

株価収益率におけるボラティリティの長期依存性に関する一考察⁽¹⁾

前 川 功 ⁽²⁾ 一*
河 合 研 一**

1. はじめに

数理ファイナンス及び計量ファイナンスでは時系列の時点ごとの変動の大きさをボラティリティと呼ぶが、これは統計学的にいえば、各時点での分散である。Black-Scholes モデルに見られるように初期の株価変動モデルでは、単純化の仮定としてボラティリティは時間を通じて一定という仮定が置かれているが、もちろん時点ごとに分散は変動すると見るほうがより現実に近い。ボラティリティは Black-Scholes の公式に端的に現れているように、オプション価格を計算するために不可欠な変数である。そのためボラティリティに関する詳細かつ膨大な先行研究がなされてきた。そしてボラティリティの時間的変動を表現するさまざまなモデルがこれまでに提示されてきた。そして多くの先行研究において、ボラティリティ変動には長期依存性、あるいは長期記憶性といわれる現象が観察された結果、ボラティリティに長期依存性が存在することはほぼ定説として定着したとあってよい。本稿の目的は、このような現象の背後に存在するであろう構造的要因を探ることである。

近年、金融時系列データにおけるいわゆる高頻度データを利用した研究が活発化している。高頻度データとは株価や為替レートの1分間隔データ、あるいは取引ごとのすべてのデータを含むきわめて観測時間間隔の短い時系列データを指す。株式市場や為替市場には以前からこのようなデータは取引記録として存在していたが、実務家にとっては過去のこれほど詳細で大量なデータを保存し再利用する必要性はほとんどなかったと思われる。したがってそのようなデータが整理された形で保存され、研究者に提供されることもほとんどなかった。しかし金融工学的な実証研究

* 広島経済大学大学院経済学研究科教授

** 大学共同利用機関法人情報・システム研究機構 統計数理研究所リスク解析戦略研究センター研究員

が活発化し、計量ファイナンスあるいはフィナンシャル・エコノメトリックスという研究分野が確立するにつれ、高頻度データの重要性が認識され始め、またパソコンの大容量化と計算速度の向上とがあいまって高頻度データに基づく研究が著しく増大しつつある。まさに高頻度データは研究者にとっていわば宝の山ともいえる情報の宝庫である。このような研究者サイドからの需要の高まりに呼応して、商業ベースで高頻度データが（非常に高価でありまた寡占的ではあるが）販売されるようになってきた。このような状況のなかでわれわれも高頻度データの収集につとめ、研究を進めてきたところである。

高頻度データは多くの情報を含んでいる反面、実証分析を行う際に、従来の計量経済学や時系列分析における方法やモデルでは処理仕切れないいくつかの問題が生じてきた。主な問題として、以下のような点が挙げられる。すなわち、

- (1) 日内季節変動の問題。これは、株式市場などに見られる現象で、市場開始直後、昼休み直前、午後の市場開始直後、市場終了直前などに特長的な変動などがあるために、日内変動にある特定のパターンが生じる。このようなパターンの処理に関する問題を指す。
- (2) マイクロ・ストラクチャーノイズの問題。これは、高頻度であるがために発生するといわれるさまざまな雑音の処理に関する問題を指す。代表的なものに「bid-ask bounce」がある。これは取引はbidで成立したりaskで成立したりするために、取引価格が真の価格から乖離し、リターンに負の自己相関が生じるというものである（渡部敏明 2007）。このようなノイズを含んだままデータ処理をおこなうと不正確な結果を得る可能性がある。
- (3) 非同期性の問題。これは複数の高頻度時系列を扱う際に生じる。複数の高頻度時系列データは同時刻に観察されるのではない。すなわちデータの発生時刻がお互いにずれている。そのような場合、相関係数や共分散を計算する際の方法論的問題。この問題に対して最近 Hayasi-Yoshida (2007) はひとつの解決法を示した。
- (4) ボラティリティの長期依存性の問題。金融時系列だけでなく経済時系列には長期依存性または長期記憶性といわれる現象がしばしば見られる。市場に何らかのショックが発生するとその影響は時間と共に減衰していくが、減衰の仕方が極端に遅い場合、長期依存性があるという。長期依存性のある時系列の中で、一定の数学的条件を備えた系列を長期記憶過程という（より形式的な定義は後述）。株価収益率の高頻度データにおける2時点間での収益率の日内2乗和として定義される実現ボラティリティ（Realized Volatility）の記

語。以後 RV と略記する。) は計量ファイナンスでは重要な役割を果たすが、この RV には長期記憶性があることが多くの先行研究で指摘されている。

さて本稿では上に述べたこれらの問題点のうち、ボラティリティや RV の長期依存性あるいは長期記憶性の問題に焦点を絞って、長期記憶性の構造的要因を、現実のデータ及び現実を模したシミュレーション実験によって論じてみたい。

2. ボラティリティ変動モデル

経済時系列においてしばしば観察される現象としてボラティリティ・クラスタリングという現象がある。これは何らかの原因で時系列の変動が大きくなるとしばらく変動の大きい時期が持続し、大きな変動が収まると変動の小さい時期がしばらく持続するという現象である。このようなボラティリティ・クラスタリングを表現するモデルとして ARCH (Autoregressive Conditional Heteroskedasticity) モデル (Engle(1982)) とそれを一般化した GARCH (Generalized Autoregressive Conditional Heteroskedasticity) モデル (Bollerslev(1986)) とがある。本稿では以下、GARCH モデルに依拠した議論を展開するが、その前にこのモデルの定式化と特徴を簡単に要約しておこう。

2.1 GARCH モデル

GARCH(p,q) モデルは以下のように定式化される。

$$\begin{aligned} r_t &= e_t, & e_t &\sim N(0, \sigma_t^2) \\ e_t &= \sqrt{\sigma_t^2} \varepsilon_t, & \varepsilon_t &\sim N(0, 1) \\ \sigma_t^2 &= \omega + \sum_{i=1}^p \beta_i \sigma_{t-i}^2 + \sum_{j=1}^q \alpha_j e_{t-j}^2 \\ \omega > 0, \quad \beta_i, \quad \alpha_i &\geq 0 (i=1, 2, \dots, p; j=1, 2, \dots, q) \end{aligned}$$

ここに係数制約は $\sigma_t^2 > 0$ および高次のモーメントが存在するための条件である。GARCH モデルが定常であるための十分条件は

$$\sum_{i=1}^p \beta_i + \sum_{j=1}^q \alpha_j < 1$$

である。

3 番目の式は、時点 t 以前の情報が与えられれば時点 t におけるボラティリティ σ_t^2 が決定されることを示している。ことからこのモデルは条件付分散不均一 (Conditional Heteroskedasticity) モデルであり、またボラティリティ σ_t^2 に関して (1

階の) 自己回帰 (Autoregressive) 型になっている。多くの先行研究により、現実のデータの挙動は GARCH(1,1) モデルで十分捉えられることが知られているので、本稿では以下、GARCH(1,1) モデルに限定して考察をすすめる。

GARCH(1,1) モデルは次のように表わされる。

$$r_t = e_t, \quad e_t \sim N(0, \sigma_t^2) \quad (1)$$

$$e_t = \sqrt{\sigma_t^2} \varepsilon_t, \quad \varepsilon_t \sim N(0, 1) \quad (2)$$

$$\sigma_t^2 = \omega + \beta \sigma_{t-1}^2 + \alpha e_{t-1}^2 \quad (3)$$

このモデルにおいて観察値は r_t のみであり、 σ_t^2 は観察できない潜在変数ではあるが、観察値データから推定することが出来る。定常性の十分条件は $\alpha + \beta < 1$ である。そして $\alpha + \beta$ が 1 に近いほどボラティリティの持続性がたかまり、 $\alpha + \beta = 1$ のときを integrated GARCH (IGARCH) モデルといい、持続性は無限大となる。

次に長期記憶性の定義と長期記憶性を持つ代表的時系列モデルである ARFIMA (p,d,q) モデル及び FIEGARCH モデルの定式化とその特徴を要約する。

2.2 長期記憶時系列モデル, ARFIMA, FIGARCH, FIEGARCH

本稿では定常な時系列を考察の対象とする。長期記憶定常時系列とは、簡潔に直感的に表現すれば、時点 t における変動の影響が非常に緩やかに減衰しつつ長期間持続し、自己相関係数がなかなか 0 に収束しないような時系列のことである。これとは逆に、短期記憶系列とは、時点 t における変動の影響が短時間で減衰し、自己相関係数が比較的短時間で 0 に収束するような時系列である。このことを数式で表現すれば、次のように書ける。ある定常時系列 $\{y_t\}$ において h 時点隔たった 2 時点間の自己相関係数を $\rho(h)$ とするとき

$$\sum_{h=0}^{\infty} |\rho(h)| < \infty$$

が成立するとき短期記憶定常時系列といい、

$$\sum_{h=0}^{\infty} |\rho(h)| = \infty$$

が成立するとき長期記憶定常時系列という。後者は自己相関係数のゼロへの収束が極めて遅いことを表している。多くの経済時系列において長期記憶性が観測されるが、そのような系列を表現するモデルとして次の ARFIMA (p,d,q) モデル (Granger and Joyeux (1980), Hoskin (1981)), FIGARCH モデル (Baillie, Bollerslev and Mikkelesen (1996)), 及び FIEGARCH モデルなどがある。

ARFIMA (p,d,q) モデルの定義：

$$\Phi(L)(1-L)^d y_t = \Theta(L)\varepsilon_t \quad (4)$$

$$\Phi(L) = 1 - \phi_1 L - \phi_2 L^2 \cdots - \phi_p L^p \quad (5)$$

$$\Theta(L) = 1 - \theta_1 L - \theta_2 L^2 \cdots - \theta_q L^q \quad (6)$$

ここで L はラグオペレータ, また $\{\varepsilon_t\}$ は分散 σ^2 のホワイト・ノイズである。ラグオペレータ L を使えば $Ly_t = y_{t-1}$, $(1-L)y_t = y_t - y_{t-1}$ (1階の階差) などのようにあらわされる。 $\Phi(L)$, $\Theta(L)$ は L に関するそれぞれ p 次と q 次の多項式を表す。また d は実数差分パラメータである。一般に時系列の階差をとる場合, 階差の回数は整数であるが, $(1-L)^d$ では非整数回階差をとることを表しており, 非整数回階差とは何を意味するか分かりにくい。この点については $(1-L)^d$ 以下のように級数に展開して, さまざまな整数階差の加重平均と考えれば分かりやすい。すなわち

$$(1-L)^d = 1 - \sum_{k=0}^{\infty} \frac{d(d-1)(d-2)\cdots(d-k+1)}{k!} (-L)^k$$

と表される。ARFIMA モデルは $d < 0.5$ のとき定常な長期記憶過程, $d \geq 0.5$ のとき非定常な長期記憶過程である。また $d=0$ のときは互いに無相関なホワイトノイズ過程となる。 d が大きいほど長期依存性が高まる。なお定常な GARCH 過程は短期記憶過程である。

FIGARCH モデルの定義：

GARCH (p,q) モデルは簡単な変形によって, r_t^2 に関する ARMA (p,q) の形に表現できる。すなわち

$$\phi(L)r_t^2 = \phi_0 + (1 - \theta(L))\tau_t$$

ここに

$$\phi(L) = 1 - \phi_1 L - \phi_2 L^2 \cdots - \phi_p L^p$$

$$\theta(L) = 1 - \theta_1 L - \theta_2 L^2 \cdots - \theta_q L^q$$

である。ここでさらに実数差分オペレーター $(1-L)^d$ を作用させて一般化したモデル

$$\phi(L)(1-L)^d r_t^2 = \phi_0 + (1 - \theta(L))\tau_t$$

を fractionally integrated GARCH モデル (FIGARCH モデル) という。このモデルは, 持続性が GARCH より長く IGARCH ほど長くない過程を表すことができる。

FIEGARCH モデルの定義：

時系列 y_t が非負であるような時系列においてはモデルのパラメータに複雑な非負条件を課さなければならない。この点を回避するために（それだけが理由ではないが）、Bollerslev and Mikkelsen (1996) は次のような FIEGARCH モデルを提案した。

$$\begin{aligned} \Phi(L)(1-L)^d \{\log(y_t) - \omega\} &= g(z_{t-1}) & (7) \\ g(z_{t-1}) &= \theta z_{t-1} + r \{|z_{t-1}| - E(|z_{t-1}|)\} \\ z_t &\sim i.i.d., E(z_t) = 0, V(z_t) = 1 \end{aligned}$$

ここに右辺の関数 $g(z_{t-1})$ は前期の攪乱項 z_{t-1} の変動の次期への影響が非対称性を持つことを示している。

この他にも確率的ボラティリティモデル (stochastic volatility model, SV モデル) に長期記憶性を付加した long memory SV モデル (LMSV モデル, Harvey (1993), Breidt, Crato, and de Lima (1994), 白石・高山 (1998)) もあるが、本稿では扱わない。

2.3 実現ボラティリティ (Realized Volatility, RV)

第 t 日における対数株価収益率の日内時系列 $\{r_{tj}, j=1, 2, 3, \dots, n\}$ における日内の 2 乗和を実現ボラティリティといい RV_t で表す。すなわち RV_t は

$$RV_t = \sum_{j=1}^n r_{tj}^2$$

で定義される。ここで n が十分大きければ RV_t は第 t 日の収益率のボラティリティ σ_t^2 の精度の良い推定値となることが知られている。 σ_t^2 は直接観測できない潜在変数であるが、高頻度データが利用可能になったことによって精度の高い推定値を得ることが出来るようになったのである。

ここで、GARCH モデルから推定されるボラティリティと RV を比較してみると興味深い。次の図 2.1a は 2006 年 3 月 1 日から 2007 年 2 月 28 日までの日経平均株価の 5 分間隔高頻度データから計算した RV とこの間の日次データを用いて推定された GARCH(1,1) モデルから計算したボラティリティ σ_t^2 の推定値の系列である。日次収益率から推定された GARCH(1,1) モデルのパラメータは

$$\hat{\omega} = 2.7 \times 10^{-7}, \hat{\alpha} = 0.0298, \hat{\beta} = 0.9646$$

であった。図 2.1b は個別銘柄の事例として 2006 年 2 月 1 日から 2006 年 12 月 28 日まで

のソフトバンク社の株価高頻度データから同様の比較を行ったものである。この場合の GARCH(1,1) のパラメータ推定値は

$$\hat{\omega}=2.1 \times 10^{-7}, \hat{\alpha}=0.0257, \hat{\beta}=0.9491$$

であった。これらの図から見て取れるようにボラティリティの推定値 $\hat{\sigma}_t^2$ は RV の変動を均した様な、あるいは移動平均のような形状を示している。言い換えれば推定値 $\hat{\sigma}_t^2$ は RV の傾向的変動を、ある程度捉えているといえる。しかし大きな変動を平滑化しすぎてしまっており、真のボラティリティを過少推定している可能性がある。

しかし上の日経平均株価の例では、真のボラティリティ σ_t^2 が分からない以上、ボラティリティ σ_t^2 の推定値として RV と $\hat{\sigma}_t^2$ のどちらがより正確であるかは判断できない。そこで真のボラティリティ σ_t^2 が既知であるような状況の下でシミュレーションによって σ_t^2 , RV の 2 者を比較してみよう。なお、その際日時収益 r_t に対して

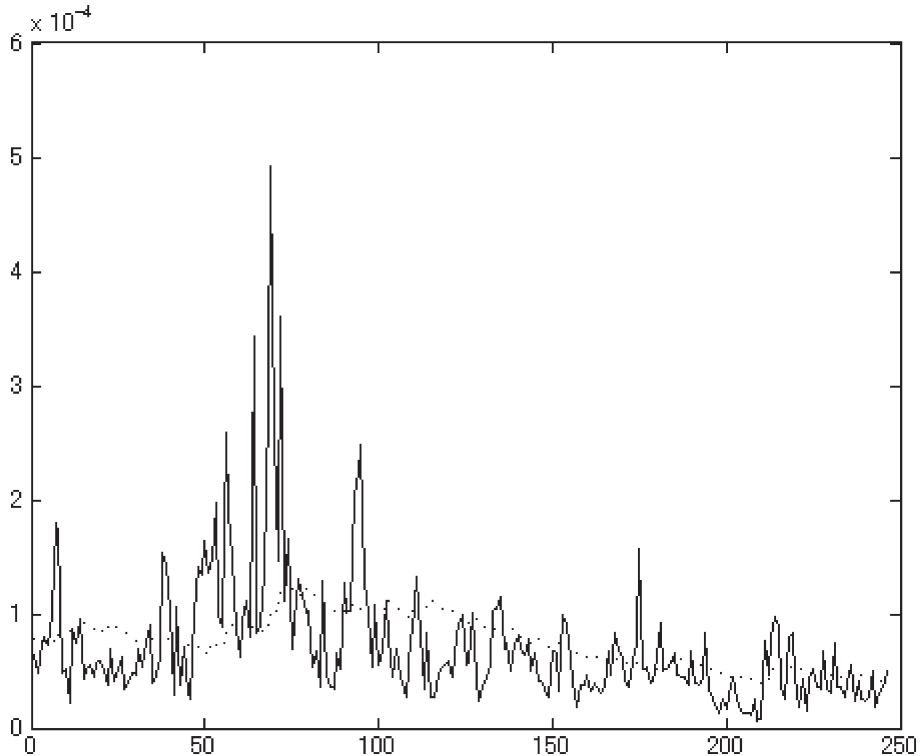


図2.1a 5分間隔の日経平均株価から計算されたRV(実線)とGARCH(1,1)ボラティリティ推定値(点線)

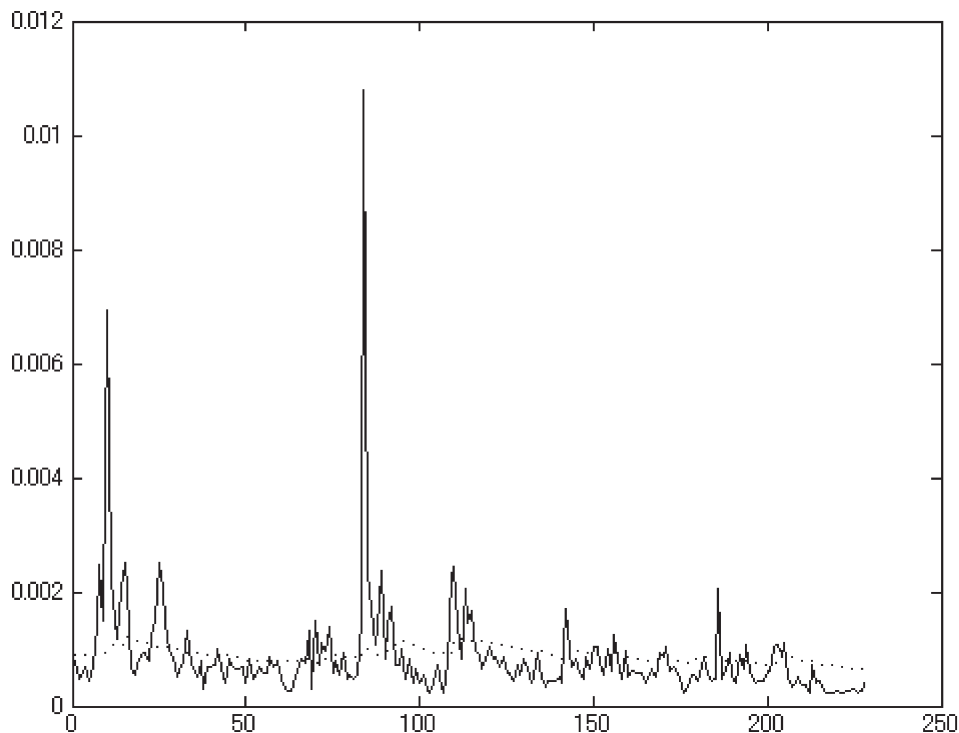


図2.1b 5分間隔のソフトバンク(証券コード9984)社の株価から計算されたRV(実線)とGARCH(1,1)ボラティリティ推定値(点線)

GARCH(1,1)から推定される日次ボラティリティ $\hat{\sigma}_t^2$ と R_t とを比較することは適切ではないので、 $\hat{\sigma}_t^2$ を計算しなかった。⁽³⁾

シミュレーションは以下のような想定と手順の下で行った。

- (1) 1分刻みの対数株価収益率の高頻度データを1000日分(60(分)×24(時間)×1000(日)=1440000時点)を人工的に発生させる。以後、第 t 日の第 j 分の1分間収益率を r_{tj} と表記する。
- (2) その際、収益率 r_{tj} の生成過程(Data Generating Process, DGP) を

$$r_{tj} = e_{tj}, \quad e_{tj} \sim N(0, \sigma_{tj}^2) \quad (8)$$

$$e_{tj}, \sqrt{\sigma_{tj}^2} \varepsilon_{tj}, \quad \varepsilon_{tj} \sim iidN(0, 1) \quad (9)$$

$$\sigma_{tj}^2 = \omega + \beta \sigma_{t,j-1}^2 + \alpha e_{t,j-1}^2 \quad (10)$$

で与えられる GARCH(1,1)過程とし、分散方程式のパラメータ値を

$$\omega=0.0001, \alpha=0.15, \beta=0.75$$

と与える。

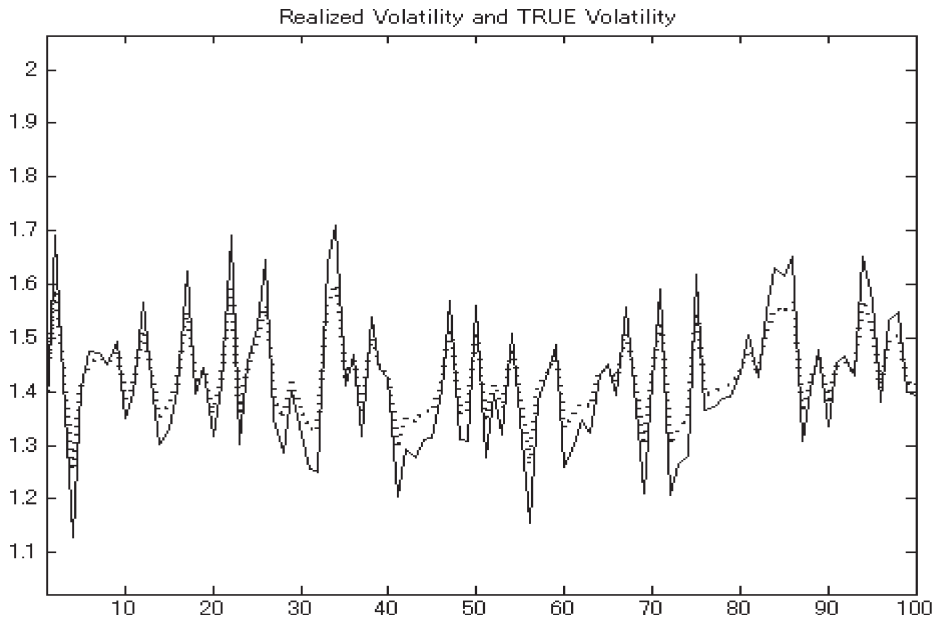


図2.2 シミュレーションによる RV_t (実線) と σ_t^2 (点線) の比較 ($t=1, \dots, 100$)

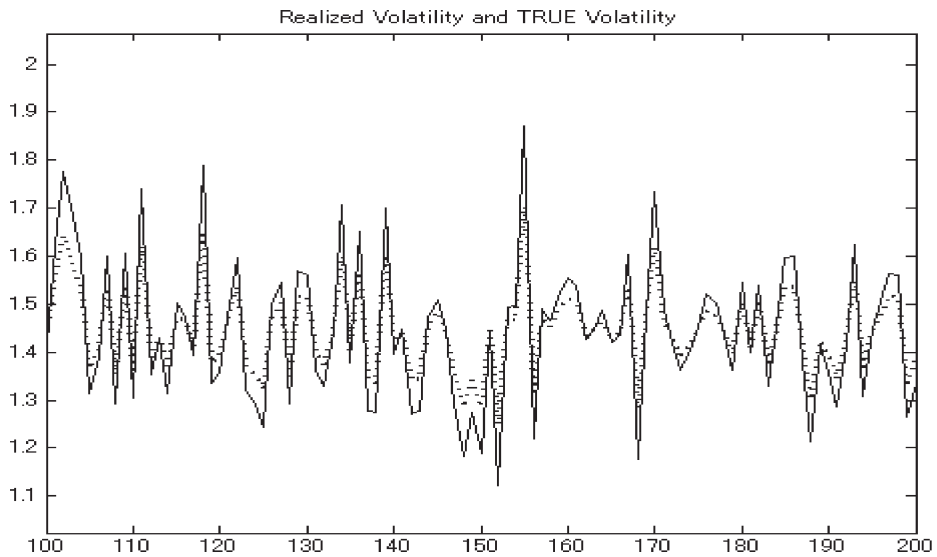


図2.3 シミュレーションによる RV_t (実線) と σ_t^2 (点線) の比較 ($t=100, \dots, 200$)

- (3) 1日分すなわち1440時点分のRVを1000日分計算する。第 t 日のRVを R_t と表記する。
- (4) 1分刻みの収益率 r_{tj} を1日分合計したものは第 t 日の収益率 r_t であるから、第 t 日のモデルの真のボラティリティは $\sigma_t^2 = \sum_{j=1}^{1440} \sigma_{tj}^2$ によって計算される。ここに σ_{tj}^2 は分散方程式(10)より計算する。

図2.2と図2.3は、以上の手順によって得られた R_t と真のボラティリティ σ_t^2 の一部を図示したものである。RVが真のボラティリティの精度の良い推定値であることが分かる。

3 東証1部上銘柄による実証分析

東洋経済新報社の2006年度版株価CD-ROMに収録されている東証1部上場の個別銘柄の収益率データに対してGARCH(1,1)モデルを当てはめて推定を行う。対象とした業種は次の通りである。

1. 水産・農林業
2. 鉱業
3. 建設業
4. 製造業：1(食品, 繊維製品, パルプ紙)
5. 製造業：2(化学, 医薬品, 石油・石炭製品, ゴム製品)
6. 製造業：3(機械, 電気機器, 輸送機器, 精密機器, その他)
7. 電気・ガス
8. 通信業
9. 小売業
10. 不動産業
11. 銀行業

使用データは2003年1月6日から2005年12月30日までの標本数737の日次データ3年間分である。但し、この3年間分のデータが存在しない銘柄あるいは、GARCH(1,1)モデルのパラメータ推定を行ったとき収束解が得られなかった銘柄は分析対象から除外した。その結果、1130銘柄が分析対象となった。これら1130の各銘柄に対してGARCH(1,1)モデルのパラメータ推定を行った。その結果得られた分散方程式

(3)式

のパラメータ推定値 $\{\hat{\alpha}_i\}_{i=1}^{1130}$ と $\{\hat{\beta}_i\}_{i=1}^{1130}$ のヒストグラムを図3.1と3.2に示す。

このとき $\{\hat{\alpha}_i\}_{i=1}^{1130}$ と $\{\hat{\beta}_i\}_{i=1}^{1130}$ の平均 $\bar{\alpha}, \bar{\beta}$ とそれらの標本分散 $V(\hat{\alpha}), V(\hat{\beta})$ はそれぞれ

$$\bar{\alpha} = \frac{1}{1130} \sum_{i=1}^{1130} \hat{\alpha}_i = 0.1522, \quad V(\hat{\alpha}) = \frac{1}{1130} \sum_{i=1}^{1130} (\hat{\alpha}_i - \bar{\alpha})^2 = 0.0201$$

$$\bar{\beta} = \frac{1}{1130} \sum_{i=1}^{1130} \hat{\beta}_i = 0.7363, \quad V(\hat{\beta}) = \frac{1}{1130} \sum_{i=1}^{1130} (\hat{\beta}_i - \bar{\beta})^2 = 0.0526$$

であった。

次に、図3.1, 図3.2の形状から分散方程式(3)式の $\hat{\alpha}, \hat{\beta}$ がそれぞれ ベータ分布に従うと仮定し、ベータ分布 $\text{Beta}(k, l)$ のパラメータ (k, l) を推定する。ここで、ベータ分布の密度関数は以下のように与えられる。

$$f_{\text{Beta}}(x) = \frac{1}{B(k, l)} x^{k-1} (1-x)^{l-1}$$

$$B(k, l) = \int_0^1 x^{k-1} (1-x)^{l-1}$$

上に得られた $\hat{\alpha}, \hat{\beta}$ それぞれの経験分布から推定されたベータ分布は

$$\hat{\alpha} \sim \text{Beta}(1.1197, 5.6328)$$

$$\hat{\beta} \sim \text{Beta}(1.8314, 0.7992)$$

であった。推定値の当てはまり具合を見るため、これらベータ分布の密度関数と

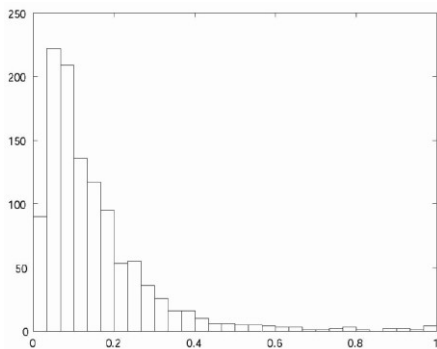


図3.1 $\{\hat{\alpha}_i\}_{i=1}^{1130}$ のヒストグラム

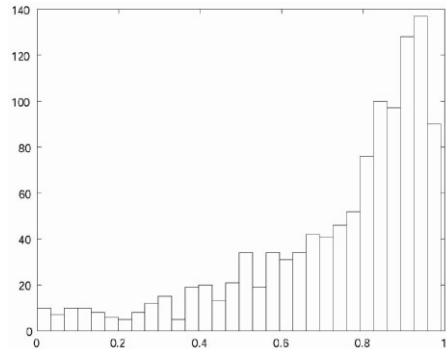


図3.2 $\{\hat{\beta}_i\}_{i=1}^{1130}$ のヒストグラム

$\{\hat{\alpha}_i\}_{i=1}^{1130}$ と $\{\hat{\beta}_i\}_{i=1}^{1130}$ の経験分布を図3.3, 図3.4に示す。

これらの図から GARCH(1,1)モデルを当てはめた結果、 $\hat{\alpha} + \hat{\beta}$ が1に近い場合が多いことを読み取ることができる。しかしながらこの推定結果では、図3.4のグラフの右端の曲線の下降部分がうまく当てはまっていない。そこで右端が下降するような曲線をパソコン画面を見ながら手動で当てはめてみたものが、図3.5, 図3.6である。このとき設定したベータ分布のパラメータは $\hat{\alpha} \sim \text{Beta}(5.4, 30)$, $\hat{\beta} \sim \text{Beta}(6, 1.6)$ である。

われわれは上記の1130銘柄について d を推定し、 $\hat{\alpha} + \hat{\beta}$ と \hat{d} の散布図(図3.7)と \hat{d} のヒストグラム(図3.8)を描いた。 d の推定は対数ペリオドグラム回帰法(たとえば矢島(2003)を見よ)によっておこなった。

次に上で得られた東証1部上場1130銘柄の GARCH(1,1)モデルのパラメータを用いて(3)式の分散方程式から1130個の分散系列をそれぞれ計算し、それらの分散系列を時点 $t=1, 2, \dots, 737$ の各時点ごとに平均し、新たな系列を計算する。この新たな平均化された系列に対して自己相関係数を計算してみた。その結果が図3.9に示されている。この図から、自己相関係数は時間が経過してもなかなか0に収束しないという長期記憶性の特徴が読み取られる。

以上の図にあらわれた特徴は以下のように要約することが出来る。

- (1) $\hat{\alpha} + \hat{\beta}$ が0.8以下では \hat{d} の値は0近辺にとどまっている。
- (2) $\hat{\alpha} + \hat{\beta}$ が0.8を越えるあたりから \hat{d} の値が徐々に上昇し0.9を越えるあたりから急上昇する。
- (3) $\hat{\alpha} + \hat{\beta}$ が1に極めて近いところではほとんどの \hat{d} が0.5を超える。
- (4) $\hat{\alpha} + \hat{\beta}$ が1に近く、また \hat{d} が0.5を超える銘柄が多く観測される。このことは非正常長期記憶過程と判定される銘柄が多いことを意味する。
- (5) 自己相関係数の下がり方が緩やか。

4. シミュレーション分析

前節の実証分析から、大部分の東証1部上場銘柄の収益率に関しては GARCH(1,1)モデルによってよく記述され、しかも係数 $\hat{\alpha}$, $\hat{\beta}$ はベータ分布に従っている様子が確認できた。このような現実を踏まえた上で、本節では以下のようなシミュレーションを行う。

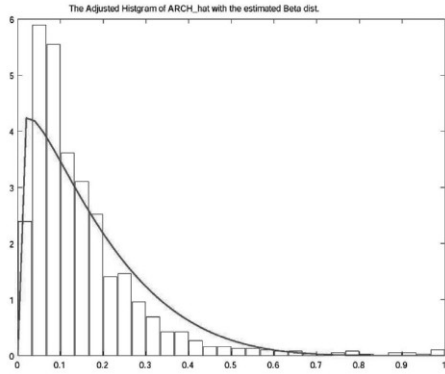


図3.3 $\{\hat{a}_i\}_{i=1}^{1130}$ のヒストグラムと密度関数

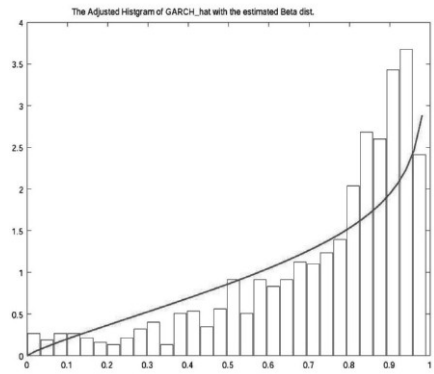


図3.4 $\{\hat{\beta}_i\}_{i=1}^{1130}$ のヒストグラムと密度関数

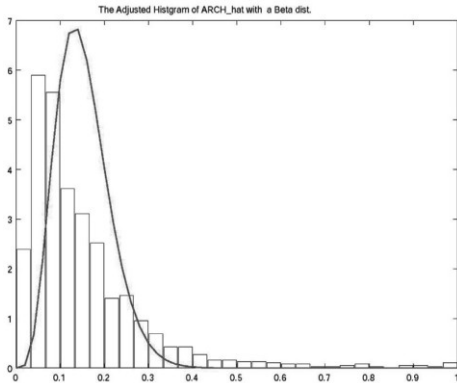


図3.5 $\{\hat{a}_i\}_{i=1}^{1130}$ のヒストグラムと密度関数

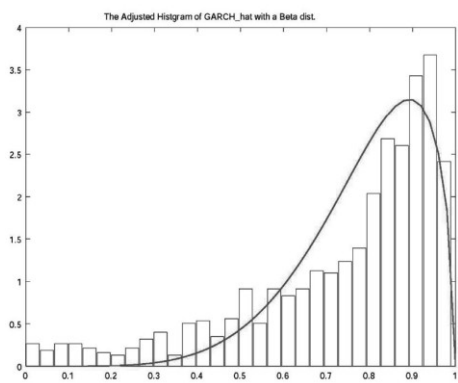


図3.6 $\{\hat{\beta}_i\}_{i=1}^{1130}$ のヒストグラムと密度関数

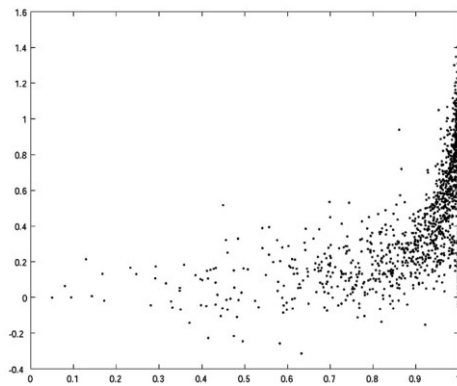


図3.7 東証1部上場1130銘柄の
 $(\hat{a}_i + \hat{\beta}_i, \hat{d}_i)$ の散布図

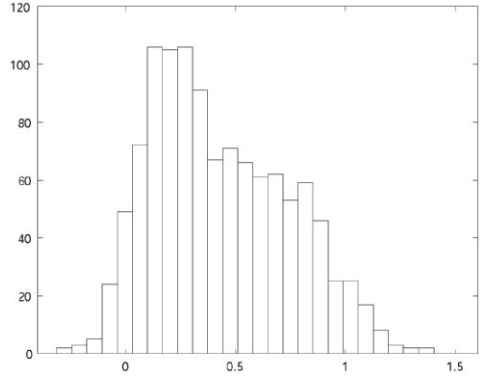


図3.8 東証1部上場1130銘柄の
 $\{\hat{d}_i\}_{i=1}^{1000}$ のヒストグラム

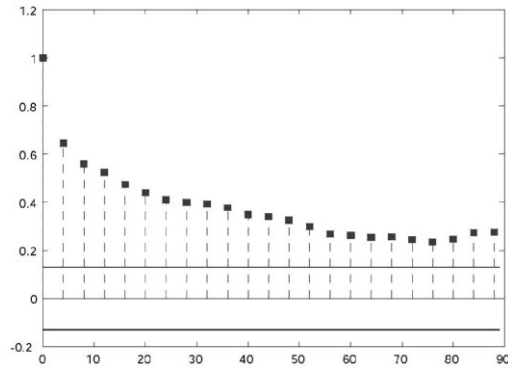


図3.9 東証1部上場1130銘柄の分散系列の平均に対する自己相関係数

分散方程式(3)式の α と β をそれぞれベータ分布から発生させ、GARCH(1,1)プロセスを生成する。シミュレーション手順として、まず分散方程式(3)式の α と β のペア (α, β) を、それぞれをベータ分布から1000組発生させ、系列の長さが1000時点のGARCH(1,1)プロセスを1000系列生成する。この1000系列は東証1部上場銘柄の動きを模した仮想現実と考えられる。次に、生成した1000系列のそれぞれにGARCH(1,1)モデルを当てはめモデルパラメータの推定を行う。最後に、長期記憶性を示す d を推定する。

図4.1には α と β の推定値の和 (横軸) と d の推定値 (縦軸) の組み合わせ

$((\hat{\alpha}_i + \hat{\beta}_i), \hat{d}_i)$ の散布図を、図4.2には $\{\hat{d}_i\}_{i=1}^{1000}$ のヒストグラムを示す。

これら1000個の \hat{d} の標本平均 $\bar{\hat{d}}$ 、標本分散 $V(\hat{d})$ 、歪度、尖度、メディアンは次の

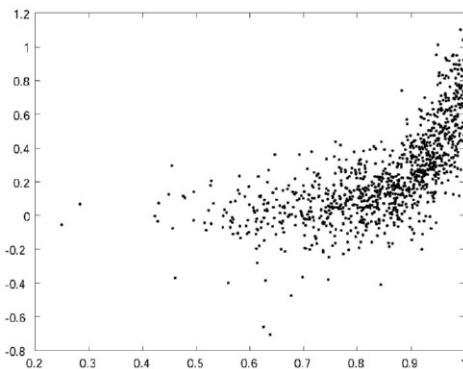


図4.1 $((\hat{\alpha}_i + \hat{\beta}_i), \hat{d}_i)$ の散布図

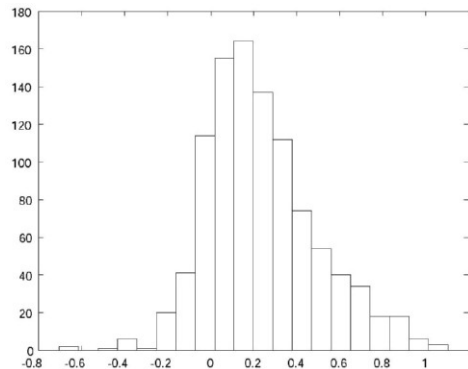


図4.2 $\{\hat{d}_i\}_{i=1}^{1000}$ のヒストグラム

通りであった。

$$\begin{aligned} \bar{d} &= 0.2346, & V(\hat{d}) &= 0.0663 \\ \text{Skewness}(\hat{d}) &= 0.5516, & \text{Kurtosis}(\hat{d}) &= 3.5025, & \text{Median}(\hat{d}) &= 0.1932 \end{aligned}$$

図4.3, 図4.4の形状は東証1部上場銘柄から計算結果と酷似しており, 上記の特徴(1)~(4)はシミュレーション結果にもほとんど当てはまる。

これらの観察結果から, われわれのシミュレーションは現実をよく反映しているといえる。

次に, 現実の推定結果では非常に多くの銘柄において $\alpha + \beta \approx 0.9$ であることを考慮して, ベータ分布から α, β を生成する際に $\alpha + \beta \approx 0.9$ となるように設定して推定を行ってみた。その結果についても $((\hat{\alpha}_i + \hat{\beta}_i), \hat{d}_i)$ の散布図と $\{\hat{d}_i\}_{i=1}^{1000}$ のヒストグラムを図4.3, 4.4に示す。また, 各種記述統計量は

$$\begin{aligned} \bar{d} &= 0.4411, & V(\hat{d}) &= 0.0664 \\ \text{Skewness}(\hat{d}) &= 0.2606, & \text{Kurtosis}(\hat{d}) &= 3.1785, & \text{Median}(\hat{d}) &= 0.4226 \end{aligned}$$

であった。

5. まとめと今後の課題

以上のシミュレーションの結果は, GARCH において $\bar{\alpha} + \bar{\beta}$ が1に近いときは, それが短期定常時系列であるにもかかわらず, 0以上の d の推定値が得られ, 長期記憶

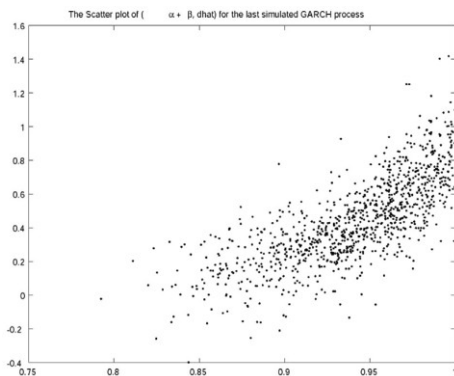


図4.3 $((\hat{\alpha}_i + \hat{\beta}_i), \hat{d}_i)$ の散布図

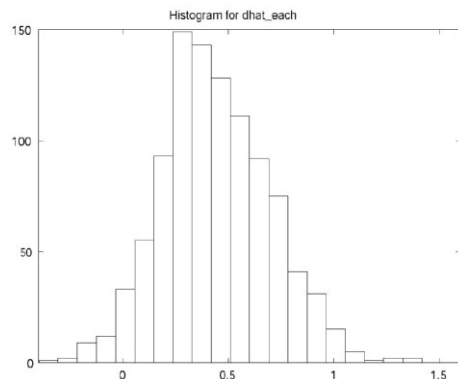


図4.4 $\{\hat{d}_i\}_{i=1}^{1000}$ のヒストグラム

定常時系列と区別することが難しいばかりか、長期記憶非定常時系列として推定されてしまう場合が多いことが示された。このことは、現実には長期記憶系列であるにもかかわらず、GARCHを適用したため $\hat{\alpha} + \hat{\beta}$ が1に近いモデルが推定されてしまうのか、または逆に現実には $\hat{\alpha} + \hat{\beta}$ が1に近い短期記憶定常な GARCH(1,1)過程であるにもかかわらず、長期記憶系列をフィットさせると、見かけ上の長期記憶系列が検出されるのか、そのどちらが正しいのかをデータから識別することが非常に困難であることを示していると考えられる。

今後、両者の精度の良い識別方法を考案することが課題として残されている。さらに両者を2社択一的なモデルとして捉えるのではなく、Granger(1980)が最も簡単な短期記憶定常時系列である AR(1)過程の係数がベータ分布に従うという仮定の下で示した、「多数の AR(1)過程の平均が長期記憶性を持つ」という結果に対応するような関係が GARCH モデルについても成立するのではないかを今後検討したい。われわれは「短期記憶定常過程である GARCH(1,1)の係数 α と β にベータ分布を仮定することによって短期記憶定常 GARCH(1,1)過程の平均が長期記憶性を持つ」という conjecture を抱いており、現在検討中である。

注

- (1) 本研究は日本学術振興会平成19年度科学研究費補助金（基盤研究(B)研究課題（課題番号：18330041）：高頻度データによる株価・為替レートの計量ファイナンス分析）の補助を受けている。
- (2) 本稿を作成するに当たり、得津康義、高石哲弥（広島経済大学）、Tee Kian Heng（岩手県立大学）、森本孝之（一橋大学）、永田修一（広島大学大学院社会科学研究所博士課程）の諸氏から貴重なコメントを頂いたことをここに感謝する。
- (3) 第 t 日の1分刻みの収益率 r_{tj} を1日分合計した第 t 日の収益率 r_t の分散 σ_t^2 は簡単な計算により

$$\sigma_t^2 = \sum_{j=1}^n \sigma_{tj}^2 = \sum (\omega + \beta \sigma_{t,j-1}^2 + a e_{t,j-1}^2) = n\omega + \beta \sum \sigma_{t,j-1}^2 + a \sum e_{t,j-1}^2$$

となる。すなわち

$$\sigma_t^2 \neq \omega + \beta \sigma_{t-1}^2 + a e_{t-1}^2$$

であるから日次収益率の系列 $\{r_t, t=1, 2, 3, \dots, 1000\}$ はもはや GARCH(1,1)に従わない。したがって日次収益率の系列に GARCH(1,1)を当てはめて第 t 日のボラティリティの推定値 $\hat{\sigma}_t^2$ を計算しても正しい推定結果は得られない。

参考文献

Baillie, R. T., T. Bollerslev and H. O. Mikkelesen(1996) Fractionary Integrated

- Generalized Autoregressive Conditional Heteroscedasticity, *Journal of Econometrics*, 74, pp.3-30.
- Bollerslev, T. (1986) Generalized Autoregressive Conditional Heteroscedasticity, *Journal of Econometrics*, 31, 307-327
- Bollerslev, T., and H. O. Mikkelsen, (1996) Modeling and Pricing Long memory in Stock Market Volatility, *Journal of Econometrics*, 73, pp.151-184
- Breidt, F.J., N. Crato and P. de Lima, (1994) On the Detection and Estimation of Long Memory in Stochastic Volatility, *Journal of Econometrics*, 83 pp.325-348
- Engle, Robert F. (1982) Autoregressive Conditional Heteroscedasticity with Estimates of Variance of United Kingdom Inflation, *Econometrica*, Vol.50 pp.987-1007
- Harvy, A.C. (1993) Long memory and Stochastic Volatility, mimeo, Statistics Department, London School of Economics.
- Hayashi, T. and N. Yoshida (2007) Asymptotic normality of a covariance estimator for nonsynchronously observed diffusion processes, *Annals of the Institute of Statistical Mathematics*
- Hosking, J. R. M. (1981) Fractional Differencing, *Biometrika*, 68, pp.165-176
- Granger, C.W.j. (1980): Long memory relationships and the aggregation of dynamic models. *Journal of Econometrics*, 14, 227-238.
- 白石典義・高山俊則 (1998) 株式収益率ボラティリティの長期依存性とロングメモリー・モデル, *ジャファイア・ジャーナル* [1998] リスク管理と金融・証券投資戦略, 東洋経済新報社
- 矢島美寛 (2003) 長期記憶を持つ時系列モデル, *統計科学のフロンティア* 8・経済時系列の統計第II部 岩波書店
- 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* Vol.58, No.4, Oct.2007

