Realized Range-based Variance と Realized varianceの比較 using Japanese data

西村　洋平

9. Dec. 2006

Summary　本稿は、最近新しく研究され始めた、Realized Range-based Variance(RR) を日本の TOPIX データで計算し、以前に研究されていた GARCH モデルや ARFIMA モデルに当てはめ、Realized Variance（RV) との比較が内容の中心となっている。GARCH モデルでは、GARCH(1,1) モデルに RR、RV を加え、尤度が上昇するかということを明らかにした。また、RR、RV の対数を取ったものを時系列モデルで分析した。それが ARFIMA モデルである。このモデルは long memory という特性を持っており、RR、RV の分析に適している。本稿では ARFIMA(0,d,0) モデルのパラメタを推定し分析、比較を行った。結論として、真の Volatility の推定値としては、RV よりも RR を用いたほうが真の値に近づくことが示された。

1　Introduction

Black F and Scholes M (1973) による、資本市場の一般均衡理論以降、ファイナンスの分野は様々な方向に発展した。しかし、彼らの論文の非常に重要な仮定である、Volatility の固定は現実にそぐわないとして、いかにして Volatility を変動させつつ、市場の分析を行うかということに注目が集まり、また、その方向に研究が発展していった。

そのような流れの中、ファイナンスの分野で注目、研究がなされている中心の一つとして、資産（株価指数、為替レート、債権など）価格変化率の 2 次モーメントをあらわす Volatility に焦点を絞って分析する研究がある。その研究としては主に 2 方向の発展がある。一つは、volatility が確率的に変動するという仮定の下、その対数値の変動を線形の ARMA モデルで記述するという Stochastic Voratility 変動モデルの分析がある。もう一方は、Engle R.F. (1982) の ARCH モデル以降、Bollerslev .T (1986) の GARCH モデル、さらにそれを拡張したモデルである。これらのモデルでは、Volatility を観測されない潜在変数と考え、日次リターンデータを用いてモデルのパラメタを推計をすることによって Volatility

の推定値や予測値を計算する。この研究では、どのモデルを使うのかで Volatility の推定値や予測値は異なる。これらは Volatility 変動モデルと呼ばれている。

しかし、Andersen, Torbin G. and Bollerslev.T (1998) で初めて、モデルに依存しない Volatility の推定量である Realized Variance(RV) が研究された。これは日中の 5 分毎のリターンの 2 乗を足し合わせたもので、資産価格の日中データが利用可能になるにつれて、研究が進められている[1]。

さらに、Christensen .K and Podolskij .M (2005) で考えられた Realized Range-based Variance(RR) という Volatility の推定量がある。これは、理論的に RV より 5 倍の有効性があることが分かっており、また実際的な応用性もあるということで注目すべきことである。本論文は、日本のデータ (TOPIX) による RR の計算、および、モデルによるパラメタ推定、それによる RV との比較が中心となっている。

本論文の構成は以下の通りである。まず第 2 節では、Realized Variance、Realized Range-based Variance の定義、その理論的特性を述べる。次に、第 3 節では、TOPIX のデータを用いて実際に RV、RR を計算し、基本統計量による分析を行う。さらに第 4 節では、GARCH モデル、ARFIMA モデルの枠組みを用いて RV と RR の比較を行う。最後に第 5 節でまとめと今後の研究の課題を挙げる。

2 Framework compared with RV

このセクションでは、RV、RR それぞれの計算方法と、理論的観点からの違い、RR の有効性について論じる。

まず、RV の計算方法は以下である。今、第 t 日の日中の n 個のリターンデータ $\{r_t, r_{t+1/n}, ..., r_{t+(n-1)/n}\}$ が与えられているものとする。このとき、それらを 2 乗して足し合わせた、

$$RV_t = \sum_{j=0}^{n-1} r^2_{t+j/m} \tag{1}$$

$$r^2_{t+j+1/m} = p_{j+1/m} - p_{j/m}$$

を RV と定義する。ここで $p_{j/m}$ とは、対数を取った、1 日の中の (m 個中)j 番目の価格である。

データを用いる際に注意すべきは、サンプリングの頻度である。上述したことから、サンプリング頻度を高めれば高めるほど真の Volatility に近づくが、市場のミクロ構造によるノイズが大きくなることが知られている[2]。よって、Andersen,

[1]例えば、Barndorff-Nielsen O. E. and Shephard N. (2002)、Koopman S. J. ,Jungbacker B. and Hol E. (2005)

[2]を参照

Torbin G. and Bollerslev.T (1998) にしたがって、5 分おきのリターンを計算し、そこから RV を導出した。

次に、RR の計算方法について述べる。RR は、

$$RR_t = \frac{1}{4\log 2}\sum_{i=1}^{I}(\log H_{t,i} - \log L_{t,i}) \tag{2}$$

のようにして表される。ここで、$\log H_{t,i}$ とは、一日のリターンデータを i 個ずつのブロックに分けたとき、そのブロック内で最も大きい価格の対数を取った値である。同様に $\log L_{t,i}$ はブロック内の最小の価格の対数値である。

ここで、RR を $\frac{1}{4\log 2}$ 倍するのは、以下の理由による。$\frac{1}{4\log 2}$ は、標準ブラウン運動の 2 次モーメントである。それはセミマルチンゲールのフレームワークで RR、RV、IV を定義するから、その分散を調整するために、RR の推定値にモーメントの逆乗をかけるのである。明確な値を求めるのは難しいので、シミュレーションで標準ブラウン運動の 2 次モーメントを求めてみると、漸近的に $4\log 2$ に近づいていくのがわかる。また、理論的には、連続時間で考えられており、実際のデータで、その 2 次モーメントを求めることは難しいが、データ数が大きければ理論的な 2 次モーメントを実証分析で用いても差し支えのないことが分かっている[3]。

以下、このようにして計算された RV、RR を基にして話を進める。[4]

RR が RV よりも理論上で優れている点として、

1. RR は全てのデータポイントの情報をもっているのに対して、RV はいくらかの情報を無視している

2. RV よりも 5 倍の有効性を持っているので、信頼区間が狭まる[5]

が挙げられる。

---

[3]Christensen .K and Podolskij .M (2005)

[4]上のようにして RV が計算される根拠としては次のような理由による。
資産価格の対数値 $\log P(s)$ が伊藤過程

$$d\log P(s) = \mu(s)ds + \sigma(s)dW(s)$$

にしたがっているとする (W(s) はウィナー過程である)。そうすると、第 t 日の真の Volatility は

$$IV_t = \int_t^{t+1} \sigma(s)^2 ds$$

と定義され、これは瞬間的な Volatility$\sigma(s)^2$ を積分したもの (Integrated volatility) となる。
(1) 式で定義される RV_t は、$n \to \infty$ とすると、IV_t に確率収束するので、n が十分大きいなら RV_t は IV_{t} の精度の高い推定量となる。

[5]これは RR の推定値の分散が、RV の推定値の分散と比較して 1/5 しかないことによる。(Christensen .K and Podolskij .M (2005))

3 Empirical Application: TOPIX

このセクションでは、TOPIX の Tick Data を用いて RV、RR を導出、その特徴について述べる。

3.1 data

データ範囲は、柴田舞 (2006) にのっとり、2000 年 1 月 6 日 9 時 1 分から 2003 年 12 月 29 日 15 時までとした。しかし、この期間のうち、年末年始は午前中しか取引が行われないので、分析データからはずす。また、以下の日は RV が非常に高い日であり、それらを異常値として分析対象からはずした。

- 2000 年 4 月 17 日：前週末のアメリカ株式相場の影響を受け、東証一部上場銘柄のうち、1236 銘柄が値下がった。
- 2000 年 4 月 21 日、24 日：日経平均株価指数に採用されている銘柄のうち、30 銘柄の入れ替えに伴い、入れ替え対象銘柄を中心に売買が活発になった。
- 2001 年 9 月 12 日：アメリカで 9 月 11 日に発生した同時多発テロの影響による混乱が生じた。

これらを除いた総データ数は 972 である。

3.2 基本統計量による分析

まず、基本統計量を用いた分析をおこなう。RV、RR と、それぞれの対数を取ったものを以下の表に載せる。

[Table 1:Basic Statistics insert here]

RV と RR の相関係数は 0.69 であり、高い相関がある。Fig1,Fig2,Fig3 はそれぞれ、RV,RR のサンプルパス、密度、自己相関関数である。これから、$\log RR$、$\log RV$ は正規分布に非常に近い分布をしている。つまり、RR、RV は対数正規分布に従っていると考えられる。

4 モデルによる比較

4.1 Estimation of GARCH(1,1) モデル

RR、RV の比較をするために、本論文では基本統計量の分析だけではなく、Koopman S. J. ,Jungbacker B. and Hol E. (2005) にのっとり、以下の GARCH(1,1) モデルのパラメタを推定し、RR、RV の比較を行う。

$$
\begin{aligned}
r_t &= \mu_0 + \mu_1 r_{t-1} + \epsilon_t \\
\epsilon_t &= \sigma_t z_t \\
z_t &\sim N(0,1), i, i, d \\
\sigma_t^2 &= \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2 + \gamma RR_{t-1}
\end{aligned}
$$

これを、最尤法を用いて推定した。推定すべきパラメタ Ψ は、$\Psi = (\mu_0, \mu_1, \omega, \alpha, \beta, \gamma)$ である。

まず、RR での、GARCH(1,1)+RV の推定結果を挙げる。

[Table 2:Estimation coefficients using RR insert here]

次に、RV での、GARCH(1,1)+RR の推定結果を挙げ、RR を用いたモデルと比較する。

[Table 3:Estimation coefficients using RV insert here]

上の比較から分かる、最も大きなことは、RV をモデルに加えることよりも、RR をモデルに加えたほうが尤度が高くなり、モデルの推定結果が良くなるということである。これは、RV よりも RR を用いる優位性があるということの一つの大きな根拠となる。

4.2 Estimation of ARFIMA model

上では、RV との比較をボラティィリティ変動モデルの枠組みを用いて行った。次に、RV、RR の特性である、long memory を定式化したモデルを用いて、RV と RR の比較を行っていく[6]。

4.2.1 ARFIMA モデル

ARFIMA モデルとは、長期記憶性をモデル化したもので、Hosking J.R.M (1981)、Beran J (1994) などによって考察されている。z_t が ARFIMA(p,d,q) に従うとすると、

[6]例えば

$$\phi(B)(1-B)^d(z_t-\mu)=\theta(B)\epsilon_t$$
$$p,q\in Z, d\in R$$
$$\phi(B)=1-\sum_{i=1}^{p}\phi_i B^i$$
$$\theta(B)=1-\sum_{j=1}^{q}\theta_j B^q$$
$$\epsilon_t \sim N(0,\sigma^2)$$
$$Bz_t=z_{t-1}$$

z_t は以上の式を満たすこととなる。

この ARFIMA モデルの特別な場合、つまり、$p=0,q=0$ のパラメタ $\Psi=(\mu,\sigma^2,d)$ を推定する。先行研究では、（擬似）最尤法などで推定されているが[7]、今回は、Markov Chain Monte Carlo 法を用いて推定した[8]。推定方法は以下である。

ベイズの定理より、

$$\pi(\Psi|\log RR)\propto f(\log RR|\Psi)\pi(\Psi)J$$

ここで、$\pi(\Psi|RR)$ は事後確率密度関数、 $f(\log(RR)|\Psi)$ は尤度関数、 $\pi(\Psi)$ は事前確率密度関数である。また、J はヤコビアンである。

本論文では、ARFIMA$(0,d,0)$ を推定する。その確率密度関数は、

$$(1-B)^d(\log RR_t-\mu)=\epsilon_t$$
$$\epsilon_t \sim N(0,\sigma)$$

推定すべきパラメタは $\Psi=(\mu,\sigma^2,d)$ である。誤差項が正規分布に従うことから、尤度は、

$$f(\log RR|\Psi)\propto(\sigma^2)^{-\frac{T}{2}}\exp\left(-\frac{1}{2\sigma^2}\sum_{t=1}^{n}\hat{\epsilon_t}^2\right)$$

ここで、T はデータ数である。

また、パラメタ d には、$0\leq d\leq 1$ の制約をつけるので、変数変換を行う。

$$d=\frac{1}{1+e^(-y)}$$
$$\therefore J=\frac{e^{-y}}{(1+e^{-y})^2}$$

[7]例えば

[8]MCMC 法については、Hastings W.K. (1970)、和合肇 (2005) などを参照

ベイズ推定の枠組みでは、パラメタの事前分布を設定する必要があるので、パラメタの事前分布を以下のように設定する。

$$
\begin{aligned}
\pi(\Psi) &\propto \pi(\mu)\pi(\sigma^2)\pi(y) \\
\pi(\mu) &\propto \exp\left(-\frac{\mu^2}{2\sigma_\mu^2}\right) \\
\pi(y) &\propto \exp\left(-\frac{y^2}{2\sigma_y^2}\right) \\
\pi(\sigma^2) &\propto (\sigma^2)^{-(\frac{\alpha_0}{2}+1)}\exp\left(-\frac{\beta_0}{2\sigma^2}\right)
\end{aligned} \tag{3}
$$

尤度は上で計算したとおりである。よって、事後確率密度関数は、

$$
\begin{aligned}
\pi\big(\Psi|\log RR\big) \propto (\sigma^2)^{-\frac{T}{2}}\exp\left(-\frac{T}{2}\sum_{i=1}^{T}u_t^2\right)\frac{e^{-y}}{(1+e^{-y})^2} \\
\times\exp\left(-\frac{\mu^2}{2\sigma_\mu^2}-\frac{y^2}{2\sigma_y^2}\right)(\sigma^2)^{-(\frac{\alpha_0}{2}+1)}\exp\left(-\frac{\beta_0}{2\sigma^2}\right)
\end{aligned}
$$

MCMC 法とは、事後確率密度関数からサンプリングをして、パラメタの書き換えを繰り返すことにより、推定値を求める方法である。このモデルでは、事後確率密度関数がよく知られた分布ではないので、MH アルゴリズムを用いてサンプリングを行う。本論文では、独立連鎖 MH アルゴリズムを用いて、推定値を発生させた[9]。5000 回の繰り返しを行って、最初の 1000 回を初期値に依存する期間 (burn-in period) として棄てた。そのアルゴリズム、推定結果は以下である。

Algorithm

Step.1 初期値 $\mu^{(0)},\sigma^{2(0)},d^{(0)}$ の設定

Step.2 σ^2 の発生　条件付事後確率密度関数を考える。σ^2 の条件付事後密度関数は、

$$
\begin{aligned}
f(\sigma^2|\log RR,\mu,y) &\propto f(\log RR|\mu,\sigma^2,y)\pi(\sigma^2) \\
&\propto (\sigma^2)^{-\frac{\alpha_1}{2}+1}\exp\left(-\frac{\beta_1}{2\sigma^2}\right) \\
&where \quad \alpha_1=\alpha_0+T,\beta_1=\beta_0+\sum_{t=1}^{T}\hat{\epsilon}_t{}^2
\end{aligned}
$$

[9]Chib S and Greenberg E (1994) を参照

よって、

$$\sigma^2 \sim IG\left(\frac{\alpha_1}{2}, \frac{\beta_1}{2}\right)$$

であるから、ギブス・サンプラーでサンプリングを行う。

Step.3 μ の発生　μ の条件付事後確率密度関数は、

$$\begin{aligned} f(\mu|\log RR, \sigma^2, y) &\propto f(\log RR|\mu, \sigma^2, y)\pi(\mu) \\ &\propto \exp\left(-\frac{1}{2\sigma^2}\sum_{t=1}^{T}\hat{\epsilon_t}^2 - \frac{\mu}{2\sigma_\mu^2}\right) \end{aligned}$$

$\hat{\mu}$ を条件付事後確率密度関数のモードとする。条件付事後確率密度関数の対数を取り、2 次テーラー展開をモードで評価すると、

$$\begin{aligned} g(\mu|\log RR, \sigma^2, y) &\equiv \log f(\mu|\log RR, \sigma^2, y) \\ &\propto -\frac{1}{2\sigma^2}\sum_{t=1}^{T}\hat{\epsilon_t}^2 - \frac{\mu}{2\sigma_\mu^2} \\ &\approx g(\hat{\mu}|\log RR, \sigma^2, y) + g_1(\mu-\hat{\mu}) + \frac{1}{2}g_2(\mu-\hat{\mu})^2 \\ &\equiv h(\mu|\log RR, \sigma^2, y) \\ where \;\; g_1 &= \left.\frac{\partial g(\mu|\log RR, \sigma^2, y)}{\partial \mu}\right|_{\mu=\hat{\mu}}, g_2 = \left.\frac{\partial^2 g(\mu|\log RR, \sigma^2, y)}{\partial \mu^2}\right|_{\mu=\hat{\mu}} \end{aligned}$$

ここで、$h(\mu|\log RR, \sigma^2, y)$ を整理すると、

$$\begin{aligned} h(\mu|\log RR, \sigma^2, y) &\propto -\frac{(\mu-b_1)^2}{2B_1} \\ where \;\; B_1^{-1} &= -g_2, b_1 = \hat{\mu} + B_1 g_1 \end{aligned}$$

よって、次点の μ を μ'、現在の点を μ と表し、μ' の提案密度 $q(\mu'|\log RR, \sigma^2, y)$ を、$N(b_1, B_1)$ とする。採択確率 α は、

$$\begin{aligned} \alpha &= \min\left\{1, \frac{\pi(\mu'|\log RR, \sigma^2, y)/q(\mu'|\log RR, \sigma^2, y)}{\pi(\mu|\log RR, \sigma^2, y)/q(\mu|\log RR, \sigma^2, y)}\right\} \\ &= \min\Bigg\{1, \exp\Bigg(g(\mu'|\log RR, \sigma^2, y) - h(\mu'|\log RR, \sigma^2, y) \\ &\qquad -g(\mu|\log RR, \sigma^2, y) + h(\mu|\log RR, \sigma^2, y)\Bigg)\Bigg\} \end{aligned}$$

となる。

Step.4 y の発生　y の条件付事後確率密度関数は、

$$
\begin{aligned}
f(y|\log RR,\mu,\sigma^2) &\propto f(\log RR|\mu,\sigma^2,y)\pi(y)J \\
&\propto \exp\left(-\frac{1}{2\sigma^2}\sum_{t=1}^{T}\hat{\epsilon_t}^2-\frac{y}{2\sigma_y^2}\right)\frac{e^{-y}}{(1+e^{-y})^2}
\end{aligned}
$$

$\hat{y}$ を条件付事後確率密度関数のモードとする。条件付事後確率密度関数の対数を取り、2 次テーラー展開をモードで評価すると、

$$
\begin{aligned}
k(y|\log RR,\mu,\sigma^2) &\equiv \log f(y|\log RR,\mu,\sigma^2,y) \\
&\propto -\frac{1}{2\sigma^2}\sum_{t=1}^{T}\hat{\epsilon_t}^2-\frac{y}{2\sigma_y^2}+\log\left(\frac{e^{-y}}{(1+e^{-y})^2}\right) \\
&\approx k(y|\log RR,\mu,\sigma^2)+k_1(y|\log RR,\mu,\sigma^2)+\frac{1}{2}k_2(y|\log RR,\mu,\sigma^2)^2 \\
&\equiv l(y|\log RR,\mu,\sigma^2) \\
where\ \ k_1 &= \left.\frac{\partial k(y|\log RR,\mu,\sigma^2)}{\partial y}\right|_{y=\hat{y}}, k_2=\left.\frac{\partial^2 k(y|\log RR,\mu,\sigma^2)}{\partial y^2}\right|_{y=\hat{y}}
\end{aligned}
$$

ここで、$l(y|\log RR,\mu,\sigma^2)$ を整理すると、

$$
l(y|\log RR,\mu,\sigma^2) \propto -\frac{(y-c_1)^2}{2C_1}
$$
$$
where\ \ C_1^{-1}=-k_2, c_1=\hat{y}+C_1k_1
$$

よって、次点の y を y'、現在の点を y と表し、y' の提案密度 $q(y'|\log RR,\mu,\sigma^2)$ を、$N(c_1,C_1)$ とする。採択確率 α は、

$$
\begin{aligned}
\alpha &= \min\left\{1,\frac{\pi(y'|\log RR,\mu,\sigma^2)/q(y'|\log RR,\mu,\sigma^2)}{\pi(y|\log RR,\mu,\sigma^2)/q(y|\log RR,\mu,\sigma^2)}\right\} \\
&= \min\left\{1,\exp\left(k(y'|\log RR,\mu,\sigma^2)-l(y'|\log RR,\mu,\sigma^2y)\right.\right. \\
&\qquad \left.\left.-k(y|\log RR,\mu,\sigma^2)+l(y|\log RR,\mu,\sigma^2)\right)\right\} \qquad (4)
\end{aligned}
$$

となる。

Step.5 Step.2 にもどる

以上のアルゴリズムを一定以上の回数繰り返し、burn-in period 以降のサンプルを保存し分析した。その結果が以下である。

[Table 4:Estimation Arfima(0,d,0) parameters using log(RR) insert here]

[Table 5:Estimation Arfima(0,d,0) parameters using log(RV) insert here]

Fig4,5,6,7,8,9 はそれぞれ、$\log RR, \log RV$ をデータとして用いたパラメタの推定値のサンプルパス、ヒストグラム、計算された自己相関である。

4.3 ボラティリティ予測

以上、いくつかのモデルを用いて、RR、RV の比較をおこなってきたが、それらのモデルをボラティリティ予測という観点から比較する。

Koopman S. J. ,Jungbacker B. and Hol E. (2005) にのっとり、RMSE(root mean squared error)、RMSPE(root mean squared parcentage error)、MAE(mean absolute error)、MAPE(mean absolute parcentage error) という指標を計測することで比較を行う。これらは、以下の式によって導かれる。

$$RMSE = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(RV_t - \hat{\sigma}^2_{t|t-1}\right)^2}$$

$$RMSPE = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\frac{\left(RV_t - \hat{\sigma}^2_{t|t-1}\right)^2}{RV_t}}$$

$$MAE = \frac{1}{T}\sum_{t=1}^{T}\left|RV_t - \hat{\sigma}^2_{t|t-1}\right|$$

$$MAPE = \frac{1}{T}\sum_{t=1}^{T}\left|\frac{RV_t - \hat{\sigma}^2_{t|t-1}}{RV_t}\right|$$

これらの指標には、基準となる値が必要である。これまでは基準値として日次リターンの 2 乗が用いられることが多かったが、Andersen, Torbin G. and Bollerslev.T (1998) によって、真のボラティリティの代理変数として日次リターンの 2 乗を用いると、日次リターンの式から z_t^2 の変動にも依存することが指摘された。彼らは、この z_t^2 の変動が大きいため基準として日次リターンの 2 乗を用いると、ボラティリティの予測パフォーマンスが正しく評価できないということを示した。また、基準値として RV を用いることを提案しており、本論文でもそれにしたがって分析を行う。Table6 は、GARCH(1,1)、ARFIMA(0,d,0) のそれぞれで、RV、RR、lnRV、lnRR を用いて、指標を計算した結果である[10]。

[Table 6: RMES,RMSPE,MAE,MAPE insert here]

5 Conclusion and Future Reserch

以上の比較から、RR を用いることの有用性が示された。

また、更なる RR の分析として、以下のことがあげられる。たとえば、日本の他のデータでも、同様な結果が得られるか、また、ここでは分析されなかったモデル、たとえば GJR モデルに RR を加えて、ボラティリティの分析を行うことも重要である。

[10]初期値 $\hat{\sigma}^2$ には、データの RV の最初の値を用いた。

次に、ARFIMA(0,d,0) の予測パフォーマンスがそれほど良くないので、Koopman S. J. ,Jungbacker B. and Hol E. (2005) などで、より良い結果が得られている ARFIMA(0,d,1) で Realized Range-based Variance を分析することが急務である。

また、ボラティリティの予測パフォーマンスを測る指標として、今回は MSE、MSPE、MAE、MAPE を用いたが、これには様々な議論があり、さらに正確にパフォーマンスを計測するには、VaR(value at risk) などを MCMC で推定することも非常に重要である[11]。

参考文献

Andersen, Torbin G. and Bollerslev.T (1998) "Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts", *International Economic Review*, Vol. 39, No. 4, pp. 885–905.

Barndorff-Nielsen O. E. and Shephard N. (2002) "Econometric analysis of realized volatility and its use in estimating stochastic volatility models", *Journal of Eoyal Statistical Society, B*, Vol. 64, No. 2, pp. 253–280.

Beran J (1994) *Statistics for long-memory process*: Chapman and Hall.

Black F and Scholes M (1973) "The pricing of options and corporate liabilities", *Journal of Political Economy*, Vol. 81.

Bollerslev .T (1986) "Generalized Autoregressive Conditional Heteroskedasticity", *Journal of Econometrics*, Vol. 31.

Chib S and Greenberg E (1994) "Bayes inference in regression models with ARMA(p,q) errors", *Journal of Econometrics*, Vol. 64.

Christensen .K and Podolskij .M (2005) "Realized range-based estimation of integrated varience".

Engle R.F. (1982) "Autoregressive Conditional Heteroskedasticity with Estimates of the Varience of United Kingdom Inflation", *Econometrica*, Vol. 54.

Hansen P.R. , Asger L. (2006) "A Forecast Comparison of Volatility Models: Does Anything beat a GARCH(1,1)?", *Journal of Econometrics*, Vol. 131, pp. 97–121.

Hastings W.K. (1970) "Monte carlo sampling methods using markov chains and thier applications", *Biometrika*, Vol. 57.

[11] RV については渡部 (2000) で分析されている

Hosking J.R.M (1981) “Fractional differencing”, *Biometrika*, Vol. 68.

Jeffrey S.P and Ravishanker N (1998) “Bayesian analysis of autoregresive fractionally integrated moving average process”, *Journal of Time Series Analysis*.

Koopman S. J. ,Jungbacker B. and Hol E. (2005) “Forcasting Daily Variability of the S&P 100stock Index using Histrical, Realized and Implied Volatility Measurements”, *Journal of Empirical Finance*, Vol. 12, pp. 445–475.

Martens .M and Dijk .D“Measuring volatility with the realized range”.

Paekinson .M (1980) “The extreme value methods for estimating the variance of the rate of return”, *Journal of Business*, Vol. 53, No. 1.

柴田舞 (2006) 「高頻度データによるボラティリティの推定:Realized Volatility のサーベイと日本の株価指数および株価指数先物の実証分析」.

渡部敏明 (2000) 『ボラティリティ変動モデル』, 朝倉書店.

渡部敏明佐々木浩二 (2006) 「ARCH 型モデルと Realized Volatility によるボラティリティ予測と Value-at-risk」.

和合肇（編） (2005) 『ベイズ計量経済分析マルコフ連鎖モンテカルロ法とその応用』, 東洋経済新報社.

Table 1:Basic Statistics

	Mean	Stdev	skeuness	kurtosis
RV	1.02	0.65	1.86	8.19
RR	0.22	0.12	1.87	9.01
lnRV	-9.36	0.58	0.11	2.70
lnRR	-10.83	0.47	0.21	2.97

Table 2:Estimation coefficients using RR

	coefficient	Std.err	Z-value	P-value
μ_0	-0.073	0.042	-1.72	0.043
μ_1	0.084	0.033	2.52	0.0059
ω	0.43	0.14	3.00	0.0013
α	-0.037	0.024	-1.55	0.060
β	0.31	0.15	2.07	0.020
γ	4.19	1.02	4.11	0.00

Log likelihood = -772.56

Table 3:Estimation coefficients using RV

	coefficient	Std.err	Z-value	P-value
μ_0	-0.072	0.043	-1.65	0.050
μ_1	0.086	0.034	2.51	0.0061
ω	0.32	0.095	3.33	0.00043
α	-0.0090	0.023	-0.38	0.35
β	0.60	0.079	7.55	0.00
γ	0.46	0.12	3.79	0.00

Log likelihood = -781.38

Table 4:Estimation Arfima(0,d,0) parameters using log(RR)

	Mean	Std.dev	95% lower quantile	Median	95% upper quantile
μ	-10.70	0.15	-10.99	-10.71	-10.39
σ^2	0.13	0.0060	0.12	0.13	0.14
d	0.43	0.026	0.38	0.43	0.48

Table 5:Estimation Arfima(0,d,0) parameters using log(RV)

	Mean	Std.dev	95% lower quantile	Median	95% upper quantile
μ	-9.36	0.12	-9.59	-9.36	-9.13
σ^2	0.26	0.012	0.24	0.26	0.28
d	0.31	0.024	0.27	0.31	0.36

Table 6: RMES,RMSPE,MAE,MAPE

	RMSE	RMSPE	MAE	MAPE
GARCH(1,1):RV	1.10	1.98	0.98	1.50
GARCH(1,1):RR	1.07	1.84	0.94	1.41
ARFIMA(0,d,0):lnRV	9.76	22.03	11.78	15.87
ARFIMA(0,d,0):lnRR	5.06	19.097	11.59	15.45

Fig 1 :Sample path of RV and RR and logarithm respectively

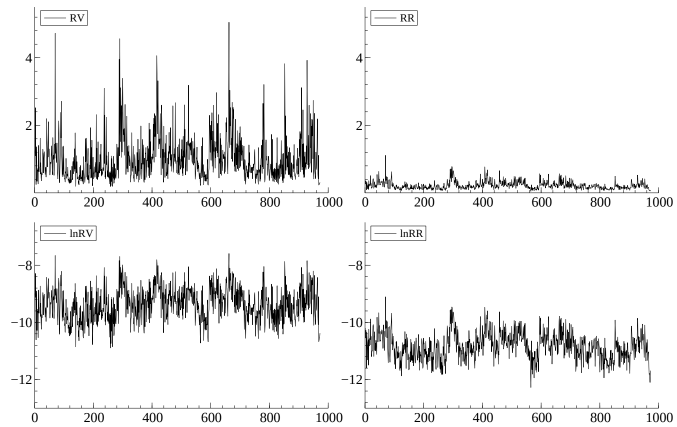

Fig 2 :Density of RV and RR and logarithm respectively

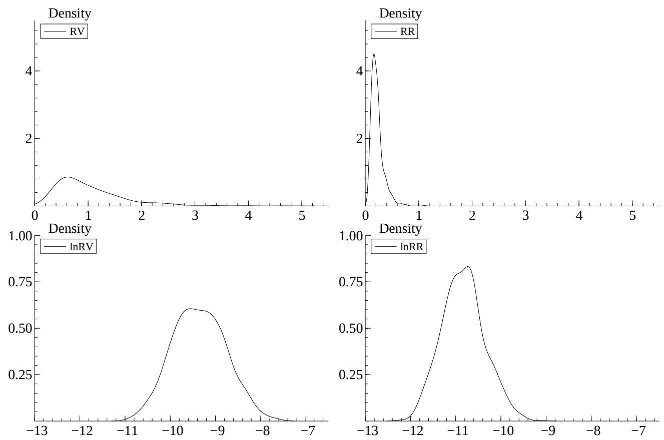

Fig 3 :Auto correlation function of RV and RR and logarithm respectively

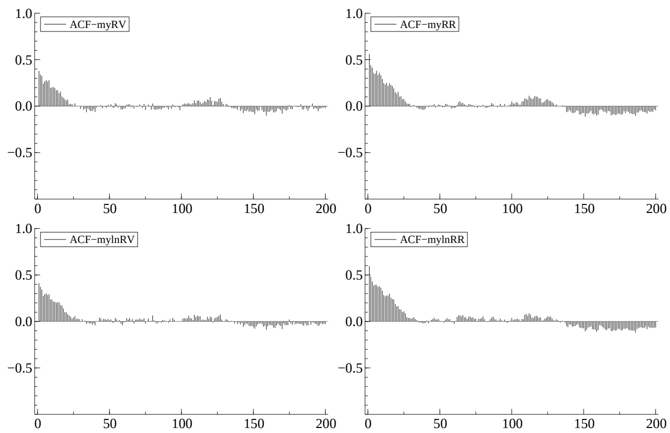

Fig 4 :Sample path of μ,σ^2,d respectively using log RR

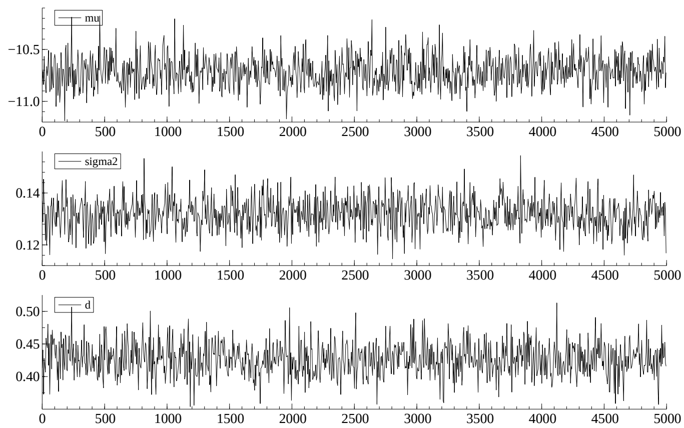

Fig 5 :Density of μ,σ^2,d respectively using log RR

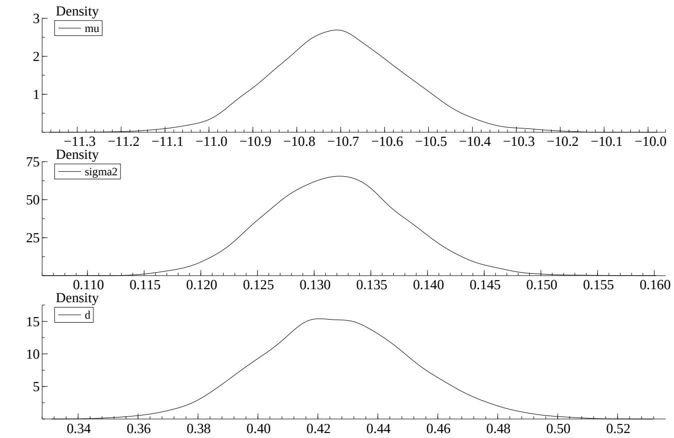

Fig 6 :Auto correlation function of μ,σ^2,d respectively using log RR

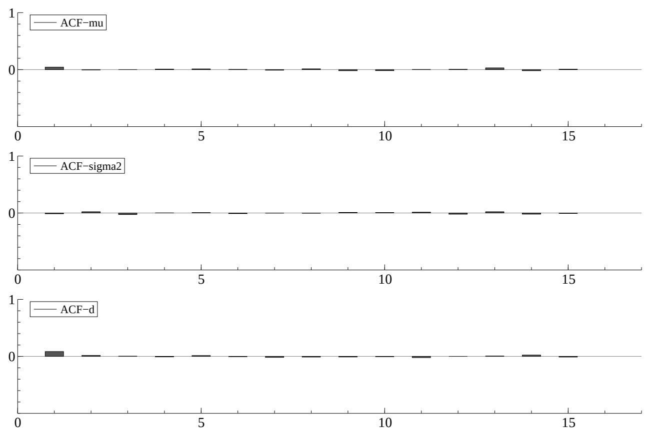

Fig 7 :Sample path of μ,σ^2,d respectively using log RV

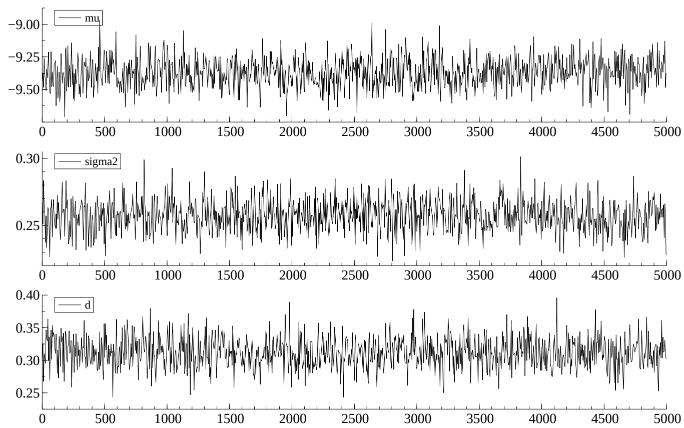

Fig 8 :Density of μ,σ^2,d respectively using log RV

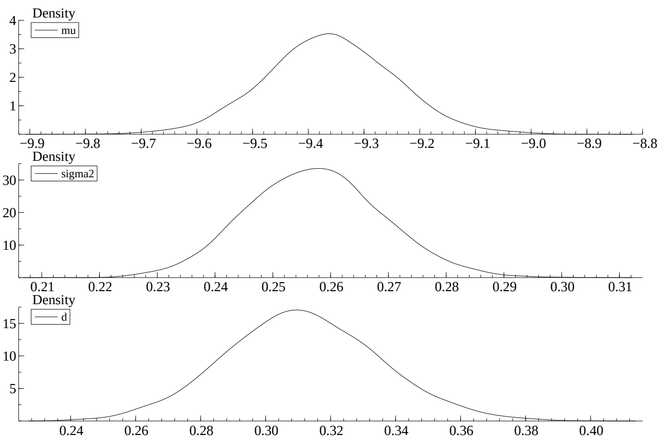

Fig 9 :Auto correlation function of μ,σ^2,d respectively using log RV