論　文

Volatility Clustering と Fat Tails を考える

原　田　康　平

目　次

はじめに

筆者は，前報[1]において，株価や為替レートの日次変動率が示す20日から30日の記憶過程，いわゆる volatility clustering について詳細な分析を行い，世界の主要国の株価指数のみならず，円ドルレート，日米の債券金利がすべてほとんど同じ記憶特性を持っていること，したがって，世界の金融市場全体に共通したダイナミクスが存在する可能性を示した。ここでは，前報の中で否定的に言及した「出来高の緩和だけで変動率の clustering を説明できるか」という点について改めて考えたい。

1. Volatility Clustering と Fat Tails

まず，volatility clustering と fat tails について簡単に解説しておきたい。

図 1 は2000年から2004年末までの TOPIX 日次終値と東証一部出来高，図 2 は日次の変動率とレンジを示している。ここで，日次変動率とレンジはそれぞれ

$$日次変動率 = \log_e(当日終値／前日終値) \times 100$$

$$レンジ = \log_e(当日高値／当日低値) \times 100$$

により求めている。

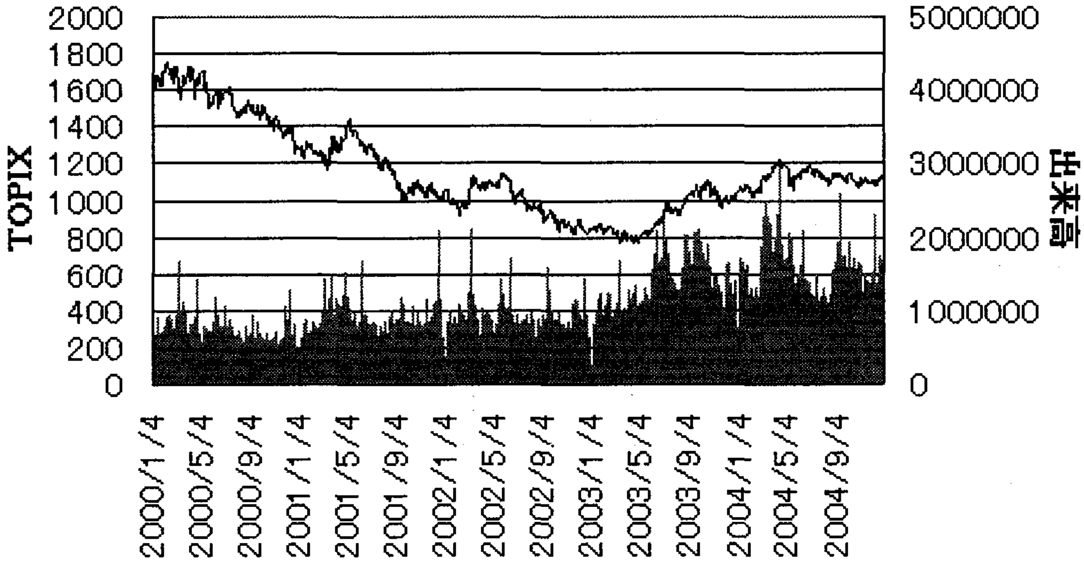

図 1　2000年 1 月から2004年12月までの TOPIX および東証一部出来高の推移

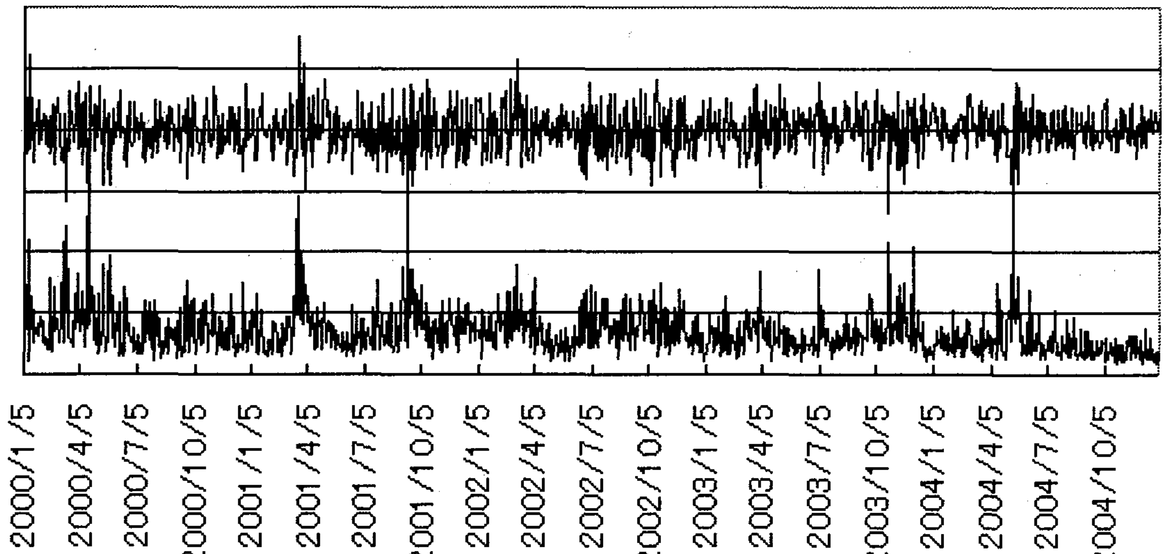

図 2　TOPIX 日次変動率（上段）とレンジ（下段）の推移

図3はTOPIX日次変動率の分布であり，一見，正規分布に近いが，尖度は1.6と1より大きく，正規分布より尖った裾の広い分布となっている。言い換えると，比較的大きい増減が正規分布から期待されるより高い割合で出現する。これがfat tailsと呼ばれる現象であって，変動率に正規性を仮定するVaR（Value at Risk）推定やオプションプレミアム式の問題点として長く指摘されてきた[2]。また，近年はもっと短い時間スケールでの変動についても，Levy分布といった可能性も議論されている[3, 4]。

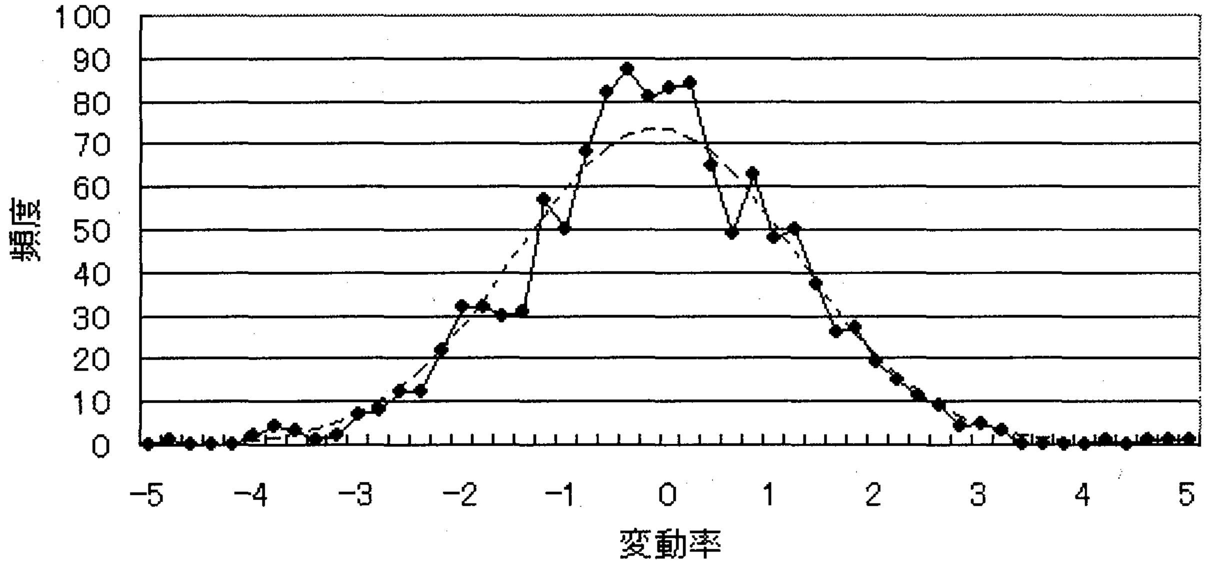

図3　TOPIX日変動率の分布（点線は正規分布曲線）

しかしながら，株価の変動率は時系列であって，時間特性抜きに分布だけを議論してもあまり意味はない。図4は変動率，図5は変動率平方およびレンジ，出来高の自己相関係数を示している。いずれも2004年12月30日までの300営業日×2区間について次式で求めた結果を平均し，3項移動平均を施したものである。

$$(1)\qquad R(\tau)=\sum_{i=1}^{n-\tau}(x_i-\bar{x})(x_{i+\tau}-\bar{x})/(n-\tau)/V$$

ここで$\bar{x}$，V，nはそれぞれ標本平均と標本分散および標本数である。明らかに

変動率の自己相関は概ねデルタ関数に等しく，この限りで，毎日の変動はまったく記憶を持たない。一方，変動率平方，レンジおよび出来高の自己相関はいずれもほぼ同じ1次の緩和を示している。

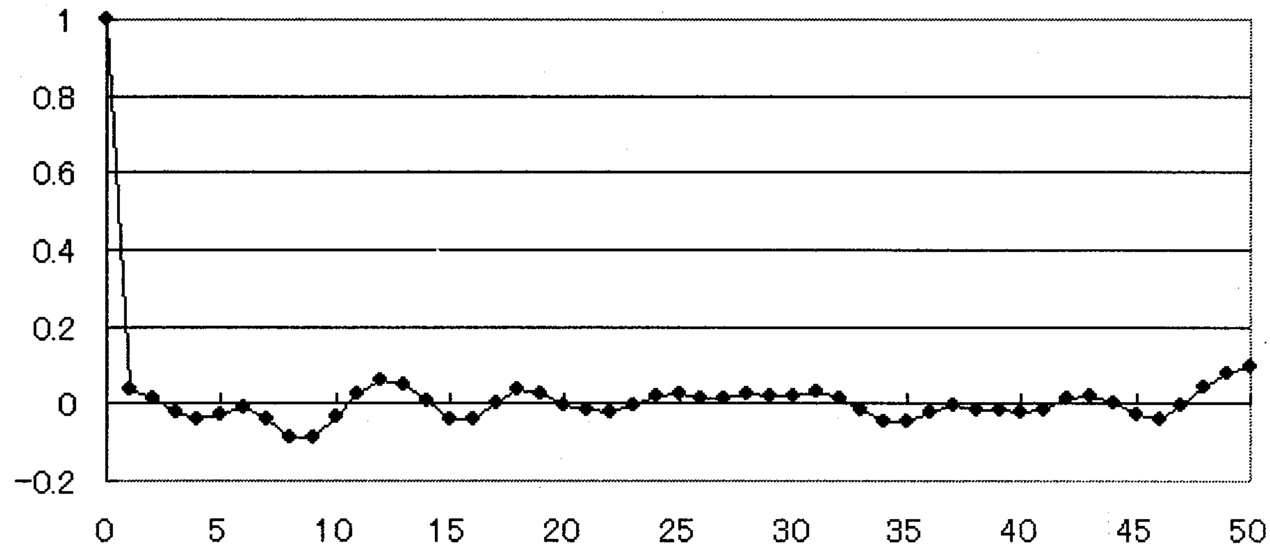

図4　TOPIX 日次変動率の自己相関係数

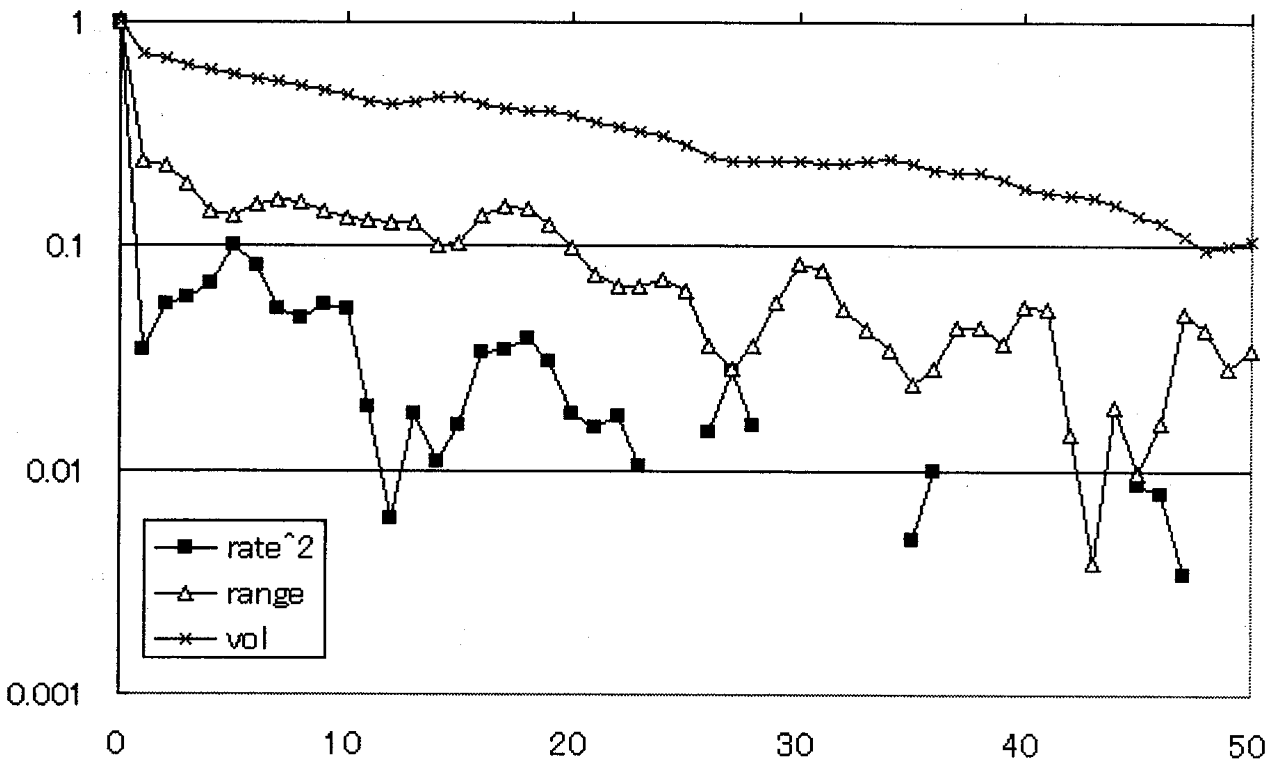

図5　TOPIX 日変動率平方（rate^2），レンジ（range）および出来高（vol）の自己相関係数

そこで，レンジの自己相関係数 $R_R(\tau)$ および出来高の自己相関係数 $R_V(\tau)$ の対数値について，ラグ１日から20日までの回帰式を求めたところ，次の結果が得られた。

（2）　$\log_{10} R_R(\tau) = -0.01271\,\tau - 0.712, \quad R^2 = 0.55$

（3）　$\log_{10} R_V(\tau) = -0.01341\,\tau - 0.163, \quad R^2 = 0.92$

回帰係数を -0.013 とすれば，自己相関が $1/e$ だけ下がるに要するラグは33日であり，したがって，レンジ，出来高ともおよそ33日の緩和時間を持つ１次の緩和過程とみなすことができる。つまり，ある平均値からの乖離を x とおき，x について，

（4）　$$\frac{dx}{dt} = -ax + f(t), \quad a > 0, \quad \langle f(t) \rangle = 0, \quad \langle f(t)f(t+\tau) \rangle = 2D\delta(\tau)$$

なる Langevin 方程式を想定する。ここで，$\langle \cdots \rangle$ は時間平均を示している。このとき，自己相関関数は

（5）　$$R(\tau) = R(0)e^{-a\tau} = R(0)e^{-\tau/\tau_c}$$

で与えられ[5]，この τ_c が33日ということである。ただし，緩和時間は推定に用いたラグの範囲にかなり敏感なため，ある程度の誤差を見込んでおく必要がある。

このような変動率平方が示す記憶が volatility clustering と呼ばれる現象であって，基本的に「変動自体は記憶を持たないが，変動の大きさ（volatility）が記憶を持つ」ものとして理解されてきた。この性質を考慮した時系列モデルが GARCH（Generalized Auto-Regressive Conditional Heteroscedasticity）モデルと SV（Stochastic Volatility）モデルである[6, 7]。GARCH は，残差系列 ε_t に対して，

（6）　$$E_{t-1}(\varepsilon_t^2) = \sigma_t^2$$

で定義された volatility が

(7) $$\sigma_t^2 = k + \sum_i \beta_i \sigma_{t-i}^2 + \sum_j \alpha_i \varepsilon_{t-j}^2$$

なる自己回帰に従うとする。一方，SV モデルは

(8) $$\log(\sigma_t^2) = \omega + \phi \log(\sigma_{t-1}^2) + \varepsilon_t$$

として，volatility の対数が自己回帰すると仮定する。容易に想像できるように，SV モデルは係数 ϕ，GARCH(1, 1) の場合は $\alpha_1 + \beta_1$ の大きさに依存して volatility が緩和する。

つまるところ，株価の変動では，比較的小さい揺らぎと大きい変動が30日前後の緩和時間で交錯して繰り返されるということであり，結果的に，前者は分布の中央を尖ったものとし，後者は広い裾を演出する。定性的には volatility の緩和によって clustering と fat tails の両方が説明されることになる。

2．何が問題なのか

前節の最後に触れた結論は周知のことであり，それだけに GARCH や SV モデルの精緻化も含めて，膨大な数の論文が提出されてきた（たとえば文献[7]に詳しい）。このような現象論的モデルは，現象の発生メカニズムそのものを直接に説明するものではないから，今後も対象を変えた同様の研究報告が量産される可能性が高い。

筆者が前報[1]で主張した内容は「緩和過程は，変動率平方（あるいは絶対値）よりレンジや出来高により明瞭に現れる」「同じ緩和過程は世界の株価指数，為替レート，金利などにほぼ共通して観測される」ということである。後者は，たとえば株価だけに注目して精緻な時系列モデルを追及してもあまり意味がないということ，前者は volatility clustering が日内の速い変動とそのサンプル点数の違いという 2 つに階層に関わることを示唆している。GARCH であれ SV であれ，

現象を自己完結的に記述しているように見えて，この２点をまったく含まず，結果的に何も説明していない。

ここでは「世界の金融市場に存在するダイナミクス」には触れない。この問題に踏み込むには，おそらく，膨大な金融データについて詳細な検証を行う必要があると思われるからである。したがって，本論では「volatility clustering はどの階層に起因するのか」に論点を絞りたい。

為替取引に関する Stanley らの実証分析[3]以降，とりわけ経済物理学などの分野で金融取引が示す短時特性に注目が集まっている。現在，「ティック間の変動はべき乗則を示す」「記憶はない」がほぼ共通した理解であって，複雑系などフラクタル性を意識したモデルシミュレーションなどが進められている。その一方で，「短時の統計特性と日次や月次の変動特性はどのように接続されるのか」という問題に言及されることはほとんどない。短時変動に記憶がなければ，中心極限定理がそのまま適用されるわけで，いうなら「正規分布に漸近する」がほぼ自明のこととして受け入れられているように見える。しかしながら，毎時あるいは毎日の取引頻度がほぼ一定であるなら，まさに中心極限定理が支配する世界となるが，現実の取引頻度はかなりの幅で変動する。図１に示した東証１部の出来高はその一例であり，しかも，出来高の変動自体が図３のような記憶特性を持っている。とするなら，日次や月次の時系列特性は，短時の特性にある種の変調がかかったものということになり，GARCH も SV も単に変調特性を観察しているに過ぎない可能性が高い。

この推論は，GARCH など変動率を対象に展開されてきた精緻な分析の意義を根底から覆しかねない内容を含んでいる。もし，volatility clustering が出来高あるいは取引頻度の緩和だけに由来するのであれば，中心極限定理が母集団の正規性を要求しないのと同じように，clustering が短時変動の統計的性質とまったく関わらない可能性も否定できない。つまり，短時特性と日次以上の特性の接続

を考慮する必要はなく，両者は単純に異なる階層の現象として分離されることになる。次節で，もう少し慎重に，この問題を考えてみたい。

3．出来高と volatility

短時の株価の変動が Wiener 過程とみなし得る，言い換えると，「期待値は 0 で，自己相関はデルタ関数」という過程であるとき，時刻 0 である値から出発した株価の確率分布は単純な拡散方程式に従い，時間 t だけ経った後の確率分布は分散が時間 t に比例する正規分布に近づいていく。これは時間軸上で表現した中心極限定理にほかならない。

ところで，株価は取引が成立するときに動く以上，株価変動の時間軸は実時間というより取引回数で考えた方が実態に近い。そう考えると，始値から終値までの動きの幅は，取引回数の多い日の方が少ない日より広がる傾向を示すはずである。ここで，1 日のうちの取引ごとの変動率の並び r_1，r_2，・・・，r_n が Wiener 過程とみなし得るなら，j 回目の取引の株価 p_j は

(8) $$p_j = p_0 e^{\Sigma r_i}$$

で与えられる。1 回ごとの変動率の平均を 0，分散を σ^2 とすると，始値に対する終値変動率の確率分布は平均 0，分散 $n\sigma^2$ の正規分布に漸近することになる。このとき，変動率平方の期待値は分散 $n\sigma^2$ に比例し，レンジはおよそ標準偏差 $\sqrt{n}\,\sigma$ に比例する。したがって，取引回数が多い日の変動率平方，レンジとも増えるはずであり，出来高が取引回数に比例するなら，出来高と変動率平方，レンジの間にも同様の関係が成り立つことが予想される。

具体的なデータを見てみよう。図 6 は，一例として，1998年 1 月 5 日から2002年12月30日までの富士通の出来高と変動率平方およびレンジの関係を示している。

明らかに，出来高が大きいほど変動率平方，レンジとも大きくなる傾向が見られる。そこで，日経平均採用銘柄225社のうち，データの連続性に難がない159社を対象として，1998年1月4日から2002年12月30日までの出来高と変動率平方，レンジの回帰式

(9)　$\log_e(\text{変動率平方}) = \alpha_1 \times \log_e(\text{出来高}) + \beta_1$

(10)　$\log_e(\text{レンジ}) = \alpha_2 \times \log_e(\text{出来高}) + \beta_2$

を調べたところ，α_1 として 0.70 ± 0.22，α_2 として 0.25 ± 0.08 が得られた。なお，ほとんどのケースで回帰は有意である。

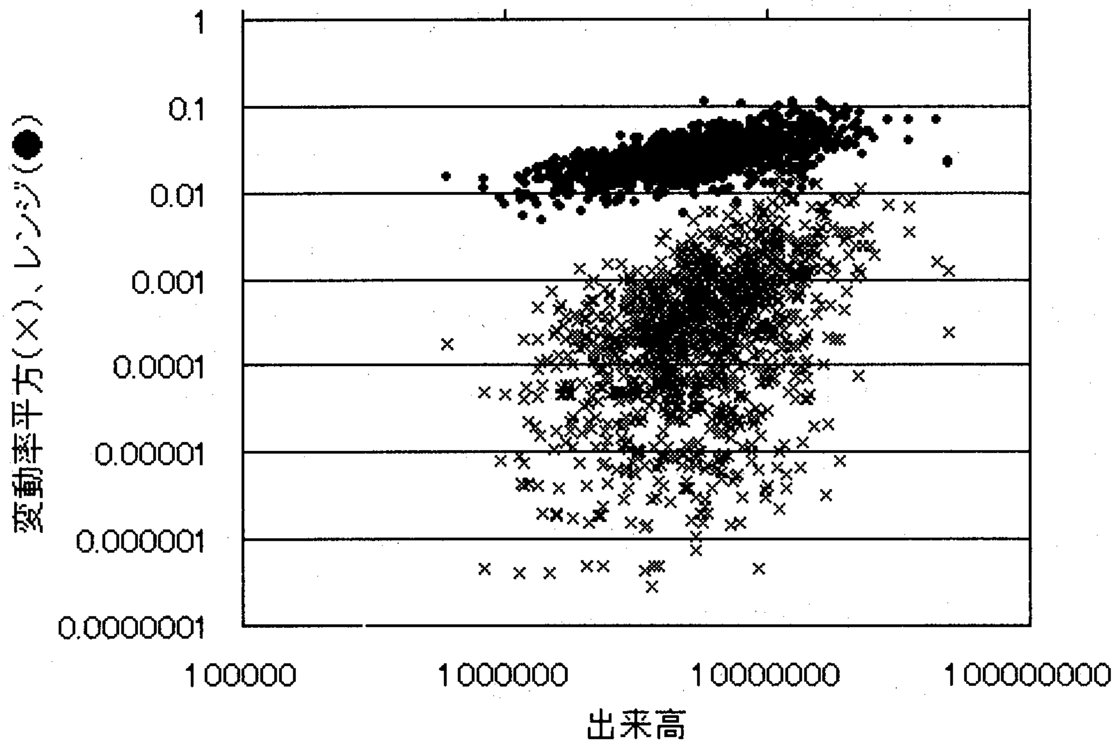

図6　1998年～2002年における富士通の出来高と変動率平方，レンジ

つまり，変動率平方 r^2，レンジ R と出来高 V の間には

(11)　$r^2 \propto V^{0.70}$,　$R \propto V^{0.25}$

が近似的に成り立つということであって，出来高が記憶を持つなら，それはほぼ

ダイレクトに変動率平方とレンジに反映されることが分かる。回帰係数が中心極限定理から期待される1.0および0.5にならない理由としては，出来高が多い日は1回あたりの取引量も大きい可能性が考えられる。

以上の結果は，変動率平方に見られるvolatility clusteringが単に出来高の記憶過程を反映しているに過ぎない可能性を強く示唆している。実際，GARCHモデルを用いてvolatility clusteringを分析したLamoureuxとLastrapesの研究[8]は，出来高を説明変数として加えると，残差平方とvolatilityの係数が有意性を失うケースが多いことを示している。

4．なぜ単純な議論が受け入れられてこなかったのか

ここまで触れてきた「出来高と株価変動率は相関している」「出来高は長期の記憶を持つ」などの点もまた，金融研究の世界では周知の事実である。株価と出来高の関係を検証した報告や，出来高の自己相関からGARCHなどに発展させた文献も数多い[9, 10]。また，Lamoureuxらの研究[8]は，volatility clusteringにおける出来高の決定的な役割を示している。にもかかわらず，なぜ多くの研究が変動率にのみ注目し，出来高の役割は限定的とみなされてきたのか。この点について，ここでは次の問題を指摘しておきたい。

まず，clusteringの証拠として，変動率絶対値や変動率平方，出来高の自己相関係数を挙げている報告は多いが，少なくとも筆者が手に取った範囲において，縦軸を対数スケールでプロットしたケースはなかった。また，平均操作や平滑化を試みたケースもなかった。これらのことは，自己相関係数の推定と解釈において，かなり重大な問題が存在することを示している。

おそらく株価や出来高の変動系列には，純然たる不規則変動，いわゆるノイズ成分とclusteringに関わる成分，そして，市場が不可避的にはらむ長期のトレン

ド成分などが含まれる。たとえば，好景気のときは出来高が膨らむのが一般的傾向であり，数ケ月あるいは1年以上にわたって上昇，下降のトレンドを示すことも少なくない。この成分は，とりわけ出来高やレンジの自己相関係数を推定するとき，きわめて大きい撹乱要因となる。変動率の場合，緩やかなトレンド成分は変換の過程でほぼ除去される。単純にいうなら，変動率は株価の時間微分に等しく，低周波成分は周波数の2乗に比例して落ちていく。これに対し，出来高やレンジの自己相関係数にはトレンド成分が除去されず，それが大きい場合，結果的に長い緩和時間として現れることになる。したがって，単純に変動率平方やレンジ，出来高の自己相関係数を推定したとして，見かけ上，これらがかなり異なる特性を示す可能性は大きい。このように，少なくとも変動率平方とレンジ，出来高の間には時間微分の有無という基本的な違いがある点を十分に念頭に置いておかなければならない。

この問題に対処するためには，推定する区間を短めに取る必要がある。この結果，推定値の信頼性は低下し，これをカバーするために，ある程度の平均操作と平滑化が不可欠となる。そもそも，自己相関係数やスペクトルは有効推定量ではないという根本的問題があり，これからも平均操作は欠かせない。つまり，自己相関係数の推定においては，区間長の設定や適切な平均操作など，経験的に検証するしかない問題が少なくないということである。

さらにもう一点，自己相関係数を対数プロットしない習慣はもっと重大であり，Granger に始まる「スペクトルをログログプロットしない」習慣（これについては文献[5]，第6章の冒頭で触れている）と重なる。基本的に，自己相関係数の解釈は「デルタ関数か，指数関数か，減衰振動型か，それ以外か」の判断から始まる。自己相関係数を対数プロットするのは，指数関数型をより正確に判読するためのテクニックに過ぎない。最初の3つは，それぞれ完全な不規則系，1次システム，2次システムに対応し，高次の線形システムはこれらの組合せとなる。

したがって，この3つに分類されないとき，その他の，たとえばカオスなどを含めた候補が探索されることとなる。

図5に示されているように，変動率平方やレンジの自己相関はほぼ指数関数型であり，これから抽出される情報は「不規則成分と緩和成分の割合」と「緩和時間」の2つだけに過ぎない。SVモデルでいうなら，回帰成分と残差の割合が前者に対応し，後者には回帰係数ϕ，GARCH(1, 1)なら$\alpha_1+\beta_1$が対応する。そして，変動率平方，レンジ，出来高の自己相関を比較するときは，「同じ指数関数型か」「そうであるなら，緩和時間は異なるのか」などが主要なチェックポイントとなる。「各指標にはほぼ共通した特性が見られる」という筆者の推論はこのようなチェックを経た後の判断であって，変動率平方とレンジ，出来高の自己相関特性，つまりダイナミカルな性質がほぼ同じであるなら，それらは同じ現象を違った角度から捉えているに過ぎないとの結論に至る。この推論が正しければ，volatilityと出来高の両方を含むGARCHやSV分析はまったくナンセンスな試みとなる。

ともあれ，従来の金融時系列分析には推定あるいは判読の点で問題があり，それがごく単純な現象を見逃し，ときに無意味となりかねない精緻化をもたらしている可能性を示した。ここでは「ある程度大雑把に眺めれば，株価や債券，為替レートに見られるvolatility clusteringは，出来高に代表される市場のアクティビティにその起源を持っている」を改めて結論としたい。今後，より精緻な検証を進めていく予定である。

おわりに

本論でも触れたが，近年，金融現象に興味を持つ物理学者や工学者が増えている。この背景には，金融取引は，経済学でいうところの「需要と供給だけで価格

が決定される均衡市場」そのものであること，比較的短時間で膨大なデータ数を確保できることなどが挙げられる。そして，たとえば数十万個というティックデータを用い，経済的合理性といった問題をまったく抜きにして，その数に由来する信頼性を武器に統計的性質だけに注目した議論が展開されている。しかしながら，金融はあくまでも社会現象であって，それぞれの数値は互いに関連しあっており，因果の関係さえも単純ではない。それだけに，全体像をおおまかに捉え得る分析が先行しなければ，精緻な議論も生きてこないと考えたい。

〔参考文献〕

[1] 原田康平：「金融市場が示す volatility clustering」，久留米大学産業経済研究，第44巻，第2号，pp227-239，2003年。

[2] たとえば，次の資料に文献も含めて詳しい。
大本隆：「Volatility Skew を考慮したプライシングモデル」，京都大学経済研究所金融工学センター応用金融工学寄附研究部門シンポジウムレジメ，2003年(www.kier.kyoto-u.ac.jp/fe-nomura/symposium/moridaira_omoto/sympo_omoto01)。

[3] R.N. Mantegna and H.E. Stanley："Scaling Behavior in the Dynamics of an Economic Index", Nature, 383, pp46-49, 1995.

[4] 高安秀樹，高安美佐子：『エコノフィジックス』，日本経済新聞社，2001年。

[5] 原田康平：『経済時系列分析再考』，第5章，九大出版会，1998年。
GARCH，SV に関しては膨大な研究があり，いくつかを挙げるにとどめる。

[6] T. Bollelev："Generalized Autoregressive Conditional Heteroscedasdicity", Journal of Econometrics, 31, pp307-327, 1986.

[7] 渡部敏明：『ボラティリティ変動モデル』，朝倉出版，2000年。

[8] C.G. Lamoureux and W.D. Lastrapes："Heteroscedasticity in Stock Return Data: Volume versus GARCH Effects", Journal of Finance, 45, pp221-229, 1990.
出来高と株価の関係，出来高の自己相関などについても[7]に詳しく，その他，いくつかを挙げるにとどめる。

[9] A.R. Gallant, P.E. Rossi and G. Tauchen："Stock Price and Volume", Review of Financial Studies, 5, pp199-242, 1992.

[10] T.G. Anderen："Return Volatility and Trading Volume: An Information Flow Interpretation of Stochastic Volatility", Journal of Finance, 51, pp169-204, 1987.