Discussion Paper Series

No.230

長期記憶過程におけるノイズの有無の検定

山口圭子

November 2007

Hitotsubashi University Research Unit
for Statistical Analysis in Social Sciences
A 21st-Century COE Program

Institute of Economic Research
Hitotsubashi University
Kunitachi, Tokyo, 186-8603 Japan
http://hi-stat.ier.hit-u.ac.jp/

長期記憶過程におけるノイズの有無の検定

山口　圭子 *

Testing for the presence of noise in long memory processes

Keiko Yamaguchi *

概 要

本論文では, ある変数が長期記憶過程に従うシグナルとホワイト・ノイズに従うノイズの和として表されるとした場合に, ノイズが存在するかどうかを検定するための新たな統計量を考えた。同様な統計量に Sun-Phillips (2003) が考えた統計量があるので, モンテカルロ実験によって, 彼らの統計量とここで新たに考えた統計量の性質を調べた。また, 長期記憶過程に従うことが指摘されている系列に Realized Volatility (RV) があるので, 日経 225 株価指数の 5 分ごとのリターンを用いて計算した RV と 1 分ごとのリターンを用いて計算した RV に対してノイズが存在するかどうか検定を行った。

In this paper, we propose a new test for the presence of noise in the long-memory signal plus white noise model. A similar test was proposed by Sun-Phillips(2003), so we conduct simulation experiments to examine and compare the finite sample properties of these two tests. It is well-known that the realized volatility(RV) follows a long memory process, so we apply these tests to the RVs calculated using the 1- and 5-minitues returns of the Nikkei 225 stock index.

Key Words and Phrases : 長期記憶性, realized volatility, 観測誤差, セミパラメトリック, 局所 Whittle モデル

JEL classification : C22

*一橋大学大学院経済学研究科, 〒 186-8601, 東京都国立市中 2-1
本論文は一橋大学 21 世紀 COE プログラム「社会科学の統計分析拠点構築」および文部科学省特別研究促進費「高頻度データを用いた日本の証券市場の計量分析」より助成を受け購入したデータ (NEEDS-TICK データ) を利用している。また, 科学研究費補助金 (特別研究員奨励費) から助成を受けている。一橋大学経済研究所の渡部敏明先生, 同大学経済学研究科の田中勝人先生からは有益なアドバイスを頂いた。記して感謝の意を表したい。

1 はじめに

本論文では, 以下のモデルのもとで, 長期記憶過程にノイズがあるかどうかの検定を考える。

$$X_t = Y_t + \eta_t, \quad t = 1, \ldots, n \tag{1}$$

ここで $\{X_t\}$ だけ観測可能で, $\{Y_t\}$ はシグナル, $\{\eta_t\}$ はノイズである。$\{Y_t\}$ の和分オーダーは d^*, $0 < d^* < 1/2$ であり, $\{\eta_t\}$ は平均 0, 分散 σ_η^2 のホワイトノイズ (以下 $WN(0, \sigma_\eta^2)$ と略) で, $\{Y_t\}$ と $\{\eta_t\}$ は互いに独立とする。ただし, 以下 d はパラメータ, d^* は真の値とする。このモデルのもとでノイズがあるかどうか検定するためには, 帰無仮説と対立仮説をそれぞれ以下のようにすればよい。

$$H_0 : \sigma_\eta^2 = 0 \quad \text{vs.} \quad H_1 : \sigma_\eta^2 > 0 \tag{2}$$

Tanaka (2002) は, $\{Y_t\}$ が ARFIMA$(0, d, 0)$ の場合の (2) の LM 検定統計量を導出した。帰無仮説はパラメータ空間の境界上であるが, LM 検定統計量は漸近的に正規分布に従う (King-Hillier (1985))。

ARFIMA(p, d, q) モデルでは, モデルの特定を誤ると一致性を失うことがあることが知られている。そこで, d をセミパラメトリックに推定する方法に関して数多くの研究が近年なされてきた。特に, 対数回帰モデルと局所 Whittle モデルが代表的である。(2) 式の検定問題もセミパラメトリックモデルで考えれば, $\{Y_t\}$ に対して短期的な変動の特定化を避けてより一般的な仮定をおける。Sun-Phillips (2003) は対数回帰モデルを拡張して (1) 式の $\{Y_t\}$ の d の推定を扱い, (2) のワルド検定統計量を導出した。ワルド検定なので統計量の漸近分布は $\sigma_\eta^2 = 0$ で mass をもってしまう。Hurvich et al. (2005) は局所 Whittle モデルを拡張して (1) において d とシグナルの大きさと σ_η^2 の比 の推定量の漸近分布を求めた。本論文では, 局所 Whittle モデルは対数回帰モデルとは違い, 目的関数が尤度関数の形をしていることに着目し, セミパラメトリックな枠組みでの (2) の LM 検定統計量を導出した。

また, Sun-Phillips (2003) ではシミュレーション実験がなされていないので, LM 検定統計量とともにシミュレーション実験を行い, パフォーマンスを比較した。両者とも size distortion が大きいことが分かった。また, 若干, LM 検定統計量のほうがサイズは良かった。

ボラティリティのノンパラメトリックで精度の高い推定量として “Realized Volatility” (以下, RV と略す) が注目を集めている。RV は日中リターンの 2 乗を足し合わせることにより求められる。RV には長期記憶性があることが知られており ARFIMA モデル (Andersen et al. (2001)) で定式化されることが多い。RV は, ノイズがない場合は, 日中リターンの時間間隔を 0 に近づけることにより真のボラティリティに収束する。しかし, 実際には, 時間間隔を小さくするとマイクロストラクチャーノイズによる影響が大きくなるというトレードオフの関係にあることが知られている。そこで先行研究では 5 分ごとのデータが使われていることが多い 。ここでは, 日経 225 株価指数の 5 分ごとと 1 分ごとのリターンを使って RV を計算し, それぞれの RV に対し (2) 式の検定を行なった。その

結果, 1 分ごとのリターンを使って計算した RV のほうが 5 分ごとのものより多く棄却された。これは, RV を計算するために用いる日中リターンの時間間隔が短いほど, RV がマイクロストラクチャーノイズの影響を大きく受けるという事実と整合的である。

本論文は以下, 次のような構成になっている。まず, 2 節でモデルを説明し, LM 検定統計量を導出する。3 節ではシミュレーションにより統計量の有限標本でのパフォーマンスを調べる。さらに, 4 節では, 日経 225 株価指数の RV について実証分析を行う。最後に 5 節で本論文をまとめるとともに, 今後の課題を述べる。

2　モデルと検定統計量

2.1　定義と記法

$\{Z_t\} \sim WN(0,1)$ とする。(1) 式で $\{Y_t\}$ を

$$Y_t = (1-L)^{-d^*} W_t, \qquad W_t = \sum_{j\in\mathbb{Z}} b_j Z_{t-j} \tag{3}$$

とする。$b(x) = \sum b_j e^{ijx}$ と定義し, b は 0 で連続, $b(0) > 0$ であると仮定する。 ここで, $\{Y_t\}$, $\{W_t\}$ のスペクトル密度関数をそれぞれ $f_Y(x)$, $f_W(x)$ とすると,

$$\begin{aligned} f_Y(x) &= |1-e^{ix}|^{-2d^*} f_W(x) \\ &= |1-e^{ix}|^{-2d^*} |b(x)|^2/(2\pi) \end{aligned} \tag{4}$$

となる。

(4) 式のようなスペクトル密度関数を持つものの例としては, 次の ARFIMA(p,d,q) モデルが挙げられる。

$$\Phi(L)(1-L)^d V_t = \Psi(L) Z_t \tag{5}$$

ここで, L はラグオペレータであり,

$$\Phi(L) = 1 - \phi_1 L - \cdots - \phi_p L^p, \qquad \Psi(L) = 1 - \psi_1 L - \cdots - \psi_q L^q$$

は $|\omega| \le 1$ に対して $\Phi(\omega) \neq 0, \Psi(\omega) \neq 0$ である。$-\frac{1}{2} < d$ ならば反転可能, $d < \frac{1}{2}$ ならば定常である。また, $0 < d$ ならば長期記憶性を持つ。

b が 0 の近傍における滑らかという仮定 (詳細後述) をおくと, $\{X_t\}$ のスペクトル密度関数 $f_X(x)$ は $x \to 0$ で ,

$$f_X(x) \sim x^{-2d^*} f_W(0) + \frac{\sigma_\eta^2}{2\pi} \equiv f_W(0) x^{-2d^*}(1 + \theta^* x^{2d^*}) \tag{6}$$

と近似できる。ここで,

$$\theta^* = \frac{\sigma_\eta^2}{2\pi f_W(0)} \tag{7}$$

である。

任意の系列 $\{V_t\}$ に対して $x_j = 2j\pi/n, j = 1, \ldots, n$ での離散フーリエ変換とピリオドグラムを以下のように表記する。

$$d_{V,j} = (2\pi n)^{-1/2} \sum_{t=1}^{n} V_t e^{-itx_j}, \quad I_{V,j} = |d_{V,j}|^2$$

以下では $\hat{d}, \hat{\theta}$ は推定値, d^*, θ^* は真の値, 何もついていない場合はパラメータを表すこととする。

2.2 局所 Whittle モデル

Hurvich et al. (2005) による局所 Whittle 尤度関数は,

$$\hat{J}_m(d, \theta) = \log \left(\frac{1}{m} \sum_{k=1}^{m} \frac{x_k^{2d} I_{X,k}}{1 + \theta x_k^{2d}} \right) + \frac{1}{m} \sum_{k=1}^{m} \log\{x_k^{-2d}(1 + \theta x_k^{2d})\} \tag{8}$$

で与えられる。ここで, m は $m = n^\alpha,\ 0 < \alpha < 1$ である。(6) 式の近似を使い, 原点付近で Whittle 尤度を評価したものである。

なお, (8) で $\theta = 0$ とすると Robinson(1995) の Gaussian Semiparametric Estimator (以下 GSE と略) に帰着する。GSE はノイズのない場合は $\alpha < 0.8$ で $\sqrt{m}$ 一致性をもつ。それに対し, Dalla et al. (2006) などで示されているように, ノイズがある場合は $\frac{4d^*}{4d^*+1} < \alpha$ のときに

$$\left(\frac{n}{m}\right)^{2d^*} (\hat{d} - d^*) \xrightarrow{p} -\theta B(d^*) \tag{9}$$

ただし, $B(d^*)$ は d^* の関数で正の値をとる, となり, $\left(\frac{n}{m}\right)^{2d^*}$ のオーダーで θ に比例する (言い換えるとノイズの大きさに比例する) 負の bias に支配されてしまう。Hurvich et al. (2005) は (8) 式のように θ も推定することにより, bias を取り除いた。

2.3 LM 検定統計量

以下, Hurvich et al. (2005) と同様の仮定をおく。

仮定 1 b は $[-\pi, \pi]$ 上で可積分, $b(-x) = \overline{b(x)}$ である。ある $\vartheta \in (0, \pi],\ \beta \in (0, 2],\ \mu > 0$ が存在して $x \in [-\vartheta, \vartheta]$ ならば

$$|b(x) - b(0)| \leq \mu |b(0)||x|^\beta, \quad \beta \in (0, 1] \tag{10}$$

$$|b(x) - b(0) - xb'| \leq \mu |b(0)||x|^\beta, \quad \beta \in (1, 2] \tag{11}$$

をみたす。

仮定 2　$\{Z_t\}$ は martingale difference 系列で, すべての t に対して $E[Z_t^4] = \mu_4 < \infty$ かつ $E[Z_t^2|\sigma(Z_s, s < t)] = 1\ a.s.$

仮定 3 $\sup_t E[\eta_t^4] < \infty$

仮定 4 m は非減少列で, ある任意に小さい $\delta > 0$ に対して,

$$\lim_{n\to\infty}(m^{-4d^*-1+\delta}n^{4d^*} + n^{-2\beta}m^{2\beta+1}\log^2(m)) = 0 \tag{12}$$

をみたす。

仮定 5 $\beta > 2d^*$

仮定 1 において, 例えば, $\{Y_t\}$ が ARFIMA 過程の場合は $\beta = 2$ である 。仮定 4 において, $m = n^\alpha$ とすると,

$$\frac{4d^*}{4d^*+1} < \alpha < \frac{2\beta}{2\beta+1} \tag{13}$$

である。$\beta = 2$ のときは $\alpha < 0.8$ である。α の下限は, (9) 式と関連していて, θ の識別に必要な仮定である。例えば $d^* = 0.5, 0.4, 0.25$ に対してそれぞれ 0.6667, 0.6153, 0.5 である。

次に, LM 検定統計量を求める。(2) の検定は $H_0 : \theta = 0 \quad \text{vs.} \quad H_1 : \theta > 0$ と同じである。

(8) より

$$m\frac{\hat{J}_m(d,\theta)}{\partial\theta}\bigg|_{H_0} = \frac{\sum_{k=1}^m \left\{x_k^{2\hat{d}} - \frac{1}{m}\sum_{l=1}^m x_l^{2\hat{d}}\right\} x_k^{2\hat{d}} I_{X,k}}{\frac{1}{m}\sum_{k=1}^m x_k^{2\hat{d}} I_{X,k}} \tag{14}$$

となる。$D_n^* = m^{1/2}\mathrm{Diag}(1, x_m^{2d^*})$ とする。Hurvich et al. (2005) の Proposition 4.1 より, $mD^{*-1}\nabla\hat{J}_m(d^*,\theta^*) \to N(0,\Gamma^*)$,

$$\Gamma^* = \begin{bmatrix} 4 & -\frac{4d^*}{(1+2d^*)^2} \\ -\frac{4d^*}{(1+2d^*)^2} & \frac{4{d^*}^2}{(1+2d^*)^2(1+4d^*)} \end{bmatrix}$$

であり, 畠中 (1996) p.264 より

$$Var\left(m\frac{\hat{J}_m(d,\theta)}{\partial\theta}\bigg|_{H_0}\right) = mx_m^{4d^*}(\Gamma_{22}^* - \Gamma_{21}^*\Gamma_{11}^{*-1}\Gamma_{12}^*) = mx_m^{4d^*}\frac{16{d^*}^4}{(1+2{d^*}^4)(1+4d^*)} \tag{15}$$

である。

以上から LM 検定統計量は

$$LM = \frac{\sum_{k=1}^m \left\{x_k^{2\hat{d}} - \frac{1}{m}\sum_{l=1}^m x_l^{2\hat{d}}\right\} x_k^{2\hat{d}} I_{X,k}}{\frac{1}{m}\sum_{k=1}^m x_k^{2\hat{d}} I_{X,k}} \Bigg/ \left(\frac{mx_m^{4\hat{d}}\, 16\hat{d}^4}{(1+2\hat{d}^4)(1+4\hat{d})}\right)^{1/2} \tag{16}$$

と求まる。帰無仮説のもとで, 漸近的に, $LM \to N(0,1)$ となる。

2.4 Sun-Phillips の検定統計量

Sun-Phillips(2003) の検定統計量は以下のとおりに計算する。(6) 式の f_X を $\{X_t\}$ のピリオドグラムに置き換えて両辺の対数をとると非線形対数回帰式が得られる。その平均二乗誤差

$$Q(d,\theta)=\frac{1}{m}\sum_{j=1}^{m}\left\{(\log I_{X,j}-\frac{1}{m}\sum_{k=1}^{m}\log I_{X,k})\right.\\ \left.+2d(\log x_j-\frac{1}{m}\sum_{k=1}^{m}\log x_k)-\theta(x_j^{2d}-\frac{1}{m}\sum_{k=1}^{m}x_k^{2d})\right\}^2 \tag{17}$$

を最小にする d,θ を求める。そして, θ の t 値

$$t_\theta=\sqrt{m}x_m^{2\hat{d}}\hat{\theta}/v(\hat{d}), \tag{18}$$

で t 検定すればよい。ただし $v(d)=\pi d^{-2}(2d+1)^2\sqrt{(4d+1)/96}$ である。t_θ は帰無仮説のもとで, 漸近的に, 1/2 の確率で 0 になり, 1/2 の確率で負の部分を切断した標準正規分布になる。

3 シミュレーション実験

次に, 以下のモデルから生成したデータを用いて, 2.3 節で提案した新たな LM 検定統計量と 2.4 節で説明した Sun-Phillips 統計量の有限標本でのパフォーマンスを分析する。

$$\begin{aligned} x_t &= y_t+\eta_t, \quad \{\eta_t\}\sim NID(0,\sigma_\eta^2), \quad t=1,\ldots,T \\ (1-\phi L)(1-L)^d y_t &= (1-\psi L)v_t, \quad \{v_t\}\sim NID(0,1) \end{aligned} \tag{19}$$

ただし, T=512, 1024, 2048 とし, 2000 回の繰り返しによりサイズと検出力を計算した。検定のサイズは 5%とした。

表 1

表 1 は $\phi=\psi=0$, (12) 式の $\alpha=0.75$ の場合の LM 検定のパフォーマンスである。$\sigma_\eta^2=0$ はサイズ を, 0 でないところは名目検出力を表す。T が大きくなると検出力は上がるが, σ_η^2 に対して単調増加になっていない。Tanaka(2002) でも同様のシミュレーション結果が報告されている。これは, ノイズ が大きくなり過ぎるとシグナルの推定が困難になるからであると考える。以下の実験ではすべて $T=2048$ とした。

表 2

表 2 では, 表 1 で $\sigma_\eta^2=0.5, 20$ のときに, GSE ((8) 式で $\theta=0$) での d の bias について調べたものである。表 2 では, bias が大きいのは α が大きいとき, $\sigma_\eta^2=20$ のときであり, これは (9) 式と整合的な結果である。表 1 で検出力が単調増加しないことを指摘したが, 検定をして帰無仮説が棄却されない場合, α をいくつか変えて GSE で d を推定し, α が小さいところより大きいところのほうが推定値が極端に小さければ, そのことは, ノイズが小さいからではなく大きすぎるからだということの目安にはなる。

表 3,4

表 3, 4 では LM 検定統計量と Sun-Phillips (SP) 統計量を比較した。$\phi = \psi = 0$ とした。表 3 では漸近的に正しい名目 5%棄却点 (=1.64) を使い, サイズと名目検出力を調べた。表 4 では実験で求めた経験 5%棄却点を使い, サイズ調整済み検出力を調べた。若干, LM 検定統計量のほうがサイズは良さそうであるが, 大幅に優れているとは言えない。また, α, d それぞれ大きいほうが検出力は大きい。

表 5,6

表 5,6 では ϕ, ψ の値を変えて LM 検定, SP 検定のサイズをそれぞれ調べた。$d = 0.4, \alpha = 0.75$ とし, ϕ, ψ はそれぞれ $(-0.8, -0.4, -0.2, 0.2, 0.4, 0.8)$ とした。これらの表から, ψ が大きいほど,ϕ が小さいほど帰無仮説を過剰に棄却し, ψ が小さいほど, ϕ が大きいほど過剰に受容することがわかる。特に ψ が大きいときの size distortion が大きいことがわかる。

表 7

以下では, この場合についてみていく。表 7 では α とサイズの関係を調べた。$\alpha = 0.75, 0.7$ とした。 α が大きいと size distortion も大きいことがわかる。表 3,4 の結果も合わせると, α を大きくすると検出力は大きいが, size distortion も大きくなるというトレードオフの関係にあるといえる。

表 8,9

表 8 では, $\alpha = 0.75, d = 0.4$ かつ $\psi = -0.8, -0.4, -0.2, 0, 0.2, 0.4, 0.8$ の場合における, 帰無仮説での局所 Whittle モデルの d の推定量の bias, standard error(SE) を調べた。$\psi = 0.8$ のときの d の推定値は 0 に近い値をとっており, その結果, LM 検定統計量 (16) 式の分母の値が不安定になっていると考えられる。(6) 式では原点の近傍において, $\{W_t\}$ のスペクトル密度関数は $f_W(0)$ で一定と近似しているが, 有限標本においてはその近似誤差により d の bais が生じる。ψ が大きいときは高周波部分が強くなるので, 負の bias が大きくなる。表 9 では $\psi = 0.8$ のときの d の bias を調べた。α を小さくすると, $f_W(0)$ で一定とした近似がより妥当になるので, bias は小さくなる。

表 2 と表 9 で, α と d の bias の関係が似ており, 区別が難しい。このような場合は判断を控えざるを得ない。

4　実証分析

ここでは, 日経 225 株価指数 RV の対数値 (sample size = 2476, 1996 年 3 月 11 日から 2006 年 3 月 31 日まで) に対してノイズの検定を行なった。まず, RV の定義, 計算方法について以下で少し述べる。

第 t 日の日中の n 個のリターンデータ $(r_t, r_{t+1/n}, \ldots, r_{t+(n-1)/n})$ が与えられたとき, それらをすべて 2 乗して足し合わせた

$$RV_t = \sum_{i=0}^{n-1} r_{t+i/n}^2 \tag{20}$$

を第 t 日の Realized Volatility (RV) という。資産価格の対数値 $\ln P(s)$ が

$$d \ln P(s) = \mu(s)ds + \sigma(s)dW(s) \tag{21}$$

に従っているとする。ここで, W(s) はウイナー過程である。すると, 第 t 日の真のボラティリティは

$$\sigma_t^2 = \int_t^{t+1} \sigma^2(s) ds \tag{22}$$

で定義される。

(20) 式で定義される RV は, $n \to \infty$ とすると, σ_t^2 に確率収束するので, n が十分大きいなら, RV_t は真のボラティリティσ_t^2 の精度の高い推定量となる。ただし, n を大きくするとRV に含まれるマイクロストラクチャーノイズ (非同時取引, bid-ask bounce, 無取引など) の影響が大きくなることが知られている。そこで, 先行研究では 5 分ごとの価格を使って RV を計算しているものが多い。なお, Bandi-Russell (2004, 2006) は平均 2 乗誤差 (mean square error) を最小にする時間間隔を選択する方法を提案している。

本論文では, 日経 NEEDS ティックデータを用いた。そこには, 日経 225 の 9:01 から前場の終わりまでと 12:30 から後場の終わりまでまでの 1 分ごとの価格が記録されている。そこで, 5 分ごとのリターンを使った RV^5 に加えて 1 分ごとのリターンを使った RV^1 も計算した。前場の終わりは 11:00, 後場の終わりは 15:00 であるが, 取引された時刻でなく, 日経 225 が計算され送信された時刻が記録されているので, 11:00 や 15:00 を超える時刻が記録されていることが多い。ここでは, 11:00 の価格には前場の終値を, 15:00 には後場の終値を用いた。また, 夜間 (前日の 15:00 から 9:00) や昼休み (11:00 から 12:30) は取引時間外であるので, 5 分ごとや 1 分ごとのリターンは得られない。夜間のリターンは 9:01 の価格の対数値と前日の終値の対数値の差をとり 100 倍することにより計算した。昼休みのリターンも同様に計算した。そして, 前場の 9:01, 9:05, . . . ,11:00 と後場の 12:31, 12:35, . . . , 15:00 を 5 分ごとの価格として取り出し, その対数階差を 100 倍することにより 5 分ごとのリターンを計算し, それらの 2 乗と夜間・昼休みのリターンの 2 乗を足し合わせることにより RV^5 を計算した。同様に 1 分ごとのリターンを計算し, それらの 2 乗と夜間・昼休みのリターンの 2 乗を足し合わせることにより RV^1 を計算した。

次に, 検定についてである。ここでは, $\ln(RV_t)$ を (1) 式の X_t, $\ln(\sigma_t^2)$ を (1) の Y_t と考えている。$\ln(RV_t)$ はリターンの時間間隔が 0 ではないことによる誤差 (Barndorff-Nielsen-Shephard(2002)) やマイクロストラクチャーノイズにより, $\ln(\sigma_t^2)$ から乖離する。そこで, それらの差 $\ln(RV_t) - \ln(\sigma_t^2)$ を (1) 式の誤差項 η_t としている。

η_t は Barndorff-Nielsen-Shephard(2002) や Bandi-Russell (2004, 2006) ではホワイトノイズと考えらてれる。RV や RV の対数値には長期記憶性があることが知られおり, そのモデル化には ARFIMA モデルがよく用いられている (Andersen et al. (2001), Koopman et al. (2005), 渡部-佐々木 (2006), 渡部 (2007))。このことと η_t はホワイトノイズであることと合わせると, $\ln(\sigma_t^2)$ は長期記憶過程に従っている可能性が高い。なお, σ_t^2 あるいは $\ln(\sigma_t^2)$ が長期記憶過程に従うような $\sigma(s)^2$ の確率過程としては, Fractional Ornstein-Uhlenbeck 過程（Comte-Renault(1996, 1998)) が挙げられる。

表 10,11,12

表 10-11 は $\ln(RV^5)$, $\ln(RV^1)$ に対してそれぞれ検定を行った結果である。d が 0.5 を超えているが, トレンドが無いときは, $d < 0.75$ の場合まで適用可能である。5%, 10% で棄却されることをそれぞれ [a], [b] と記す。表中の d は, LM 検定では H_0 , SP 検定では H_1

での推定値である。表から LM 検定, SP 検定ともに $\ln(RV^1)$ のほうが帰無仮説がより多く棄却されていることが分かる。表 12 は $\ln(RV^5)$ と $\ln(RV^1)$ の GSE の d の推定値であり, 表 9 のような問題はおきていないと考えられる。

ノイズがないという帰無仮説が棄却されているのは, RV の計算に用いた日中リターンの時間間隔が 0 ではないこと, 昼休み, 夜間のリターンの時間間隔が長いこと, またマイクロストラクチャーノイズが存在することなどから, RV が真のボラティリティに収束しておらず, ノイズを含んでいるためであると考えられる。RV の計算に用いる日中リターンの時間間隔を短くすると, RV はマイクロストラクチャーノイズの影響を強く受けることが知られており (Bandi-Russell (2004, 2006)), $\ln(RV^5)$ より $\ln(RV^1)$ のほうがより多く棄却されているのは, それと整合的である。

5 まとめと今後の課題

セミパラメトリックモデルである局所 Whittle モデルを用いて, 長期記憶過程にノイズが含まれるかどうかの LM 検定統計量を導出し, シミュレーション実験を行い Sun-Phillips (SP) 検定統計量と LM 検定統計量を比較を行なった。

また, 長期記憶過程に従うことが指摘されている系列に RV があるので, 日経 225 株価指数の 5 分ごとのリターンを用いて計算した RV と 1 分ごとのリターンを用いて計算した RV に対して LM, SP 検定を行った。

本論文では, (1) 式のノイズには自己相関がなく, またシグナルとも無相関であると仮定した。しかし, Hansen-Lunde (2006) は RV に含まれるマイクロストラクチャーノイズが自己相関を持つ場合や真のボラティリティと相関を持つ場合についても分析を行っているので, 今後の課題として, こうした仮定をはずした場合への拡張は重要である。また, (1) 式で d が非定常な場合へ, tapering(Velasco(1999), Hurvich(2000)) を利用することなどにより (13) 式をみたす範囲で拡張することも検討の余地がある。

参考文献

Andersen, T.G., Bollerslev, T., Diebold, F.X. and Ebens, H. (2001). The distribution of realized stock return volatility, Journal of Financial Economics, 61, 43–76.

Barndorff-Nielsen, O. E., and Shephard, N. (2002). Econometric analysis of realised volatility and its use in estimating stochastic volatility models, Journal of the Royal Statistical Society, Ser. B, 64, 253–280.

Bandi, F. and Russell, J. (2004). Microstructure noise, realized volatility, and optimal sampling, Manuscript, University of Chicag.

Bandi, F. and Russell, J. (2005). Separating microstructure noise from volatility, Journal of Financial Economics, 79, 655–692.

Comte, F. and Renault, E. (1996). Long memory continuous time models, Journal of Econometrics, 73, 101–149.

Comte, F. and Renault, E. (1998). Long memory in continuous-time stochastic volatility models, Mathematical Finance, 8 291–323.

Dalla, V., Giraitis, L. and Hidalgo, J. (2006). Consistent estimation of the memory parameter for nonlinear time series. Journal of Time Series Analysis, 27, 211–251.

Hansen, P.R. and Lunde, A. (2006). Realized variance and market microstructure noise, Journal of Business and Economic Statistics, 24, 127–161.

畠中道雄 (1996). 計量経済学の方法, 創文社.

Hurvich, C. M. and Chen, W. (2000). An efficient taper for potentially overdifferenced long-memory time series, Journal of Time Series Analysis, 26, 155–180.

Hurvich, C. M., Moulines, E. and Soulier, P. (2005). Estimation of long memory in volatility, Econometrica, 73, 1283–1328.

King, M.L. and Hillier, G.H. (1985). Locally best invariant test of the error covariance matrix of the Journal of the Royal Statistical Society, Series B, 47, 98–102.

Koopman, S.J., Jungbacker, B. and Hol, E. (2005). Forecasting daily variability of the S&P100 stock index using historical, realised and implied volatility measurements, Journal of Empirical Finance, 12, 445–475.

Robinson, P. M. (1995). Gaussian semiparametric estimation of long range dependence, Annals of Statistics, 23, 1630–1661.

Sun, Y. and Phillips, P.C.B. (2003). Nonlinear log-periodogram regression for perturbed fractional processes, Journal of Econometrics, 115, 355–389.

Tanaka, K. (2002). A unified approach to the measurement error problem in time series models, Econometric Theory, 18, 278–296.

Velasco, C. (1999). Non-stationary log-periodogram regression, Journal of Econometrics, 91, 325–371.

渡部敏明 (2007). Realized Volatility-サーベイと日本の株式市場への応用-, 経済研究, 58, 352–373.

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, 金融研究, 25, 39–74.

表 1: $\phi = \psi = 0, m = n^{0.75}$ のときの LM 検定のサイズ と 検出力 (%)

	サイズ	検出力				
σ^2_{eta}	0	0.5	1	5	10	20
T=512	3.70	8.95	11.75	15.10	13.75	18.30
1024	4.90	12.50	19.60	29.60	26.10	19.80
2048	5.75	17.10	27.95	50.15	45.65	34.30

表 2: ノイズの大小による GSE での d の推定 bias

α	0.75	0.7	0.6	0.5	0.4
$\sigma^2_\eta = 0.5$	-0.0503	-0.0393	-0.0237	-0.0169	-0.0155
20	-0.3075	-0.2928	-0.2203	-0.2203	-0.1796

表 3: サイズと名目検出力 (%)

			サイズ	検出力			
σ^2_η			0	1	5	10	20
α	d						
0.75	0.4	LM	5.75	27.95	50.15	45.65	34.30
		SP	3.95	25.65	42.60	41.40	33.60
	0.25	LM	4.60	12.55	12.60	12.65	20.55
		SP	6.80	14.45	17.00	13.45	8.65
0.7	0.4	LM	4.05	15.65	34.25	33.30	27.20
		SP	1.80	15.25	33.45	33.80	29.05
	0.25	LM	3.10	7.45	8.45	11.85	20.90
		SP	6.25	12.60	14.75	13.20	9.40

表 4: サイズ調整済み検出力 (%)

σ_η^2			1	5	10	20
α	d					
0.75	0.4	LM	26.25	47.30	43.55	32.65
		SP	27.65	45.50	43.70	35.75
	0.25	LM	13.65	13.55	13.40	20.80
		SP	11.60	14.00	11.20	6.65
0.7	0.4	LM	18.35	37.95	36.75	30.20
		SP	22.90	43.05	42.15	37.05
	0.25	LM	10.65	13.35	15.60	23.20
		SP	11.45	13.20	11.95	7.95

表 5: ARFIMA$(1, d, 1)$ のときの LM 検定のサイズ (%)

ψ	-0.8	-0.4	-0.2	0.0	0.2	0.4	0.8
$\phi =$ -0.8	6.55	8.10	11.80	19.45	42.65	83.35	100.00
-0.4	4.85	6.50	9.15	16.35	38.05	79.95	100.00
-0.2	3.40	4.45	6.50	12.60	31.80	75.55	100.00
0	0.95	1.70	2.95	5.75	19.55	64.20	100.00
0.2	0.00	0.15	0.35	0.95	6.55	37.95	100.00
0.4	0.00	0.00	0.00	0.00	0.15	6.60	98.75
0.8	0.00	0.00	0.00	0.00	0.00	0.00	6.60

表 6: ARFIMA$(1, d, 1)$ のときの SP 検定のサイズ (%)

ψ	-0.8	-0.4	-0.2	0	0.2	0.4	0.8
ϕ = -0.8	3.80	4.90	7.50	15.65	38.80	77.95	11.40
-0.4	2.75	3.80	5.50	12.45	33.55	74.70	14.35
-0.2	1.80	2.45	3.80	8.80	27.10	69.50	18.95
0	0.50	0.80	1.70	3.95	16.10	58.20	30.70
0.2	0.00	0.10	0.20	0.40	3.80	34.45	56.80
0.4	0.00	0.00	0.00	0.00	0.05	3.80	93.65
0.8	0.00	0.00	0.00	0.00	0.00	0.00	3.80

表 7: α とサイズの関係

ψ		-0.8	-0.4	-0.2	0.2	0.4	0.8
α							
0.75	LM	0.95	1.70	2.95	19.55	64.20	100.00
	SP	0.50	0.80	1.70	16.10	58.20	30.70
0.7	LM	1.35	3.40	3.90	9.05	21.80	98.80
	SP	0.95	1.10	1.30	4.05	17.10	94.80

表 8: ψ を変化させたときの局所 Whittle モデルの d のパフォーマンス

ψ	-0.8	-0.4	-0.2	0	0.2	0.4	0.8
bias	0.025	0.020	0.013	-0.003	-0.029	-0.083	-0.365
SE	0.031	0.031	0.031	0.031	0.031	0.031	0.032

表 9: $\psi = 0.8$ のときの d の bias

α	0.75	0.7	0.6	0.5	0.4
bias	-0.365	-0.292	-0.142	-0.050	-0.018

表 10: 日経 225 株価指数の $\ln(RV^5)$ の検定結果

α		0.75	0.725	0.7	0.675
LM	stat	2.527^a	1.774^a	0.852	0.678
	d	0.478	0.500	0.523	0.530
SP	stat	1.847^a	1.369^b	0.364	0.000
	d	0.606	0.608	0.576	0.561

表 11: 日経 225 株価指数の $\ln(RV^1)$ の検定結果

α		0.75	0.725	0.7	0.675
LM	stat	3.718^a	2.955^a	1.755^a	1.232^a
	d	0.457	0.482	0.512	0.529
SP	stat	3.624^a	3.300^a	2.200^a	1.93
	d	0.651	0.644	0.62	0.612

表 12: $\ln(RV^5)$ と $\ln(RV^1)$ の GSE の d の推定値

α	0.75	0.725	0.7	0.675	0.6	0.5	0.4
$\ln(RV^5)$	0.478	0.500	0.523	0.530	0.575	0.519	0.478
$\ln(RV^1)$	0.457	0.482	0.512	0.529	0.597	0.629	0.563