists of daily prices for DAX index from January 8, 1996 to January 8, 1997. There are all together 251 trading days. The average daily return for this period is 0.000895224. Intraday returns are obtained by sampling from the initial grid of one-minute prices on January 8, 1997, from 8:30 a.m. to 17:05 p.m. GMT+1, and also calculated at 5 minute, 15 minute and 30 minute return period. To annualize the realized volatility for any given day, we have to multiply it by the square root of the number of trading days in a year.

2.2. Realized volatility estimators

The classical volatility estimator is defined as a standard deviation of the daily close to close returns:

$$\sigma_{cc} = \sqrt{\frac{1}{n-1}\sum_{t=1}^{n}(r_t - \bar{r})^2}\ , \tag{3}$$

where $r_t = \ln\frac{C_t}{C_{t-1}}$ is a price return, C_t is a close price on trading day t, $t = 1:n$, $\bar{r} = \frac{1}{n}\sum_{t=1}^{n} r_t$.

Parkinson estimator [3] is based on the daily log price range, which is defined as the difference between the daily high and low log-price, that is:

$$\sigma_p = k\sqrt{\frac{1}{n}\sum_{t=1}^{n} {p_t}^2}\ , \tag{4}$$

where $p_t = \ln\frac{H_t}{L_t}$, L_t and H_t are respectively the highest and lowest prices on day t, $t = 1:n$, $k = \frac{1}{2\sqrt{\ln 2}} \approx 0.601$.

It was proven that the range estimator of daily volatility (4) is approximately five times more efficient than the estimator based on squared daily close price returns (3).

Using also additional information embedded in daily open prices, Garman and Klass [4] have improved efficiency and suggested the following volatility estimator:

$$\sigma_{gk} = \sqrt{\frac{1}{n}\sum_{t=1}^{n}\left[\frac{1}{2}p_t^2 - (2\ln 2 - 1)q_t^2\right]}\ , \tag{5}$$

where $q_t = \ln\frac{C_t}{O_t}$, O_t is an open price on day t, $t = 1:n$. The estimator (5) is more than seven times more efficient than (3).

Andersen [5] has shown that the accuracy of volatility estimation can be also more improved by exploiting high frequency data. The proposed estimator ("integrated volatility") was constructed as the sum of squared intraday returns,

$$\sigma_{\mathrm{int},t} = \sqrt{\sum_{j=1}^{m} R_{t,j}^2}\ , \tag{6}$$

where $\sigma_{\mathrm{int},t}$ is a realized volatility on day t, $R_{t,j}^2$ is a squared intraday return, and m is a number of samples per day. To avoid a bias problem, it's reasonable to filtrate tick-by-tick price series and take a time interval of 5-15 minutes.

2.3. Monte Carlo simulation

The underlying asset price process is assumed to follow a one-dimensional diffusion process (1). Applying the Itô's Lemma, we can write the logarithm of the process as

$$d\ln S(t) = \left(\mu - \frac{\sigma^2}{2}\right)dt + \sigma\, dW(t). \tag{7}$$

A discrete approximation for (7) has a form

$$\begin{aligned}\ln S(t_i) - \ln S(t_{i-1}) = \frac{1}{M}\left(\mu - \frac{\sigma_i^2}{2}\right) + \\ + \sigma_i\left(W(t_i) - W(t_{i-1})\right),\end{aligned} \tag{8}$$

where M is a number of time intervals, $i = 1:M$, $W(t_i) - W(t_{i-1}) \sim \left(0, \dfrac{1}{M}\right)$. In the Monte Carlo procedure [6], we employ the equation (8), with $M = 1000$.

A Monte Carlo simulation with 600 produced paths gives us then a benchmark. The annual volatility is chosen to be a constant $\sigma_i = \sigma$, equal 16%, and close to the actual volatility for DAX index during the observed period. Using annual return value, the drift coefficient can be calculated via $\mu = 0.2238 + \dfrac{\sigma^2}{2} = 0.2366$.

2.4. "Telegrapher" process

This model of stochastic evolution was considered first in [7]. A described process can be viewed as following: a point is running on the real line with a constant velocity v; the point's motion is controlled by the Poisson process $N(t)$ with parameter λ. That means that the point starts to move from the origin in one and then changes instantaneously the direction at the moment of the Poisson event coming; moves during a random time interval until the next Poisson event, and so on. The related process

$$\xi(t) = v \int_0^t (-1)^{N(s)}\, dt \qquad (9)$$

giving the position of the point, is called "telegrapher" process, or Kac process. This mathematical model was studied also in [8, 9]. Furthermore, an application of the considered stochastic process modification in mathematical finance was suggested in [10]. So, there are discrete, random points in the time where the changes occur, what is close to the interpreting market price movements. Let's note also, that if a frequency of point's direction changes is high ($\lambda \to \infty$), $v \to \infty$, and $\dfrac{v^2}{\lambda} \to c = const$, $c > 0$, the process $\xi(t)$ asymptotically is a Brownian motion.

An approach based on the Monte Carlo simulation described in the previous section, is conducted for the process (9) as well. Discrete approximation of the corresponding process provides, analogously to (8), a formula

$$\ln S(t_i) - \ln S(t_{i-1}) = \frac{1}{M}\left(\mu - \frac{\sigma^2}{2}\right) + \\ + \sigma\left(\xi(t_i) - \xi(t_{i-1})\right), \qquad (10)$$

where as before $\sigma = 0.23$, $\mu = 0.25025$, and we ran a simulation procedure also 600 times. The parameters λ and v are taken to be equal 50 and 5 correspondingly. The further results comparison with other volatility estimators is carried out.

3 Empirical analysis

First of all, in Table 1 we present a descriptive statistics of the returns for historical data, both daily (close-close prices) and intraday (intervals 1, 5, 15, 30 and 60 minutes).

Table 1. Statistics of returns

Value	Daily	Intraday, with interval				
		1 min	5 min	15 min	30 min	60 min
Mean	8.95E-04	7.25E-06	3.63E-05	1.16E-04	2.32E-04	2.25E-04
St. dev.	8.03E-03	1.83E-04	4.04E-04	6.76E-04	1.01E-03	1.68E-03
Max	2.31E-02	1.03E-03	1.59E-03	1.62E-03	2.14E-03	2.48E-03
Min	-4.14E-02	-1.11E-03	-9.98E-04	-1.87E-03	-2.35E-03	-3.19E-03
Skew-ness	-9.08E-01	-4.04E-01	4.15E-01	-3.18E-01	-5.47E-01	-7.64E-01
Kurtosis	6.557	9.730	4.754	3.797	3.733	2.990

Mean returns of the index increase as interval duration becomes longer. The mean estimates for intraday data provide a reasonable proxy for daily data: the daily mean is approximately four times greater than the intraday means for intervals 30 and 60 min, eight times greater than the intraday mean for interval 15 min, twenty five times greater than the intraday mean for interval 5 min, and hundred twenty three times greater than the intraday mean for interval 1 min. There is the same tendency for different time intervals standard deviation: the estimate for 60 min interval is about 21% of the daily standard deviation, for intervals 30, 15, 5 and 1 min these magnitudes are 13%, 8.4%, 5% and 2.3% correspondingly. Maximal and minimal daily magnitudes are by an order greater than intraday ones. Returns, as a rule, are slightly negatively skewed, and for the case of 15 min interval are not far from normality. The intraday returns taken with 5 min interval have a positive skewness. Kurtosis values show, that returns in general are "fat-tailed", but for intraday data they decline as a length of time interval increases. For 15, 30 and 60 min the value of kurtosis is close to the normal case. The kurtosis for daily data is twice greater than the normal one.

Table 2 reports daily volatility estimates derived from classical, Parkinson and Garman-Klass estimators, and also from simulation results for the cases, when underlying security is modelled in frameworks of both standard Brownian motion and "telegrapher" (Kac) process. As a comparison criterion, we use an efficiency of each estimator. In the capacity of benchmark we take daily volatility obtained from the standard deviation approach (close-close prices estimator). Volatility values obtained from Monte Carlo simulation procedures for diffusion and Kac processes (1) and (9) are calculated as a mean from 600 realizations. The efficiency of arbitrary estimator is defined then by the ratio of the variance of known estimator to the variance of arbitrary estimator:

$$Eff\left(estimator\right)=\frac{Var\left(CC-Estimator\right)}{Var\left(estimator\right)}.$$

The magnitude Eff larger than 1 means the variance of the considered estimator is smaller than the variance of the benchmark one. Hence, the investigated estimator is more efficient.

Table 2. Daily volatility estimates. Simulation results and range estimators.

Value / Estimator	Daily volatility	Efficiency
MC-simulation (GBM)	0.00506201	2.51567428
MC-simulation (Kac process)	0.00451555	3.16139743
Classical	0.00802879	1
Parkinson	0.00328147	5.98636966
Garman-Klass	0.00309812	6.71589560

The efficiency of the benchmark estimator is supposed to be equal 1. Comparing with theoretical value based on Monte Carlo simulation for geometric Brownian motion, the efficiency is about $2{,}5$ times smaller. The higher efficiency (more than 3 times) is achieved by application of the "telegrapher" (Kac) process instead of the standard Brownian motion. Including high/low prices in estimator provides an estimator, which is almost 6 times more efficient (Parkinson estimator), and additional use of open/close prices – an estimator, which is $6{,}7$ times more efficient (Garman-Klass estimator), than the benchmark. In general, the range-based estimates are downward biased. It's shown also via the Monte Carlo study: value of Parkinson estimator is only 65% of Monte Carlo estimate for standard Brownian motion and 72% for Kac process. In the case of Garman-Klass estimator these values equal 61% and 69% correspondingly.

Volatility estimates based on intraday information, are presented in Table 3. This approach definitely improves an efficiency of estimation relative to the classical one.

Table 3. Daily volatility estimates. Intraday prices based (Andersen) estimator.

Value / Interval	Daily volatility	Efficiency
1 minute	0.00414452	3.75276982
5 minutes	0.00407293	3.88585432
15 minutes	0.00388249	4.27641373
30 minutes	0.00403514	3.95897895
60 minutes	0.00419757	3.65851224

Squared intraday returns are biased upwards, and the bias value is larger when the data are sampled more frequently (1 and 5 min intervals). The highest efficiency is attained in the case of 15 min interval, it's about $4{,}3$ times larger than for the classical close-close price estimator. The 30 and 60 min intervals estimates still remain less biased because of smaller number of data, but have the greater variances, which make the efficiency vanishing. The estimate magnitude for 15 min interval makes 77% of Monte Carlo estimate for standard Brownian motion and 87% for "telegrapher" process.

4 Conclusion

In this study, we considered several empirical approaches to the realized volatility measure for the DAX index prices. The several estimation methods are employed – classical close-close price estimator, range estimators (Parkinson and Garman-Klass) and Andersen estimator. The empirical results show, that the range-based estimators are highly efficient, but are downward-biased, that was established through a Monte Carlo study. Our Monte Carlo results also demonstrate that by using the alternative "telegrapher" process instead of standard Brownian motion, we can improve estimation efficiency. It is concluded that using the highest available frequency of intraday data (1 and 5 minutes in our case) leads to upward-biased daily volatility estimates; effect of superior estimating takes place also for the 30 and 60 min intervals due to the larger variance values.

An obvious direction for future research would be a further theoretical and empirical study of the "telegrapher" process and its modifications, analysis and comparison with various types of estimators, and the development of volatility forecasting techniques on the basis of proposed process with taking into account the market microstructures for concrete financial securities.

References:

[1] P. A. Samuelson, Rational Theory of Warrant Pricing, *Industrial Management Review*, Vol. 6, 1965, pp. 13-31.
[2] Black F. and Scholes M.S. (1973): The pricing of options and corporate liabilities. Journal of Political Economy 81, pp. 637-659.
[3] M. Parkinson, The extreme value method for estimating the variance of the rate of return. *Journal of Business*, Vol. 53, 1980, pp. 61-68.
[4] M. Garman and M. Klass, On the estimation of security price volatilities from historical data, *Journal of Business*, Vol. 53, 1980, pp. 67-78.
[5] T. G. Andersen, Some reflections on analysis of high-frequency data, *Journal of Business & Economic Statistics*, Vol. 18, 2000, pp. 146-153.
[6] P. Glasserman, Monte Carlo methods in financial engineering, Springer Verlag, 2003, 596 p.
[7] M. Kac, A Stochastic Model Related to the Telegrapher.s Equation. *Rocky Mountain Journal Math.*, Vol. 4, No. 3, 1974, pp. 497-509.
[8] E. Orsingher, A planar random motion governed by the two-dimensional telegraph equation, *Journal of Applied Probability*, Vol. 22, 1986, pp. 385-397.
[9] A. D. Kolesnik and A. F. Turbin, A.F., Symmetrical random evolutions in R^2, *Doklad. Akadem. Nauk Ukraine*, Vol. 2, 1990, pp. 11-14 (in russian).
[10] J. Bondarenko and N. Branger, An alternative model of financial assets prices dynamics, *accepted by Computing and Visualization in Science*, 2004.