

ファイナンス時系列における長期記憶系列の 発生メカニズムについて*

前 川 功 一**

はじめに

いわゆる長期記憶性または長期依存性と呼ばれる現象を最初に指摘した研究は水門学の分野における Hurst (1951) のナイル川の最小水位データの解析である。Mandelbrot and Hudson (2009) の第9章に Hurst の短い伝記的紹介とエジプトでの研究と生活が紹介されている。後述する Hurst 指数に彼の名が残されている。その後、多くの自然現象や社会現象のデータにも長期記憶性の存在が指摘され、現在では統計学の一つの分野が形成されており、非常に多くの理論的及び実証的研究がなされている。さらに最近では、計量ファイナンス、中でもファイナンス時系列の分野においても長期記憶性に関する研究が活発に行われている。本論では先行研究及び筆者と共同研究者らとの研究を引用しつつこの分野の現状を、特に長期記憶性発生メカニズムを中心に紹介する。本論は次のような構成になっている。まず第1節で時系列分析における長期記憶性の定義を述べる。第2節では長期記憶性を表す2, 3のモデルを紹介し、第3節では、東証第1部上場個別銘柄の株価から長期記憶性パラメータ d (後述) の推定結果を示す。第4節は、長期記憶性の発生メカニズムに関する Granger (1980) の論文を紹介したのち、その拡張を試みる。最後に第5節では、モデルの想定誤差や構造変化によってもたらされる見せかけの長期記憶性について考察する。

* 本論は2009年12月23, 24日に明治大学で開催された日本金融・証券計量・工学会 (JAFEE) 2009冬季大会で筆者が行った会長講演の一部に加筆修正したものである。また同年12月26日開催の広島経済大学共同研究費による研究集会における筆者の同内容の報告に対して一橋大学田中勝人教授ら参加者から有益なコメントをいただいたことに感謝する。またここに共同研究者の得津康義 (広島経済大学)、河合研一 (別府大学)、森本孝之 (関西学院大学)、永田修一 (関西学院大学)、Lu Xinhong (中国人民銀行金融研究所) の諸氏に共同研究成果の一部を掲載することを了承いただいたことに謝意を表す。

** 広島経済大学大学院経済学研究科教授

1. 長期記憶性の定義

通常, 時系列 $\{x_t\}$ において, x_t と x_{t-h} の自己相関係数 $\rho(h)$ (または自己共分散 $\gamma(h)$) は時差 h の増加にともない 0 へ収束する。この収束の速度が非常に緩慢な時系列を長期記憶系列という。より正確にはある時系列 $\{x_t\}$ において自己相関係数を $\rho(h)$ が

$$\sum_{h=0}^{\infty} |\rho(h)| = \infty \quad (1.1)$$

となるとき, この系列を長期記憶時系列という。他方

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

となる系列を短期記憶時系列という。この他にもスペクトル密度関数 $f(\omega)$ が

$$f(\omega) \sim |\omega|^{-2d} \quad (1.2)$$

を満たす場合, または自己相関係数が

$$\rho(h) \sim h^{2d-1} \quad (1.3)$$

となる場合を長期記憶時系列という場合もある。あるいは時系列 $\{x_t\}$ がホワイトノイズ ε_t の線形過程

$$x_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} \quad (1.4)$$

または

$$\varepsilon_t = \sum_{j=0}^{\infty} \pi_j x_{t-j}$$

と表されるとき, これらの係数が

$$\psi_j \sim j^{d-1}, \text{ または } \pi_j \sim j^{-d-1}$$

となる場合を長期記憶性の定義とすることもある。これらの定義は必ずしも同値ではないが, Wilfredo Palma (2007) はこれらの定義をやや一般的に次のように定義し, 定義間の関係を示している。4つの定義とは

1. $\sum_{h=-\infty}^{\infty} |\gamma(h)| = \infty$
2. $\gamma(h) \sim h^{2d-1} l_1(h)$
3. $f(\omega) \sim |\omega|^{-2d} l_2(1/|\omega|)$

4. $\psi_j \sim j^{d-1} l_3(j)$

ここで $l_k(\cdot)$ は $c > 0$ に対して $x \rightarrow \infty$ のとき $l(cx) / l(x) \rightarrow 1$ となる緩慢に変化する関数をあらわす。

そして1ならば2が成立すること, また2ならば1と3が成立することを示している。さらに部分和 $S_T = \sum_{t=1}^T x_t$ を用いて

5. $\text{Var}(S_T) = O(T^{2d+1})$ を定義とするものもある (Barndorff-Nielsen and Cox, 1989)。

長期記憶性に関する統計理論と経済やファイナンス及びそれら以外の分野への応用研究成果の紹介などについては, 田中 (2006) や矢島 (2003) にすぐれた入門的解説がある。

2. 長期記憶過程モデル

長期記憶性を表現するモデルとして ARFIMA (p, d, q) モデル, FIGARCH モデル, FIEGARCH モデルなどがある。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)) もあるが, ここでは扱わない。ここでは時系列モデルの ARFIMA (p, d, q) モデルと確率過程モデルとしての FGN モデルのみを簡単に述べる。

ARFIMA (p, d, q) モデルは次の (2.1) 式のように表現される。

$$\Phi(L)(1-L)^d y_t = \Theta(L)\varepsilon_t \tag{2.1}$$

ここに

$$\begin{aligned} \Phi(L) &= 1 - \phi_1 L - \phi_2 L^2 \dots - \phi_p L^p \\ \Theta(L) &= 1 - \theta_1 L - \theta_2 L^2 \dots - \theta_q L^q \end{aligned}$$

である。

ここで L はラグオペレータ, また $\{\varepsilon_t\}$ は平均0, 分散 σ^2 のホワイトノイズである。 $\Phi(L)$ と $\Theta(L)$ 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 は非整数差分パラメータである。この d を長期記憶パラメータと呼ぶ。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 $(1-L)^d$ を以下のように級数に展開して, さまざまなラグの加重平均と考えれば分かりやすい。すなわち

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

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。またの $d=0$ 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 ARFIMA モデルは FARIMA モデルと呼ばれることもある。

長期記憶過程を表す確率過程モデルとして FGN (Fractional Gaussian Noise) と呼ばれる次のようなモデルがある。確率過程 $\{B_H(t)\}$ が期待値 0 の正規過程に従い、任意の連続時間 $0 \leq t, s < \infty$ に対してある数 $H(0 < H < 1)$ が存在し

$$E(|B_H(t) - B_H(s)|^2) = \sigma^2 |t-s|^{2H}$$

を満たすとき, この過程をフラクショナル・ブラウン運動という。 $B_H(t)$ の 1 階の階差を

$$X_t = B_H(t) - B_H(t-1), \quad t = 1, 2, \dots$$

と置くととき, X_t を FGN モデルという。 X_t の自己相関関数は $H \neq \frac{1}{2}$ のとき

$$\rho(h) \sim h^{2H-1}$$

となることが示される。ここで $d = H - \frac{1}{2}$ と置けば自己相関関数は (1.3) と同じ形に書け, X_t が長期記憶性を持つことが分かる。 H は Hurst 指数と呼ばれる。

3. 長期記憶系列: 2, 3 の例

この節では株価収益率の高頻度データから計算された Realized Volatility (RV) における長期記憶時パラメータ推定の 2, 3 の例を示す。

(1) 東証株価 (個別銘柄) における RV の実証例: (得津・永田・前川 (2007))

得津・永田・前川 (2007) は 2006 年 3 月 31 日から 2007 年 2 月 28 日までの東証 1 部上場 1,625 銘柄の 5 分足収益率から計算された各個別銘柄の RV を用いて, ARFIMA (0, d , 0) モデルのパラメータ d を疑似最尤法とウェーブレット法を用いて推定を行った。この節では疑似最尤法による結果のみを示す。ただし RV の計算に当たっては Hansen and Lunde (2005) による修正 RV が用いられている。その結果を図 3.1 と 3.2 に示す。推定値 \hat{d} の半数近くは 1/2 を超えており, この結果は採用された銘柄

の半数近くは非定常長期記憶過程であることを示唆している。非定常の長期記憶過程における \hat{d} の分布の漸近正規性は保証されていないものの、図1.1からほぼ正規性が成り立つことが予想される。仮に正規性を仮定して、対象銘柄すべてについて帰無仮説： $d = 0$ を検定すると、すべての銘柄において有意水準5%でこの帰無仮説は棄却された。

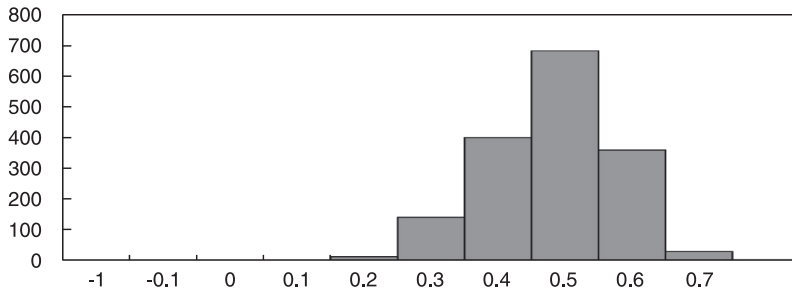
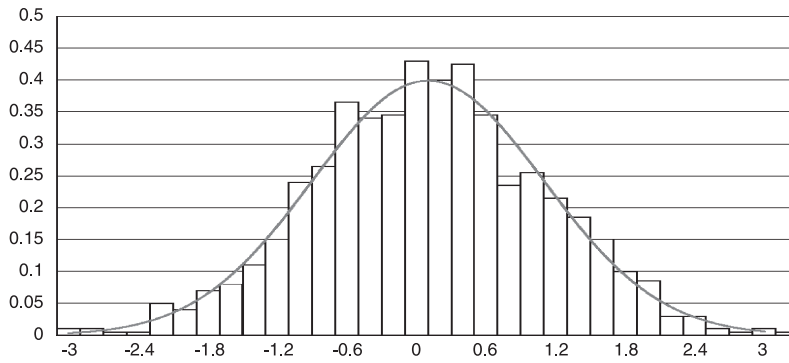


図3.1 MLEによる東証1部1,625銘柄のRVにおける \hat{d} のヒストグラム

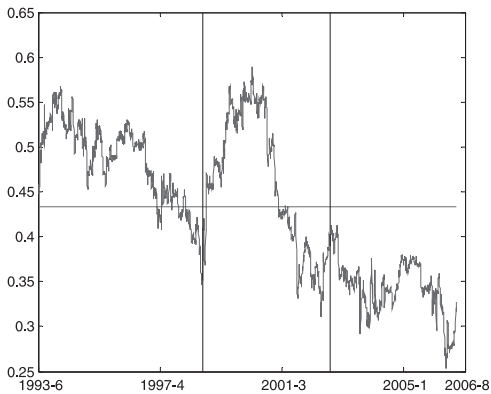
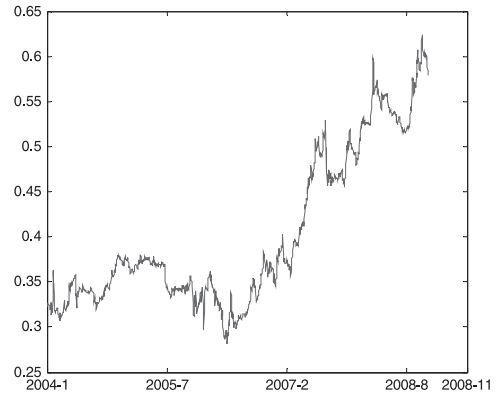


(ヒストグラムはシミュレーション結果。実線は $N(0, 1)$ の密度関数)

図3.2 $d = 0$ の下での標準化された疑似最尤推定量の分布

(2) 長期記憶パラメータ d の不安定性 (為替レートの場合: Maekawa and Lu (2009))

Maekawa and Lu (2009) は、ロシア経済危機前後の期間における為替収益率の高頻度データから計算されたRVの長期記憶性を検証した。図3.3はローリング・ウィンドウ法で推定した長期記憶パラメータ推定値 \hat{d} の変遷である。図3.3の中央部分がロシア経済危機の期間であるが、 \hat{d} の値が大きく跳ね上がっている。図3.4はローリング・ウィンドウ法で推定された2008年の世界金融危機直前から2008年末までの \hat{d} の推移である。やはり金融危機以降の期間で \hat{d} が大きく変動している。特に2008年夏以降の \hat{d} が0.5を超えていることは、この時期は非定常長期記憶過程で

図3.3 ロシア経済危機前後の \hat{d} の推移図3.4 世界金融危機と \hat{d} の推移

あった可能性を示している。図3.3, 3.4から、大きな経済的な変動が d の値に大きな影響を与えることを見取ることができる。

4. 長期記憶性の発生メカニズム

ファイナンス時系列だけではなく、他の分野においても長期記憶性が観察されている。しかしそれぞれの分野において長期記憶性が発生する理由が個別分野の科学的理論によって十分説明されているわけではない。ただし長期記憶性発生のメカニズムが統計理論的に説明されている場合がいくつかある。一つは Granger (1980) が示した集計された経済時系列の場合。他の一つは構内情報通信網 (Local area network: LAN) におけるパケット通信に関するオン・オフデータの場合である。後者については本報告では触れないが、例えば矢島 (2003) に簡潔な説明がある。

4.1 Cross sectional aggregation : Granger (1980) の要約

N 本の時系列 $\{x_{tj}\}_{t=1, \dots, T} : j = 1, 2, \dots, N$ が次のような AR(1) モデルに従っているとす。

$$\begin{aligned} x_{tj} &= \alpha_j x_{t-1, j} + \varepsilon_{tj}, \quad j = 1, 2, \dots, N \\ \varepsilon_{tj} &\sim iid(0, \sigma_\varepsilon^2), \quad \varepsilon_{tj} \text{ と } \alpha_j \text{ は独立} \end{aligned} \quad (4.1)$$

このとき x_{jt} のスペクトラムは次式で与えられる。

$$f_j(\omega) = \frac{\text{var}(\varepsilon_{tj})}{2\pi |1 - \alpha_j z|^2}, \quad z = e^{-i\omega}. \quad (4.2)$$

ここで、 α_j ($j = 1, 2, \dots, N$) のクロスセクショナルな分布は次の「特殊」なベータ分布

$$dF(\alpha) = \frac{2}{B(p, q)} \alpha^{2p-1} (1-\alpha^2)^{q-1} d\alpha, \quad 0 \leq \alpha \leq 1, \quad (4.3)$$

に従うと仮定する。このとき集計量（平均） $\bar{x}_t = \frac{1}{N} \sum_{j=1}^N x_{tj}$ のスペクトラムは

$$\bar{f}(\omega) = \frac{1}{N} \sum_{j=1}^N f_j(\omega) \equiv \frac{1}{2\pi} E[\text{var}(\varepsilon_{jt})] \int \frac{1}{|1-\alpha z|^2} dF(\alpha) \quad (4.3)$$

で与えられる。ここでクロスセクションに関する算術平均を α のクロスセクションにおける分布の期待値に置き換えている。Granger (1980) は (4.3) を特殊なベータ分布と述べているが、実はこの分布は、 $\xi_j \sim \text{Beta}(p, q)$ のときの $\alpha_j = \xi_j^{1/2}$ の分布を表している。すなわち 2 乗するとベータ分布になるような確率変数の確率密度関数を表している。Granger 自身はこの仮定を、数式展開を容易にするための単なる便宜的な仮定であると述べているが、実はこの仮定はすぐ後に見るように、現実のデータから見て不自然な仮定ではない。

集計されたスペクトル密度 $\bar{f}(\omega)$ に含まれる z^k の係数は \bar{x}_t の k 次の共分散 $\gamma_{\bar{x}}(k)$ である。 α のクロスセクションにおける分布の期待値を取ることによって、それを求めると

$$\gamma_{\bar{x}}(k) = \frac{2}{B(p, q)} \int_0^1 \alpha^{2p+k-1} (1-\alpha^2)^{q-2} d\alpha \equiv Ak^{1-q}$$

となる。ここで右辺の近似は、スターリングの公式から $\frac{\Gamma(k+a)}{\Gamma(k+b)} \sim k^{a-b}$ となることを用いて求められる。ここで長期記憶パラメータ d を使えば、(1.3) 式より $2d-1 = 1-q$ であるから $d = 1 - \frac{q}{2}$ となる。 $q > 1$ のとき $d < 1/2$ となる。以上より $\bar{x}_t \sim I(1 - \frac{q}{2})$ を得る。

なお、(4.3) のような特殊なベータ分布ではなく普通のベータ分布を仮定しても同じ結果が得られる。附録 1 に特殊なベータ分布を用いた場合を、附録 2 に普通のベータ分布を用いた場合の導出を詳細に示しておく。

4.2 Granger (1980) の ARCH(1) への拡張 (Kawai and Maekawa (2010))

次に Granger (1980) を ARCH(1) の場合に拡張する。 x_t が ARCH(1) に従うとき x_t^2 の系列は AR(1) に従うので、 x_t^2 の AR(1) 過程に Granger (1980) の方法を応用することが考えられる。計量ファイナンスでは x_t^2 から計算される RV に長期記憶

性があることが指摘されているので、 x_t^2 の長期記憶性を考察することは意義がある。

そこで次のような状況を考えよう。個別銘柄の収益率 $\{x_{tj}\}_{t=1, \dots, T}$, $j = 1, 2, \dots, N$, が ARCH(1) に従うとする。すなわち

$$x_{tj} = \sigma_{tj} \varepsilon_{tj} \\ \sigma_{tj}^2 = \alpha_{0j} + \alpha_{1j} x_{t-1, j}^2, \quad \{\varepsilon_{tj}\} \sim NID(0, 1)$$

このとき x_{tj}^2 は次のような AR(1) の形に書くことができる。

$$x_{tj}^2 = \alpha_{0j} + \alpha_{1j} x_{t-1, j}^2 + u_{tj}, \quad u_{tj} = x_{tj}^2 - \sigma_{tj}^2$$

x_{tj}^2 は必ずしも定常ではないが、 $\alpha_{1j} < \frac{1}{\sqrt{3}} = 0.557 \dots$ のとき x_{tj}^2 は定常であることが知られている。また $v_{tj} = x_{tj}^2 - E(x_{tj}^2)$ とおけば

$$v_{tj} = \alpha_{1j} v_{t-1, j} + u_{tj}$$

と書ける。 x_{tj}^2 の定常性を仮定すれば v_{tj} のスペクトル密度関数は

$$f_j(\omega) = \frac{\sigma_u^2}{2\pi} \frac{1}{|1 - \alpha_{1j} z|^2}, \quad \text{ここに } \sigma_u^2 = \frac{1 + \alpha_{1j}}{1 - \alpha_{1j}} \frac{2\alpha_{0j}^2}{1 - 3\alpha_{1j}^2}$$

したがって v_{tj} の集計量 (平均) \bar{v}_t に関しても、 α_{1j} にベータ分布を仮定して Granger (1980) とパラレルな導出方法により

$$\bar{v}_t \sim I(1 - \frac{q}{2})$$

が得られる。この結果は x_{tj}^2 が定常であれば、 \bar{x}_t^2 が長期記憶性を持つこと意味する。では実際のデータを見てみよう。

4.3 東証株価収益率データの分析

ここでは以下のデータを用いる。

- データ期間：2003年1月6日から2005年12月30日までの3年間分の東証1部上場の株価データ (価格ベース)
- 標本数：736 (=736日)
- 対象となった銘柄数：960銘柄 (データ期間の全ての時点において観測値が得られ、かつ推定値が収束した銘柄)

このデータを用いて次の手順で計算を行った。まず、個別銘柄の株価から収益率

を計算する。次に、収益率系列に対して上記 ARCH(1) モデルを当てはめ、パラメータ α_{1j} の推定値 $\hat{\alpha}_{1j}, j = 1, 2, \dots, 960$ の経験分布をヒストグラムで示し、その分布にベータ分布を当てはめた。その結果を図4.1に示す。図4.1の中で、実線は $\hat{\alpha}_{1j}$ の経験分布から推定されたベータ分布 Beta(1.7884, 5.5670) である。

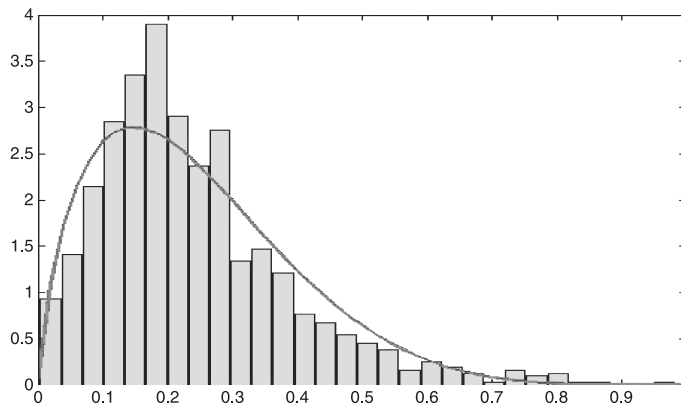


図4.1 東証1部上場銘柄から得られた $\hat{\alpha}_{1j}$ 分布

収益率の2乗の集計系列に関する長期記憶性検定

株価収益率 r_{jt} の2乗系列の j に関する集計系列 $\overline{r_t^2}$ の長期記憶性を Gewek and Porte-Hudak (1983) の対数回帰法によって検証した。対数回帰法を用いる際に使用する周波数帯の上限と下限は、 $n =$ 系列の長さ、 $l = n^a, m = n^b$ とするとき、上限 $\lambda_l = 2\pi l/n$ と下限 $\lambda_m = 2\pi m/n$ で与えられる (矢島 (2003))。ここでは $a = 0$ とし、 b をいろいろ変えて推定した。その結果を表4.1に示す。

表4.1 $\overline{r_t^2}$ 系列に対する長期記憶性の検定

a	b	\hat{d}
0.0	0.50	0.440
0.0	0.60	0.426
0.0	0.70	0.479
0.0	0.75	0.348

以上の推定結果から、

1) 図4.1より、ARCH(1) モデルのパラメータ推定値 $\hat{\alpha}_{1j}, j = 1, 2, \dots, 960$ の分布はベータ分布でうまく近似できる。

2) 対数回帰法による検定結果 (表1) は a, b に依存するが、株価収益率 $r_{j,t}$ の2乗系列の j に関する平均系列 $\overline{r_t^2}$ はほぼ長期記憶性を持つことが示された。

なお個別株価収益率の2乗 r_{jt}^2 の中には、ARCH(1) の係数推定値が定常性の条件

$\alpha_1 < 0.577$ を満たさないものが若干あるので、この条件を満たす銘柄に限定して平均 $\overline{r_t^2}$ の長期記憶性を調べたが、結果は表4.1とほとんど変わらなかった。

4.4 Granger (1980) の GARCH (1, 1) への拡張 (Kawai and Maekawa (2010))

Kawai and Maekawa (2010) は Granger (1980) の AR(1) の場合を GARCH(1, 1) に拡張した。彼らが考察の対象とした状況は次のようなものである。個別銘柄収益率 $x_{tj}, j = 1, 2, \dots, N$, が GARCH(1, 1) 過程に従うとする。すなわち

$$\begin{aligned} x_{tj} &= \sigma_{tj} \varepsilon_{tj}, \quad \varepsilon_{tj} \sim NID(0, 1) \\ \sigma_{tj}^2 &= \omega_j + \beta_j \sigma_{t-1, j}^2 + \alpha_j x_{t-1, j}^2 \end{aligned}$$

このとき x_{tj}^2 のクロスセクション集計値 (平均) $\overline{x_t^2}$ の長期記憶性を導きたい。そのために Granger (1980) のアイデアを発展させて、 α_j と β_j ($\phi_j = \alpha_j + \beta_j < 1$) は次のような確率密度関数 $f(x, y)$ を持つ 2 変量ベータ分布に従うと仮定する。

$$f(x, y) = \frac{\Gamma(\theta_1 + \theta_2 + \theta_3)}{\Gamma(\theta_1)\Gamma(\theta_2)\Gamma(\theta_3)} x^{\theta_1-1} y^{\theta_2-1} (1-x-y)^{\theta_3-1}$$

ここで $\theta_i > 0, i = 1, 2, 3, x, y \geq 0, x + y < 1$ 。この分布の原点の周りのモーメント $\mu'_{r, s} = E(x^r y^s)$ は以下のように与えられる。

$$\mu'_{r, s} = \frac{\Gamma(\theta_1 + r)\Gamma(\theta_2 + s)\Gamma(\theta_1 + \theta_2 + \theta_3)}{\Gamma(\theta_1 + \theta_2 + \theta_3 + r + s)\Gamma(\theta_1)\Gamma(\theta_2)}$$

ここで $u_{tj} = x_{tj}^2 - E(x_{tj}^2)$, $v_{tj} = x_{tj}^2 - \sigma_{tj}^2$ と置くと v_{tj} はホワイトノイズである。このとき

$$u_{tj} = \phi_j u_{t-1, j} + v_{tj} - \beta_j v_{t-1, j}$$

と ARMA(1, 1) 表現ができる。また u_{tj} のスペクトル密度は

$$f_j(\lambda) = \frac{\sigma_v^2 |1 - \beta_j z|^2}{2\pi |1 - \phi_j z|^2}$$

となる。ここに $\sigma_v^2 = \text{Var}(v_{jt})$, $z = e^{-i\lambda}$ である。このとき全銘柄の u_{tj} の集計値 (平均) $\bar{u}_t = \frac{1}{N} \sum_{j=1}^N u_{tj}$ のスペクトル密度は、Granger (1980) と同様に、 α_j と β_j に関する平均を取って

$$\bar{f}(\omega) = \frac{1}{N} \sum_{j=1}^N f_j(\omega) \equiv \frac{\sigma_v^2}{2\pi} \int \frac{|1 - \beta_j z|^2}{|1 - \phi_j z|^2} dF(\alpha, \beta)$$

によって近似する。このスペクトル密度の z^k の係数は \bar{u}_t の k 次の自己共分散 $\gamma(k)$

である。自己共分散は α, β の関数であるから、以後 $\gamma_k(\alpha, \beta)$ と書くことにする。つぎに $\gamma_k(\alpha, \beta)$ の平均的挙動を調べるために $\gamma_k(\alpha, \beta)$ を 2 変量ベータ分布で積分する。積分を容易にするためにさらに

$$\phi^n = (\alpha + \beta)^n = \sum_{k=0}^n \binom{n}{k} \alpha^k \beta^{n-k}, \quad \frac{1}{1-\phi^2} = \frac{1}{2} \left(\frac{1}{1-\phi} + \frac{1}{1+\phi} \right),$$

$$\frac{1}{1+\phi} = 1 - \phi + \phi^2 - \dots = (1-\phi) + O(\phi^2),$$

などを代入し、上の z^k 係数を α, β, ϕ で表す。そうすると z^k の係数は無数に多くの項が出てくるが、それらの項すべてを積分評価する必要はない。自己共分散 $\gamma_k(\alpha, \beta)$ において k の増加とともに減衰速度が最も遅い項（減衰速度を支配する項）だけ見ればよい。そのような項は上の 2 項展開において $k = n$ のときであることが分かるから、 α^k を含む項だけ調べればよい。その結果、AR(1) の場合と同様に

$$\gamma_j(\alpha, \beta) = \text{const.} \times j^{1-\theta_2-\theta_3}$$

の形に表されることが示される。先と同様に長期記憶パラメータで表すと $d = 1 - \frac{\theta_2 + \theta_3}{2}$ となる。この場合も集計値 \bar{u}_t は $I(d)$ の長期記憶系列となることが示される。そして $\theta_2 + \theta_3 > 1$ のとき $d < 0.5$ の定常長期記憶系列になる。 x_{ij}^2 が定常性の条件を満たせば、 \bar{x}_t^2 も長期記憶過程 $I(d)$ となる。Kawai and Maekawa (2010) は詳細な導出を与えているが、そこで用いられる導出の基本的方針は附録 2 に示された方針とほぼ同様である。

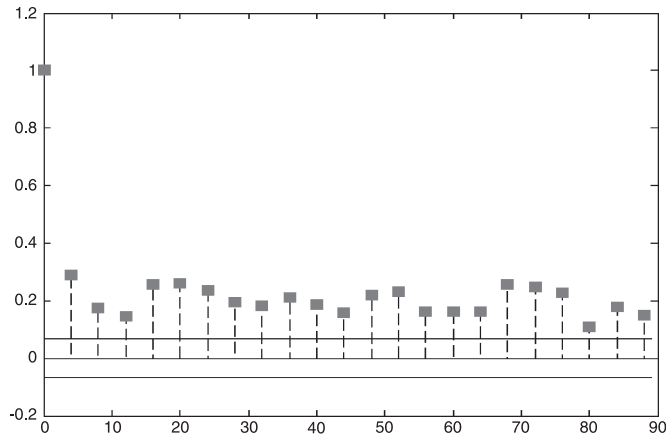
以上の理論的結論をシミュレーションによって検証してみよう。GARCH(1, 1) の係数 α と β を 2 変量ベータ分布 $f(u, v)$ からの乱数で与え、 $v = \alpha_j, u = \beta_j, j = 1, 2, \dots, 500$ を得る。これらの α と β を使って上の GARCH(1, 1) モデルから $x_{ij}, j = 1, 2, \dots, N$ を発生させ、その 2 乗の平均過程が長期記憶過程になっているかどうかを調べる。シミュレーション手順は以下のとおりである。

- [1] パラメータ (4.5, 0.7, 0.8) と設定し、2 変量ベータ分布 Beta(4.5, 0.7, 0.8) に従う確率変数 (u, v) を 500 個生成する ($j = 1, 2, \dots, 500$)。
- [2] 上記の (u, v) を $v = \alpha_j, u = \beta_j$ と対応させ、それぞれを GARCH(1, 1) モデルの分散方程式の ARCH 項と GARCH 項のパラメータとする GARCH(1, 1) プロセスを 500 系列生成する。
- [3] 上記 [2] で生成した GARCH(1, 1) プロセスから計算された 500 個本の長さ 1000

の系列 x_{tj} を作る。 ($j = 1, 2, \dots, 500, t = 1, 2, \dots, 1000$)

[4] この $x_{tj}, t = 1, 2, \dots, 1000, j = 1, 2, \dots, 500$ の時点 t における j についての集計値 (平均) \bar{x}_t^2 に関して, 対数回帰法によって長期記憶パラメータ d を推定する。

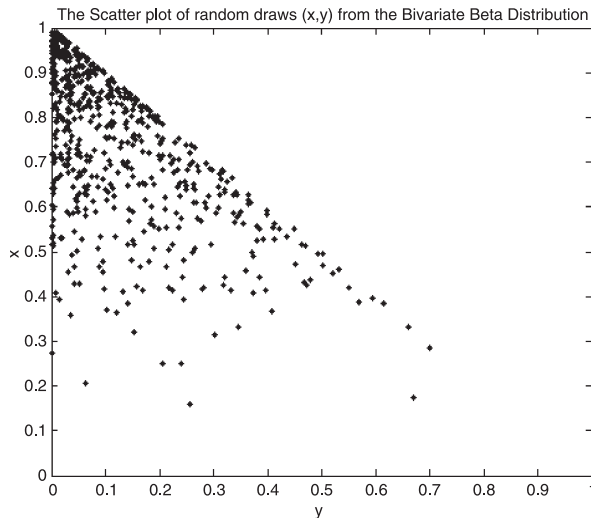
図4.2は, このように生成された \bar{x}_t^2 の自己相関係数の減衰を示したのものである。自己相関は最初急激に減少するが, その後低水準が延々と続きなかなか0に収束しないという長期記憶過程の特徴が表れている。



横軸はラグ次数, 縦軸は標本自己相関係数

図4.2 GARCH(1, 1) に従う系列の2乗 x_{tj}^2 の平均 \bar{x}_t^2 系列に対する標本自己相関係数

下の図4.3は [1] で発生させた2変量ベータ分布の乱数の散布図である。



(横軸, 縦軸) = (ARCH 項 α_j , GARCH 項と β_j)

図4.3 2変量ベータ乱数の散布図

前川・河合 (2008) は、東証1部上場1,130銘柄から GARCH(1,1) の係数推定値 $\hat{\alpha}_j$ と $\hat{\beta}_j, j=1, 2, \dots, 1130$ を計算した。図4.4はその結果を散布図に示したものである。図4.3, 4.4はよく似ており、 α_j と β_j が2変量ベータ分布に従うという仮定はかなり現実をよく描写しているといえる。

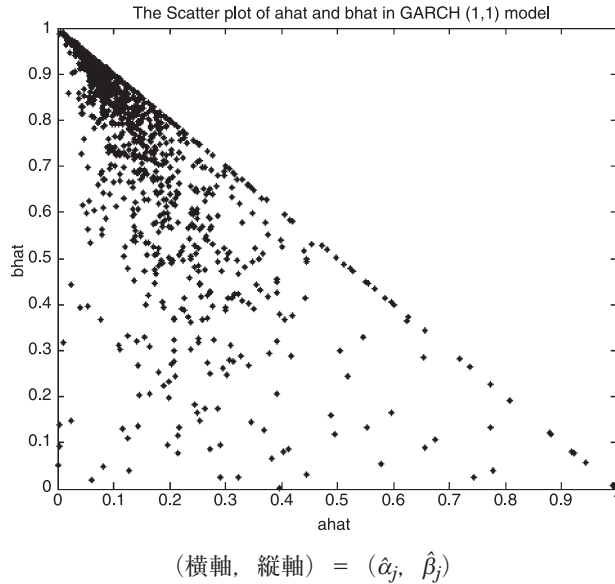


図4.4 GARCH(1,1) パラメータ推定値の散布図 (東証1部上場1,130銘柄)

[4] で得られた平均系列の長期記憶パラメータ d の推定値 \hat{d} は、先に述べた a, b を $a=0, b=0.5\sim 0.75$ とした場合、以下のようになった。

表4.2

a	b	\hat{d}
0.0	0.50	0.649
0.0	0.60	0.436
0.0	0.70	0.356
0.0	0.75	0.246

4.5 その他の長期記憶性発生メカニズム

以上、長期記憶性発生メカニズムに関する Granger (1980) の紹介と拡張を試みたが、この他にも長期記憶性発生メカニズムに関する論文がいくつかある。Lippi and Zaffaroni (1999), および Zaffaroni (2004a) はベータ分布の仮定をより弱いノンパラメトリックな場合について論じている。Zaffaroni は Granger (1980) を

より一般化した次のようなクロスセクション時系列モデルを扱い、攪乱項を個別効果と共通効果に分解するモデルを導入した。すなわち

$$\begin{aligned} x_{tj} &= \alpha_j x_{t-1, j} + \rho_j u_t + \varepsilon_{tj}, \quad j = 1, \dots, N \\ u_t &: \text{すべての単位に共通なホワイトノイズ} \\ \varepsilon_{tj} &: \text{個別単位 } j \text{ に固有なホワイトノイズ, 単位間無相関} \\ u_t, \varepsilon_{tj} &: \text{互いに独立} \end{aligned}$$

ここでは、AR パラメータ α_j 特定の分布を仮定せずノンパラメトリックな分布を想定する。すなわちある数 $b \in (-1, \infty)$ に対して次のようなノンパラメトリックな分布 $B(\alpha; b)$ を仮定する。すなわち

$$B(\alpha; b) \sim c_b(1-\alpha)^b \quad \text{as } \alpha \rightarrow 1^- \quad \text{with } 0 < c_b < \infty$$

この分布には、ベータ分布、一様分布などが含まれる。Zaffaroni は多次元の場合にも容易に拡張できると述べているが、その場合の分布については明示的に示していない。また Zaffaroni らはベータ分布を仮定しない代わりに、 α が下から 1 に近づいた場合の極限としてノンパラメトリックな分布を仮定している。現実のデータでは α が 1 に近くない場合も多いので、この仮定は問題なしとしない。Zaffaroni は、 α が 1 に近いという制約を課するという代償を払って特定の分布に依存しないという自由を得たといえよう。

この他に Ding and Granger (1996) は、クロスセクションにおける集計とは別の意味の集計量が長期記憶性をもたらすことを指摘した。彼らのモデルは以下のような GARCH(1, 1) モデルの一種の拡張版である。

$$\begin{aligned} x_t &= \sigma_t \varepsilon_t, \quad \varepsilon_t \sim i. i. d. D(0, 1), \\ \sigma_t^2 &= \sum_{i=1}^N w_i \sigma_{it}^2, \quad \sum_{i=1}^N w_i = 1 \\ \sigma_{it}^2 &= \sigma^2(1 - \alpha_i - \beta_i) + \alpha_i x_{i,t-1}^2 + \beta_i \sigma_{i,t-1}^2, \quad i = 1, 2, \dots, N. \end{aligned}$$

このモデルの特徴は、分散方程式にある。彼らは従来の GARCH モデルの分散方程式では長期記憶性を十分に捕らえることができないので、分散方程式が多数の分散要素から構成されているようなモデルを提案した。このように分散が多数の要素の集計量である場合に長期記憶性が現れることを示した。

しかし Zaffaroni (2005) は Ding and Granger のモデルから自己相関係数の緩やかな (notexponentially な) 低減は生じるものの長期記憶性は現れないと主張し

ている。筆者は Zaffaroni の主張を十分精査していないが、この点は、本論の冒頭に述べた長期記憶性の定義とモーメントの存在条件に深くかかわっていると思われる。今後の検討課題としたい。

5. 構造変化と長期記憶性

われわれは先に、大きな経済変動の時期に長期記憶パラメータ d が大きく変動することを見た。このことは構造変化と長期記憶性は無関係ではないことを示している。構造変化を見落とすという想定誤差が見せかけの長期記憶性をもたらすことを指摘する研究もある

Granger and Hyung (2004) はレベルがランダムにジャンプする次のようなモデルを取り上げ、レベルのシフトが長期記憶性をもたらすことを示した。

$$y_t = m_t + \varepsilon_t, \quad t = 1, \dots, T$$

$$m_t = m_{t-1} + q_t \eta_t, \quad \eta_t \sim i. i. d. (0, \sigma_\eta^2)$$

$$q_t = \begin{cases} \text{確率 } 1-p \text{ で } 0 \\ \text{確率 } p \text{ で } 1 \end{cases}$$

ここで仮定

$$p \rightarrow 0 \text{ のとき } T \rightarrow \infty \text{ かつ } \lim_{T \rightarrow \infty} Tp \rightarrow \text{ゼロでない定数}$$

を置くと、 y_t の自己相関係数は急激に (exponential に) は減少せず、緩やかに (hyperolic に) 減少し、典型的な長期記憶過程の自己相関の減少と同様な振る舞いをすることを示した。

Diebold and Inoue (2001) も類似のモデルを用いて同様な現象を指摘した上で、たとえジャンプが見せかけの長期記憶性の原因であるとしても、依然として長期記憶モデルが予測に有効でありうるので、長期記憶かジャンプかという単純な二者択一に懐疑的な態度を取っている。

このほかに Mikosh and Starica (2000) も構造変化を伴う IGARCH の下で構造変化を無視すると長期記憶性が現れることを指摘している。

これらの研究は、見せかけの長期記憶性かどうかを見極めるためには構造変化の検定が必要であることを示唆している。長期記憶過程における構造変化に関してもさまざまな研究がなされている。Berkes et al (2006) は長期記憶過程モデル (Fractional Brown motion) と平均シフトモデル

$$X_i = \begin{cases} \mu + Y_i, & 1 \leq i \leq k^* \\ \mu + \Delta + Y_i, & k^* \leq i \leq n \end{cases} \quad \text{ここで } k^* \text{ は未知のジャンプ時点}$$

との選択という観点から、cusum テストを提案している。ここに Y_i は平均 0 の j 階の自己共分散 γ_j と 4 次のキユムラントをもつ定常過程である⁽¹⁾。しかし彼らのシミュレーションによるとこのテストのパフォーマンスはあまり良くない。また Ling (2007) はかなり一般的な観点から、変化点を検出する Wald タイプのテストの漸近理論を示し、その結果を使って ARFIMA モデルの構造変化点を検定する方法を提案している。しかし彼は、小標本におけるシミュレーションによるとこのテストのパフォーマンスはよくないと述べている。

最近 O. Na, J. Lee, and S. Lee (2009) は d の変化点を検出する次のような Cusum テストを提案した。彼らを取り上げたモデルは ARFIMA(p, d, q) モデル

$$\Phi(L)(1-L)^d y_t = \Psi(L)\epsilon_t$$

である。ここに

$$\begin{aligned} \Phi(L) &= 1 - \phi_1 L - \phi_2 L^2 \cdots - \phi_p L^p, \\ \Psi(L) &= 1 - \psi_1 L - \psi_2 L^2 \cdots - \psi_q L^q \end{aligned}$$

である。

のパラメータベクトルを $\theta = (d, \phi_1, \phi_2, \dots, \phi_p, \psi_1, \psi_2, \dots, \psi_q)$ とする。このとき cusum テストとして

$$T_n = \max_k \frac{k^2}{n} (\hat{\theta}_k - \hat{\theta}_n)' \hat{\Omega}_n (\hat{\theta}_k - \hat{\theta}_n)$$

を定義する。ここに $\hat{\theta}_k$ は k 時点までのデータを使って計算された θ_0 の推定値、 $\hat{\theta}_n$ はすべてのデータから計算された θ_0 の推定値である。またこのモデルから誤差項は $\epsilon_t(\theta) = \Psi^{-1}(L)\Phi(L)(1-L)^d y_t$ と (無限級数で) 書けるが、これを有限時点で打ち切ったものを $e_t(\theta) = \sum_{j=0}^{t-1} \alpha_j(\theta) y_{t-j}$ とする。ここで $\hat{\Omega}_n$ はある種の共分散行列である。このときいくつかの条件の下で $n \rightarrow \infty$ のとき T_n は $p+q+1$ 次元の Brownian bridge の 2 乗の \sup すなわち $\sup_s \|w_{p+q+1}^0(s)\|^2$ に弱収束することを示した。そして $\sup_s \|w_{p+q+1}^0(s)\|^2$ の分布表から T_n がある臨界値を超えれば時点 k をパラメータの変化時点とする。彼らは、Hurst (1951) が解析したナイル川の最小水位の 622 年から 1284 年までのデータに ARFIMA(0, d, 0) をあてはめ、彼らの Cusum テストを実行したところ 732 年に d の値が大きく変化したことを突き止めている。

構造変化と長期記憶過程の問題に関しては、この他にもおびただしい数の論文がある。筆者が知る限り、現時点でこの分野に関する最新のサーベイは Andreou and Ghysels (2009) である。

終 わ り に

以上の本論において、長期記憶性が発生する 2, 3 の状況を示した。このほかにも本論で取り上げなかった発生メカニズムについては別の機会にサーベイしたい。今後これら個別の発生メカニズムの背後に存在するであろう共通の原因とそのメカニズムを統一的に表すモデルを模索したい。

附 録 1

ここで Granger (1980) の (誤植を訂正しつつ) 導出を詳細に示しておく。まず

$$\frac{1+\alpha z}{1-\alpha z} = 1 + 2 \sum_{j=1}^{\infty} (\alpha z)^j, \quad z = e^{-i\omega} \frac{1+\alpha \bar{z}}{1-\alpha \bar{z}} = 1 + 2 \sum_{j=1}^{\infty} (\alpha \bar{z})^j, \quad \bar{z} = e^{i\omega}$$

であるから

$$\frac{1+\alpha z}{1-\alpha z} + \frac{1+\alpha \bar{z}}{1-\alpha \bar{z}} = 2 + 2 \sum_{j=1}^{\infty} (\alpha z)^j + 2 \sum_{j=1}^{\infty} (\alpha \bar{z})^j$$

となる。また

$$\frac{1+\alpha z}{1-\alpha z} + \frac{1+\alpha \bar{z}}{1-\alpha \bar{z}} = \frac{2(1-\alpha^2)}{|1-\alpha z|^2}$$

であるから

$$\frac{1}{|1-\alpha z|^2} = \frac{1}{2(1-\alpha^2)} \left\{ \frac{1+\alpha z}{1-\alpha z} + \frac{1+\alpha \bar{z}}{1-\alpha \bar{z}} \right\} = \frac{1}{(1-\alpha^2)} \left\{ 1 + \sum_{j=1}^{\infty} (\alpha z)^j + \sum_{j=1}^{\infty} (\alpha \bar{z})^j \right\}$$

となる。全く些細な誤植であるが Granger (1980) では上式の $\frac{1}{(1-\alpha^2)}$ の部分が $\frac{1}{(1-\alpha)^2}$ になっているため、そのまま読み進めると混乱が生じる。Granger (1980) は「特殊」なベータ分布

$$dF(\alpha) = \frac{2}{B(p, q)} \alpha^{2p-1} (1-\alpha^2)^{q-1} d\alpha, \quad 0 \leq \alpha \leq 1$$

を仮定している。この分布を使えば

$$\begin{aligned} \int \frac{1}{|1-\alpha z|^2} dF(\alpha) &= \int \frac{1}{|1-\alpha z|^2} \frac{2}{B(p, q)} \alpha^{2p-1} (1-\alpha^2)^{q-1} d\alpha \\ &= \frac{2}{B(p, q)} \int \frac{1}{(1-\alpha^2)} \left[1 + \sum_{j=1}^{\infty} (\alpha z)^j + \sum_{j=1}^{\infty} (\alpha \bar{z})^j \right] \alpha^{2p-1} (1-\alpha^2)^{q-1} d\alpha, \quad 0 \leq \alpha \leq 1 \end{aligned}$$

となる。次にこの積分における z^j の係数を求めれば、

$$\frac{2}{B(p, q)} \int \frac{1}{(1-\alpha^2)} \alpha^{2p+j-1} (1-\alpha^2)^{q-1} d\alpha = \frac{2}{B(p, q)} \int \alpha^{2(p+\frac{j}{2})-1} (1-\alpha^2)^{q-2} d\alpha$$

となる。ここで $\alpha^2 = x$ (すなわち $\alpha = \sqrt{x}$) という変数変換を行う。 $\frac{dx}{d\alpha} = 2\alpha$ (すなわち $d\alpha = \frac{dx}{2\sqrt{x}}$) に留意すれば、上の積分は

$$\begin{aligned} \frac{2}{B(p, q)} \int x^{p+\frac{j}{2}-1} (1-x)^{q-2} \frac{dx}{2\sqrt{x}} &= \frac{1}{B(p, q)} \int x^{p+\frac{j}{2}-\frac{1}{2}} (1-x)^{q-2} = \frac{B(p+\frac{j}{2}, q-1)}{B(p, q)} \\ &= \frac{\Gamma\left(p+\frac{j}{2}\right)\Gamma(q-1)}{\Gamma\left(p+\frac{j}{2}+q-1\right)} \times \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} \equiv K \times j^{1-q} \equiv K \times j^{p-(q-1)} \end{aligned}$$

となる。ここでスターリングの公式から

$$\frac{\Gamma(k+a)}{\Gamma(k+b)} \sim k^{a-b}$$

となることを用いた。

附 録 2

次に「特殊」なベータ分布を用いない場合の導出を示す。附録1と同様に積分

$$\int \frac{1}{|1-\alpha z|^2} dF(\alpha)$$

を求めたい。「特殊」ベータ分布を仮定しない場合は

$$\begin{aligned} \int \frac{1}{|1-\alpha z|^2} dF(\alpha) &= \frac{1}{B(p,q)} \int \frac{1}{|1-\alpha z|^2} \alpha^{p-1} (1-\alpha^2)^{q-1} d\alpha \\ &= \frac{1}{B(p,q)} \int \frac{1}{(1-\alpha^2)} \left\{ 1 + \sum_{j=1}^{\infty} (\alpha z)^j + \sum_{j=1}^{\infty} (\alpha \bar{z})^j \right\} \alpha^{p-1} (1-\alpha^2)^{q-1} d\alpha \end{aligned}$$

となる。ここで

$$\frac{1}{(1-\alpha^2)} = \frac{1}{2} \left(\frac{1}{1-\alpha} + \frac{1}{1+\alpha} \right) = \frac{1}{2} \left(\frac{1}{1-\alpha} + 1 - \alpha + \alpha^2 - \dots \right)$$

と展開しておく。

先と同様に z^j の係数に着目し、上の展開を代入すると

$$\begin{aligned} &= \frac{1}{B(p,q)} \int \frac{1}{(1-\alpha^2)} \alpha^{p+j-1} (1-\alpha)^{q-1} d\alpha \\ &= \frac{1}{B(p,q)} \int \frac{1}{2} \left(\frac{1}{1-\alpha} + 1 - \alpha + \alpha^2 - \dots \right) \alpha^{p+j-1} (1-\alpha)^{q-1} d\alpha \end{aligned}$$

となる。右辺の第1項 ($\frac{1}{1-\alpha}$ に関する項) は

$$\frac{1}{B(p,q)} \int \alpha^{p+j-1} (1-\alpha)^{q-2} d\alpha = \frac{B(p+j, q-2)}{B(p,q)} = A \times j^{1-q}$$

となる。右辺の第2項以降 (級数展開の第 k 項 ($k = 0, 1, 2, \dots$)) は

$$\frac{1}{B(p,q)} \int \alpha^{p+j+k-1} (1-\alpha)^{q-2} d\alpha = B \times j^{-q}$$

となる。したがって j が大きいときは第2項以降は第1項に比べて j^{-1} のオーダーだけ小さい。すなわち z^j の係数は

$$A \times j^{1-q} + O(j^{-1}) \cong A \times j^{1-q}$$

である。ここに A, B は j に依存しない定数である。

注

- (1) より正確には Berkes et al (2006) は Y_i の過程を、以下のように定式化している。
 $\{Y_i\}$ は、平均0で j 階の自己共分散 $\gamma_j = Cov(Y_0, Y_j)$ と4次のキュムラント $\kappa(h, r, s) =$

$E[Y_k Y_{k+h} Y_{k+r} Y_{k+s}] - (\gamma_h \gamma_{r-s} + \gamma_r \gamma_{h-s} + \gamma_s \gamma_{h-r})$ を持ち、次の条件を満たす過程である。

条件 1 : ある数 σ に対して $n^{-1/2} \sum_{1 \leq j \leq nt} Y_j \xrightarrow{d} \sigma W(t)$

ここに $W(t)$ は区間 $[0, 1]$ 上のウィーナープロセス。

条件 2 : $\sum_j |\gamma_j| < \infty$

条件 3 : $\sup \sum_{r,s} |\kappa(h, r, s)| < \infty$

参 考 文 献

- Andreou, E. and Eric Ghysels (2009) Structural Breaks in Financial Time Series, In Handbook of Financial Time Series, Springer, 839–870
- Balakrishnan, N, and Chin-Diew Lai (2009) Continuous Bivariate Distributions (Second Edition), Springer
- Berkes, I., L. Horvath, and P. Kokoszka (2006) On Discriminating between Long-range Dependence and Changes in Mean, The Annals of Statistics Vol. 34, No. 3, 1140–1165
- Diebold, F. X. and A. Inoue (2001) Long Memory and Regime Switching, Journal of Econometrics 105, 131–159
- Ding and Granger (1996) Modeling Volatility Persistence of Speculative Returns: A New Approach, Journal of Econometrics 73, 185–215
- Geweke, J., Proter-Hudak, S. (1983) The Estimation and Application of Long Memory Time Series Models. Journal of Time Series Analysis 4, 15–39.
- Granger, C. W. J. (1980) Long Memory Relationships and the Aggregation of Dynamic Models, Journal of Econometrics, 14, 227–238
- Granger, C. W. J. and N. Hyung (2004) Occasional Structural Breaks and Long Memory with an Application to the S&P 500 Absolute Stock Returns, Journal of Empirical Finance 11 399–421
- Hansen, P. R., and A. Lunde (2005) A Forecast Comparison of Volatility Models: Does Anything Beat a GRACH(1, 1)?, Journal of Applied Econometrics, 20(7), 873–889
- Hurst, H. E. (1951) Long-term Storage Capacity of Reservoirs. Trans. Amer. Soc. Civil Engineers, 116, 770–799
- Kawai, K. and K. Maekawa (2009) Aggregation of Squared GARCH(1, 1) Process and Long Memory, Discussion Paper, Hiroshima University of Economics
- Lippi, M. and P. Zaffaroni (1998) Aggregation of Simple Linear Dynamics: Exact Asymptotic Results, Discussion Paper, London School of Economics
- Ling, S. (2007) Testing for Change Points in Time Series Models and Limiting Theorems of NED Sequences, The Annals of Statistics, Vol. 35, No. 3, 1213–1237
- Maekawa, K. and X. Lu (2009) Long Memory in the Realized Volatility of Returns on the Yen/US\$ Exchange Rate during the Three Financial Crises, HUE Journal of Economics and Business, Vol. 31, No. 4, 41–69
- Mandelbrot, B. B., and R. L. Hudson (2004) The (mis) Behavior of Markets, A Member of Preseus Books Group, New York (『禁断の市場』高安英樹監訳, 東洋経済新報社, 2008)

- Mikosch, T. and C. Starica (2004) Nonstationarities in Financial Time Series, the Long-range Dependence and the IGARCH Effects, *The Review of Economics and Statistics*, 86(1) 378–390
- Mikosch, T. and C. Starica (2000) Change of Structure in Financial Time Series, Long Range Dependence and the GARCH Model, Working Paper, University Aarhus
- Na, O., J. Lee, and S. Lee (2009) Constancy Test for FARIMA Long Memory Processes, Discussion Paper, Seoul National University
- Palma, W. (2007) *Long-Memory Time Series*, John Wiley & Sons
- Tokutsu, Y., S. Nagata, and K. Maekawa (2008) A Comparison of Estimators for Long — Memory Process — Simulation and Empirical Study —, FEMES2008, Singapore Management University
- Zaffaroni, P. (2004) Contemporaneous Aggregation of Linear Dynamic Models in Large Economies, *Journal of Econometrics* 120, 75–102
- Zaffaroni, P. (2005) Contemporaneous Aggregation of GARCH Process, *Journal of Time Series Analysis* 28, 4, 512–544.
- 田中勝人 (2006) 『現代時系列分析』 岩波書店
- 得津康義, 永田修一, 前川功一 (2007) 「RV の長期記憶性に関するウェーブレット分析」 第28回ジャフィー大会 中央大学
- 前川功一, 河合研一 (2008) 「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」, 『広島経済大学経済研究論集』 第30巻 3・4号 pp. 53–69
- 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』 pp. 103–202 岩波書店