数理解析研究所講究録
1066巻 1998年 221-238

情報理論的数理経済の試み

東京理科大学　松岡 隆志

概要

資本市場の計量分析、及びそのファイナンス理論、あるいはゲーム理論などの分野に情報理論の立場から新しい解析の視点、及び新しい指標を導入しようとする我々の試みを紹介する。ここでは、現在のファイナンス理論を中心にその主な基礎概念、解析手法を簡単に振り返り、そこに情報量という数理概念の必要性とその適用の可能性を論じる。また、実際にいくつかの情報量をファイナンス解析に適用した具体例を紹介する。

1. 序章 ー不確実性下の経済学ー

現実の社会における人間の行動様式において重要な問題の一つは、頻繁に直面する不確実な事態に対して、如何にその行動を決定するかという問題である。その不確実性は、主に状況に対する人間の不完全な知識と不完全な情報から生じるものと考えられる。自然科学が、ニュートンの決定論的な自然観の導入にもかかわらず、統計力学という確率的なものの見方を必要としたのはまさにその理由によるが、人間の行動やその結果から生じる事象ははるかに複雑であり、それ故に予想し難いものである。それ故、経済現象を数理的に解析していくということは、その経済現象に伴なう不確定さ、複雑さを数理的に解析していくということもできる。現在、こうした数理的経済研究と呼べる試みは、様々な視点から精力的に行われており、それらには、確率論や統計的手法をベースとする最近のファイナンス理論[1]やゲーム理論を用いたミクロ及びマクロ経済における経済主体の行動原理の研究[2]、あるいはカオス理論を適用した非線形動力学的な資本市場解析[3]などが挙げられる。

ところで、そうした最近の多くの数理的経済研究においても、「情報」という概念は非常に重要なものと位置づけられている。例えば、市場の経済主体に知られる情報の不均等さが個々の経済主体の利得（効用）にどういった影響を及ぼすかを考察するゲーム論的な市場の情報構造の研究などは、まさに市場における情報の価値を数理的に評価しようとする試みである。しかしながら、こうした研究の多くにおいては、あくまで「情報」は常識的な言葉の意味にとどまっており、そこに情報そのものを数量的に把握し解析するという考え方は導入されていないように思われる。そこで、我々が提案する情報理論的アプローチとは、情報理論の基本概念である幾つかの情報量（エントロピー、相互エントロピー etc.）を、さまざまな経済現象がもつ確率的特徴（不確定さ、複雑さ）を捉えうる指標として活用しようという企てである。

C. E. Shannon に始まる情報理論においては、情報の量的把握という抽象は、エントロピーという数理概念を用いてなされる。彼は、ある不確定さを伴なう系に対して、その系の持つ不確定さの解消がその系からの情報の獲得だと考え、不確定な事象系が有する情報

量（エントロピー）を、その不確定さ（その事象系が持つ確率分布）に依存する形で与えた。周知のように、エントロピーという概念は、熱現象の非可逆性を表す統計力学的な状態量としての側面に加え、Shannon の仕事をきっかけに、確率事象系の不確定さを計る情報量としての二つの側面から、その概念的基礎が深められ、今日、様々な分野に適用されている[4,5]。よって、この情報理論の汎用性の高さから、不確実性下の経済現象においても、その不確定さをエントロピーやいくつかの情報量を用いて定量的に解析できる可能性が存在するのである。

それでは、この情報理論を経済現象の解析にどのように適用できるだろうか。まず、情報理論の重要な基本概念図を以下に示し、その簡単な説明を試みてみよう。

情報理論の基本概念図

	入力状態 （初期状態）	→	チャネル	→	出力状態 （終状態）
	ρ	→	Λ^*	→	$\Lambda^*\rho$
情報量；	$S(\rho)$				$I(\rho;\Lambda^*)$

Shannon の情報理論は、雑音のある通信路を通して如何に効率よく情報を伝送するかという通信の問題に端を発しているが、その問題を議論するためには、二つの確率事象系（例えば、入力系と出力系）を用意し、また、上述のエントロピーに加え、二つの事象系の間の情報のやり取りを示す相互エントロピーという情報量が重要となる。このとき、入力系が持つ情報量が、入力状態ρ（古典系の通信システムを議論する場合は、入力情報源の状態を表現する確率分布が対応する。）のエントロピー$S(\rho)$であり、その入力状態ρは二つのシステム間の伝送システムを表現するチャネルΛ^*を通して出力状態$\Lambda^*\rho$へと変化し、相互エントロピー$I(\rho;\Lambda^*)$は、チャネルΛ^*を通して入力系から出力系へどれだけ正確に情報が伝わったかを示す量である。すなわち、$S(\rho)$と$I(\rho;\Lambda^*)$の関係を考察することでチャネルΛ^*の通信効率などが議論される。ここで、例えば、入力状態ρを何らかの物理的なシステムの初期状態と考えれば、チャネルΛ^*はその物理系の内在的な力学変化の要因や外部からの影響を考慮したダイナミクスを表すものとして用意することができ、よって、この情報理論の基本的枠組みを用いて、様々な系の状態変化のダイナミクスを情報量の変化という観点から解析することが可能となる。

例えば、株式変動を考察の対象に持つような経済システムにおいては、状態ρは株価変動の確率分布であり、チャネルΛ^*はマーケットの外生的、内生的要因などを反映するような推移確率行列が対応する（＊そういうチャネルを特定することは、非常に難しい問題ではあるが）。そのとき、株価変動の確率的特性を計量する指標としてエントロピー$S(\rho)$や相互エントロピー$I(\rho;\Lambda^*)$を計算することができる。

現在、我々は経済の問題に対して、まず大きく分けて、次の三つの視点から情報理論に

よるアプローチを試みようと考えている。

(1)　金融資産市場の不確実性（資産変動に伴う市場の確率的特性）を情報理論的観点から実証的に検証し[6,7,8]、資本市場の機能と資産の価格形成のメカニズムの解析を行い、ポートフォリオなどへの適用が可能な情報論的時系列モデルを考案する[9,10]。（いわゆるファイナンス理論の分野は、高精度な多量のデータが使用できる分野であり、数理的な研究の対象として非常に適している。）

(2)　金融資産の変動は、市場を構成する非常に多くの様々な経済主体の意思決定の結果によって生じる現象である。よって、その変動の因果関係は、経済主体の複雑な相互依存の関係のあり方に強く依存すると考えられるが、従来の多くの数理的手法においては、効率的市場仮説（ランダムウォーク）などの単純化（理想化）された仮説が用いられている。資産変動の理論的根拠を考察するためにも、従来のゲーム論的な市場モデル（例えば、Cournot のモデルなど）において、情報理論的な観点から経済主体の相互依存の有り様を解析する[11]。

(3)　一般均衡理論は、市場における多くの経済主体の相互作用によって、市場が一つの均衡状態を形成し、それに伴う均衡価格が形成されるということを厳密に数学的に示すものである。では、そうした均衡価格を具体的にどうやって見つけるかという問題は、非凸-非線型最適化問題の一つと考えることができるが、その一つの解法が 1960 年代に H. Scraf[12]等によって考案されている。ただし、Scraf のアルゴリズムには均衡解が複数存在する場合、その全てを求めることはできないという欠点がある。そこで情報理論の観点から、均衡解を求める効率的なアルゴリズムを開発する[13,14]。

(1) と (2) は非常に関係が深いが、それに比べれば (3) は現在のところ個別の問題と捉えることができる。

以下、（1）を中心にポートフォリオ問題を例に取りあげ、我々の情報論的アプローチの要点を解説してみよう。

今、時点t までのデータ（情報）を用いてあるポートフォリオP_tを組んだとすると、時点$t+1$における利得$R_{t+1}(P_t)$に対する効用の期待値は、効用関数をUとして次で与えられる。

$$E_{t+1}\left[U\left(R_{t+1}(P_t)\right)\right]=U\left(E_{t+1}\left[R_{t+1}(P_t)\right]\right)+F\left(\rho(P_t)\right)$$

ここで、$E_{t+1}[\cdot]$は時点$t+1$における期待値を時点t までの情報を基に計算することを示している。また、$\rho(P_t)$はポートフォリオP_tのリスクを評価する指標であり、$F\left(\rho(P_t)\right)$は$\rho(P_t)$の適当な関数である。ポートフォリオ戦略の決定は、目標期待利得を$E_{t+1}\left[R_{t+1}(P_t)\right]=M$と適当に設定し、その条件下で、期待効用$E_{t+1}\left[U\left(R_{t+1}(P_t)\right)\right]$が最大になるように$F\left(\rho(P_t)\right)$を最小にする$P_t^*$を選択することによって行われる。H. Markowitz は、ポートフォリオセレクションに関するその先駆的な研究において、投資家はその期待利得の最大化だけを基準にその行動を決定すべきでなく、分散にも注目すべきであることを指摘した[15]。彼は、ポートフォリオのリスク評価として分散を用いるためには、複数の証券がいかなる相関をもって変動するかを計測することが有効であると主張し、そのために共分散という指標を

導入したのである。すなわち、彼の方法に準じれば、$\rho(P_t)$はポートフォリオを構成する銘柄間の共分散行列、あるいは銘柄間の相関係数によって与えられる。

さて、以上のような従来のモデルは、次のような市場の確率的仮説を前提とする。すなわち、株価や収益率などの変動は正規分布に従う独立な確率過程（ランダムウォーク）に従うというものである。よって、我々は市場の効率性という仮定に基づいて、金融資産が従う分布を近似的に特定しその解を導く。しかしながら、果たして実際の資本市場においてランダムウォークは常に成立する仮定といえるだろうか？

この問いに対して、我々は、エントロピーやそれとは異なる情報量として状態のフラクタル次元[16,17]という指標を用いた新しい統計的解析手法を導入し、実証的な検証を行っている[6,7,8,9,10]。現在までに得られた結果（詳しくは3節を参照）は、日経225などのインデックスといわれる指標、あるいはSONY、NECなどの個別銘柄いづれにおいても、その株価変動には明らかな時系列相関が存在するというものである。すなわち、上記の指標を用いると、金融資産の変動に存在するある種の時系列相関をその不確定さの度合いに応じて数量的に計測することが可能になる。

よって、我々流の情報理論的なポートフォリオモデルへのアプローチとは次のような問題を解く事であると考えることができる。

(a) $E_{t+1}[\cdot]$を計算する確率分布を時点tまでの時系列相関の有り様から推定する。
(b) 各銘柄の時系列相関の類似性を評価できる指標として、$\rho(P_t)$を情報量を用いて新たに導入する。
(c) (a),(b)を踏まえ、情報論的ポートフォリオモデルを考案する。

金融資産の変動はその背後に多数の人間の意思決定というメカニズムを有するものであり、もし、我々が万能な神のごとき情報処理能力を持つとすれば、必要なパラメータを全て含んだ非線形な微分方程式を解く事によって、決定論的にその系列相関の有り様を理解し将来を予測することが可能である。しかし、それは現実においては当然不可能なことであり、よって、その一つの妥協案として経済の分野においても、自然科学同様、確率的な手法が大きな力を発揮する。系列相関の影響を最小限に押さえるという点で、ランダムウォークの仮説は確率的手法の便宜性を最大限に活用するものには違いないが、ここで、我々は情報量を用いれば、従来の確率的モデルの中にその系列相関の影響をより的確に反映したモデルの構築が可能となり、よってそれはより確からしい経済予測を与えるものになり得ると考えている。

以下、本稿では、状態のフラクタル次元を用いたファイナンス解析の結果[6,7,8,9,10]を幾つか報告する。

2. 状態のフラクタル次元の定式化

一般に、資本市場のカオス解析の一つの指標として採用されているフラクタル次元（相関次元）は、周知のように Mandelbrot が構築したフラクタル幾何学[18]において、幾何

学図形の持つ複雑さを定量的に示し得る指標とみなせるが、彼のフラクタル次元は図形に自己相似という性質を付加しない限り厳密には計算することはできない。すなわち、Mandelbrot のフラクタル次元を用いて現象を厳密に解析しようとすると、その現象に対応した自己相似集合を用意することが必要になる。よって、自然の持つフラクタルという性質を幾何の範疇の中だけで取り扱うのでなく、より多くの理論の中で議論していくためには、その概念の適切な一般化が有用となる。この目的のために、M. Ohya は、A. N. Kolmogorov の確率変数のε-エントロピー[19]を一般の量子力学系の状態に拡張することによって、Mandelbrot のフラクタル次元の概念を一般の状態空間上で定式化した[16]。

2. 1 定義

以下、離散的な確率空間（古典離散系）上の状態のフラクタル次元の定式化を簡単に振り返るが、最初に幾何学図形のフラクタル次元の一つである容量次元$d_C(X)$を、次に示しておこう。

$$d_C(X) \equiv \lim_{\varepsilon \to 0} \frac{\log N(\varepsilon)}{\log(1/\varepsilon)} \qquad (2.1)$$

ここで、$N(\varepsilon)$は、集合X（$\subset R^d$）を直径εの凸集合で被覆するのに必要なその凸体の最小個数である。情報理論の観点からすれば、(2.1) 式において、分子の$\log N(\varepsilon)$は距離空間上のε-エントロピー[20]と呼ばれる。確率空間上の状態のフラクタル次元は、確率分布に対するε-エントロピーを定式化することによって与えられる。

古典離散系はn個の事象からなる集合$X=\{x_1,x_2,\cdots,x_n\}$と、その事象の生起する確率分布$P=\{p_1,p_2,\cdots,p_n\}$$\left(\sum p_i=1, p_i\geq 0\right)$の組$(X,P)$で表される。$(X,P)$を完全事象系と呼び、確率分布$P$をその系の状態と呼ぶ。いま、二つの完全事象系$(X,P)$, (Y,Q)の複合事象系$X\times Y$の合成状態(i.e., 同時確立分布) Φを$\Phi=\{r(i,j);1\leq i,j\leq n\}$とする。ここで、入力空間$(X,P)$から出力空間$(Y,Q)$へ状態が変移するという視点に立てば、初期状態$P$から状態$Q$への推移確率行列$(p(j|i))$を用いて同時確率分布を表すことが可能である（i.e., $\Phi=\{r(i,j)\}=\{p(j|i)p_i\}$）。このとき、$(p(j|i))$は$P$を$Q$へ変換するチャネル$\Lambda^*$と考えることができ（i.e.,$Q=\Lambda^*P$ ($q_j=\sum_{i=1}^{n}p(j|i)p_i$))、初期状態$P$とチャネル$\Lambda^*$に関する相互エントロピー$I(P;\Lambda^*)$は、次で与えられる。

$$I(P;\Lambda^*)=\sum p(j|i)p_i \log \frac{p(j|i)\cdot p_i}{p_i \cdot q_j}$$

相互エントロピーは、入力状態Pに含まれる情報がチャネルΛ^*を通してどれだけ正確に出力状態Qに伝送されるかを示す情報量である。以下簡単のため、$n=m<\infty$として議論する。

いま、Cをチャネル全体の集合とし、状態Pに関するチャネルΛ^*の同値類$C(P;\Lambda^*)$を

$$C(P;\Lambda^*)\equiv\{\Gamma^*\in C;\Gamma^*P=\Lambda^*P\}$$

と与える。このとき、状態Pのε - エントロピー$S(P;\varepsilon)$は、

$$S(P;\varepsilon) \equiv \inf\left\{ J\left(P;\Lambda^*\right); \Lambda^* \in C, \left\| P - \Lambda^* P \right\| \leq \varepsilon \right\}$$

ただし、$\|P-Q\| \equiv \sum_{i=1}^{n} |p_i - q_i|$かつ、

$$J\left(P;\Lambda^*\right) = \sup\left\{ I\left(P;\Gamma^*\right); \Gamma^* \in C\left(P;\Lambda^*\right) \right\}$$

$J\left(P;\Lambda^*\right)$は極大相互エントロピーと呼ばれるが、状態$P$を状態$Q$に変化させるチャネルはただ一つとは限らないため、その相互情報量が最大になるもので与えられる。通信理論では、相互エントロピーは入力状態Pから出力状態Qへ移すことのできる情報量を示すものとしてその通信効率などが議論されるが、よって、我々はその移すことのできる情報量は大きい方を良しとするのである。また、ε-エントロピー$S(P;\varepsilon)$は、状態Pからε近傍の別の状態Q（i.e., $\|P-Q\| \leq \varepsilon$）に最低限移すことのできる情報量を表わしている。

この状態のε-エントロピーを用いて古典離散系の状態のフラクタル次元が定義される[16]。

［定義 2-1］

（1）オーダーεの容量次元；　$d_C(P;\varepsilon) \equiv \dfrac{S(P;\varepsilon)}{\log \frac{1}{\varepsilon}}$

（2）オーダーεの情報次元；　$d_I(P;\varepsilon) \equiv \dfrac{S(P;\varepsilon)}{S(P)}$

ただし，$S(P)$は状態Pのエントロピー（i.e. $S(P) = -\sum_{i=1}^{n} p_i \log p_i$）。

さて，いま，チャネルΛ^*に対応する推移確率行列$(p(j|i))$で，$\left\|P - \Lambda^* P\right\| = \varepsilon$のとき，

$$\sum_{i=1}^{n} p(j|i) p_i = \begin{cases} q_k = p_k + \dfrac{\varepsilon}{2} (1 \leq k \leq n) \\ q_j = p_j - \varepsilon_j (j \neq k) \end{cases} \qquad (2.2)$$

$$\left(\sum_{j \neq k} \varepsilon_j = \varepsilon,\ 0 \leq \varepsilon_j \leq \varepsilon/2 \right)$$

と与えられるものを考える。

＜定理 2.2＞[21,22]　チャネルΛ^*を式(2.2)で与えられるチャネルの集合に限り、かつ距離εが$\varepsilon \leq \min\{p_i\}$を満たすとき、

$$S(P;\varepsilon) = S(P) - \left(p_{\max} + \frac{\varepsilon}{2}\right) \log\left(p_{\max} + \frac{\varepsilon}{2}\right) + p_{\max} \log p_{\max} + \frac{\varepsilon}{2} \log \frac{\varepsilon}{2(n-1)} \qquad (2.3)$$

ここで、$p_{\max} = \max\{p_1, \cdots, p_n\}$。

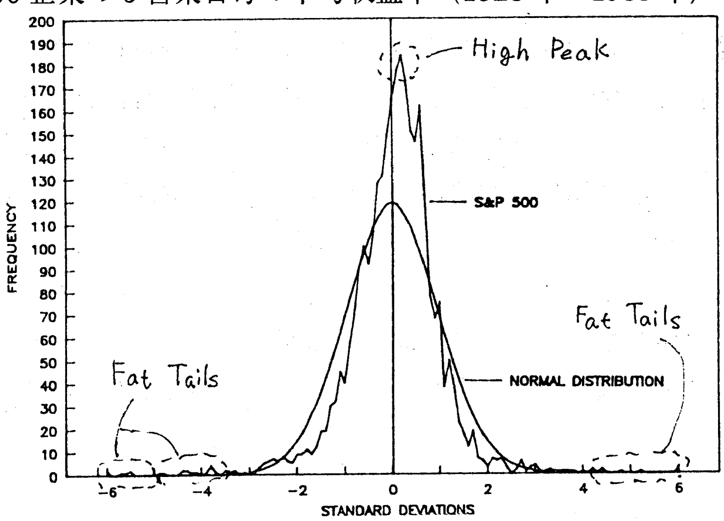
3．資本市場の確率的特性

株価などの確率変数が、ランダムウォークに従うとする理論は、効率的市場仮説を前提

とする。効率的市場仮説とは、市場を構成するプレーヤー（投資家など）が合理的であるという前提に基づき、基本要因に関する情報や過去の価格データなどの公開情報が全て瞬時かつ適切に市場に反映され均衡価格が形成されるというものである。よって、今日の価格の変化は予期できない今日のニュースにのみ依存し、過去の値に依存しない独立な変動とみなせる。すなわち、収益率の時系列データは十分な数のデータが収集されれば、中心極限定理からその極限において、それは正規分布に従う独立な確率過程となるのである。1900 年の L. Bachelier[23] に始まるとされるランダムウォークの理論は、1940 年代まではその先見性ゆえの停滞を見せたものの 1950 年以降多くの実証的研究も提出され[24]、1980 年代には効率的市場仮説は一般的な事実として広く承認されるようになった。実際、様々な経済変数がランダムウォークに従うという仮定は、その線形性ゆえに統計的解析手法を容易にし大きな力を発揮する。

その一方で、収益率の時系列データから得られる確率分布の正規性、独立性に関して疑問を投げかける実証的な研究もまた数多く存在する。B. B. Mandelbrot は、1880 年から 1958 年にわたるアメリカの木綿市場を調べ、その価格は Levy law の下で変動する可能性を示した[25]。彼は、そうした実証的ないくつかの研究を背景に、確率変数が独立な正規分布に従わないとする場合のボラティリティーの評価を R/S 解析という手法を用いて明らかにできることを提言し、収益率の確率分布が無限分散という特徴を持つ安定パレート分布のグループに属する可能性を示唆した[26]。Mandelbrot 以降、Levy law に関する実証的な研究には、日次の収益率に対する E. J. Fama の解析[27]を始めとして、最近のものでは R. B. Olsen のグループ[28]等が挙げられる。こうした研究の多くは、資本市場の確率分布に平均近くの高いピーク（high peak）と分布の袖のより多くの観察値（fat tail）が存在することを示すものであり、市場の収益率の分布が近似的に正規でかつ独立であるという仮定には、疑問の余地があることを示唆しているようにも思われる。例えば、図 3-1 は、E. E. Peters によるアメリカ市場の 500 企業の平均収益率（今日の株価を過去（例えば前日）の株価で割った値）の頻度グラフ（1928 年 1 月～1989 年 12 月）と正規分布の違いを示したグラフである[3]。

図 3 -1 500 企業の 5 営業日毎の平均収益率（1928 年～1989 年）の度数



実際の収益率の頻度には、平均付近での high peak と平均から大きく離れた fat tail が見て取れる。Peters は、R/S 解析によってこのような確率分布の特徴を持つ収益率の変動には 4 年間という長期の相関が存在する可能性を示した。彼はこれらの検証を通して、市場はある種のトレンドに従って長期的に循環するシステムであると主張している。

以上、我々は二つの視点に立って資本市場における数理研究の流れを概観してきたが、効率的市場仮説の是非に関して、そのどちらかの立場だけを強硬に主張することには、あまり多くの実りがあるとは思われない。従来の効率的市場仮説を支持する研究者の中においては、より柔軟な視点として、収益率の分散が無限であることは認めないものの、分布の high peak という特徴は正規分布の関連性に於いて既存の理論の枠内で検討される必要性を認める立場もある。また、最近の数理ファイナンス理論では、市場の確率的特性をより忠実に再現するモデルの構築が様々な視点から行われており、必要に応じて正規性の仮定と非正規性の特徴をうまく組み合わせるといった試みの有用性も示されつつあるように思われる。要は、市場に存在する確率的な特徴を如何に計量的に適切に把握し、それを実際の証券市場や資産運用においてどれだけ有用に適用できるかということにある。

以下、我々は状態のフラクタル次元という指標を用いて、収益率分布の high peak 等に代表される特徴を直接定量的に解析した結果を紹介する [6,7,8] 。

3. 1 状態のフラクタル次元から捉えうる株価変動の時系列相関と階層構造[6,7,8]

本稿で対象とするデータは、次の２種類。

（１）日経 225 および、NEC、TOYOTA、SONY の８３年１月から９５年１２月までの毎日の終値。

（２）東証１部上場からランダムに選出した４０銘柄の８７年４月から９７年１１月までの毎日の終値。

以上のデータに対して、次で定義される対数収益率を計算し、その頻度分布を特定する。

日次対数収益率＝Log（今日の終値／前日の終値）

日次対数収益率は、今日の終値と前日の終値を比較して得られるが、その比較する過去の株価との計測期間を、例えば、５日前とすればそれを週次対数収益率と呼び、その計測期間は、一般に、月次（２０日前株価の終値との比較値）のものが多く使われているようである。これは、効率的市場仮説を前提とすれば、その計測期間の違いに関わらず、得られる分布は同一の正規分布と近似的に了解できる事、および、日次、週次クラスのデータの量は非常に膨大になるという２点が挙げられる。しかし、前述したように、効率的仮説が常に成立するとは限らないものとすれば、計測期間の違いという視点も必要になってくるはずである。そこで、我々は対数収益率に対し、その頻度分布を次の二つの視点から特定した。

＊分布を特定する期間（標本数）を変化させ、分布の持つ情報量（情報次元）の推移を考察する。
＊対数収益率の計測期間を変化させ、分布の持つ情報量（情報次元）の推移を考察する。

今回の解析では、上記の方法で得られた分布に対して、式(2.3)で与えられる ε - エントロピーの評価式を用い、収益率のフラクタル次元としてオーダー ε の情報次元を計算した。

図 3.2 は計測期間 20 日の日経 225 の収益率分布（i.e.,月次分布）の標本数を 83 年 1 月から 95 年 12 月まで、順次 1 年間づつ増やしていったときの分布変化に対する情報次元の推移、および（平均 0、分散 1）の正規乱数を独立に発生させた系列の累積分布の標本数変化に対する情報次元の推移を示したグラフである。ここで、正規乱数分布の 1 年間に相当する標本数は、日経 225 の累積分布の標本数との対応から 260 個としている。また、グラフの縦軸は情報次元の絶対値ではなく、各標本数の情報次元の値から平均値を計算し、その平均値からの差として表している。また、今回、我々が得られた全ての収益率分布において、定理 2.2 の条件 $\varepsilon \le \min\{p_i\}$ を満たす ε の値の最大値は、$\varepsilon = 0.000229$ であったので、以下で示される情報次元の値は、全てオーダ $\varepsilon = 0.000229$ の情報次元の値である。

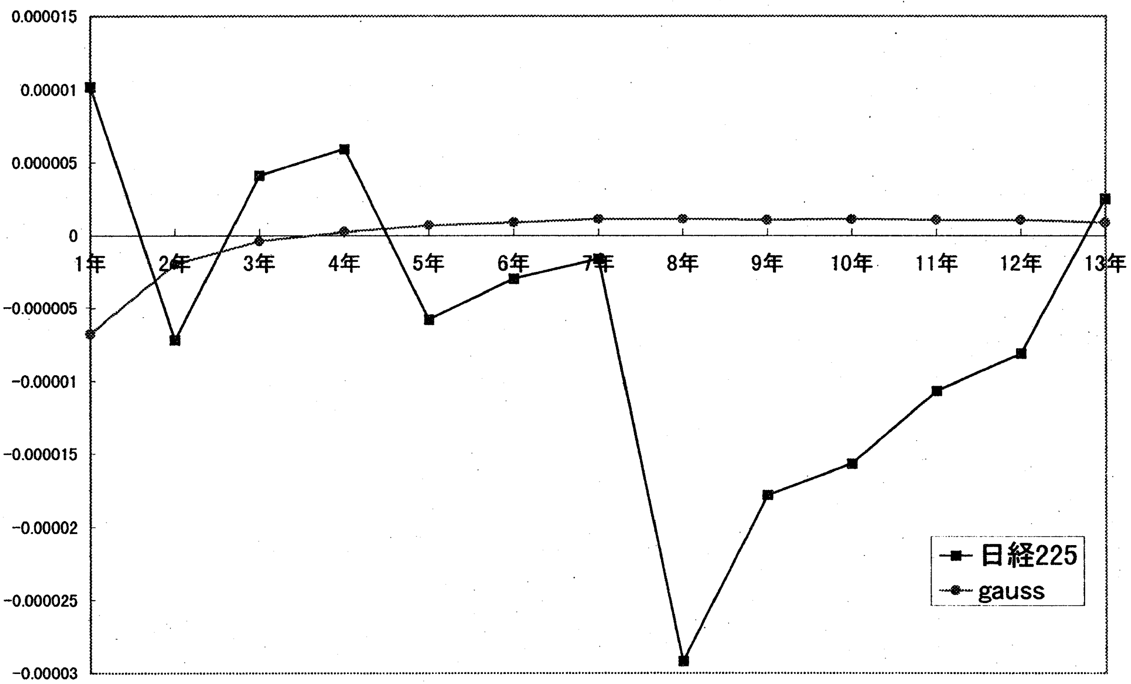


図 3.2　日経 225 および正規乱数の標本数変化に対する情報次元の推移

正規乱数から得られる分布の情報次元は、標本数の増加に伴ってほぼ単調に増加し、その変動幅は小さくなっていく。この結果は、情報次元が確率分布の持つ複雑さを定量的に示すものだと考えれば、非常に自然な結果である。つまり、分布を特定する標本数を増やすことによって次第に分布の偏りが平均化され、分布の持つ不確定さは増大し、よって情

報次元も増大する。さらに、その値がある変動幅に収まるのは、中心極限定理が示すように、標本数の増加に伴いその累積分布は正規乱数が従う正規分布に収束していくからと考えられる。仮に、市場の価格変動が同一の正規分布に従う独立な確率過程（以下、i.i.d.と呼ぶ。）であるならば、日経 225 から得られる収益率分布に対しても、その情報次元は正規乱数と同様の変化を示すはずである。ところが、日経 225 の情報次元の推移は正規乱数とは明らかに異なる変化をしている。よって、日経 225 の系列を i.i.d.とみなすことは難しいように思える。そこで、日経 225 系列の非独立性をより確かに検証するために、次のようなスクランブルテストを行った。株価の観察値を実際に生じた時系列の順序と全く異なるようにランダムに混ぜ合わせ、そのランダムな系列から月次の収益率分布を特定し情報次元を計算した。その結果を正規乱数のものと比較したのが図 3.3 である。このグラフも、図 3.2 と同様にその値は平均値からの差で示してある。

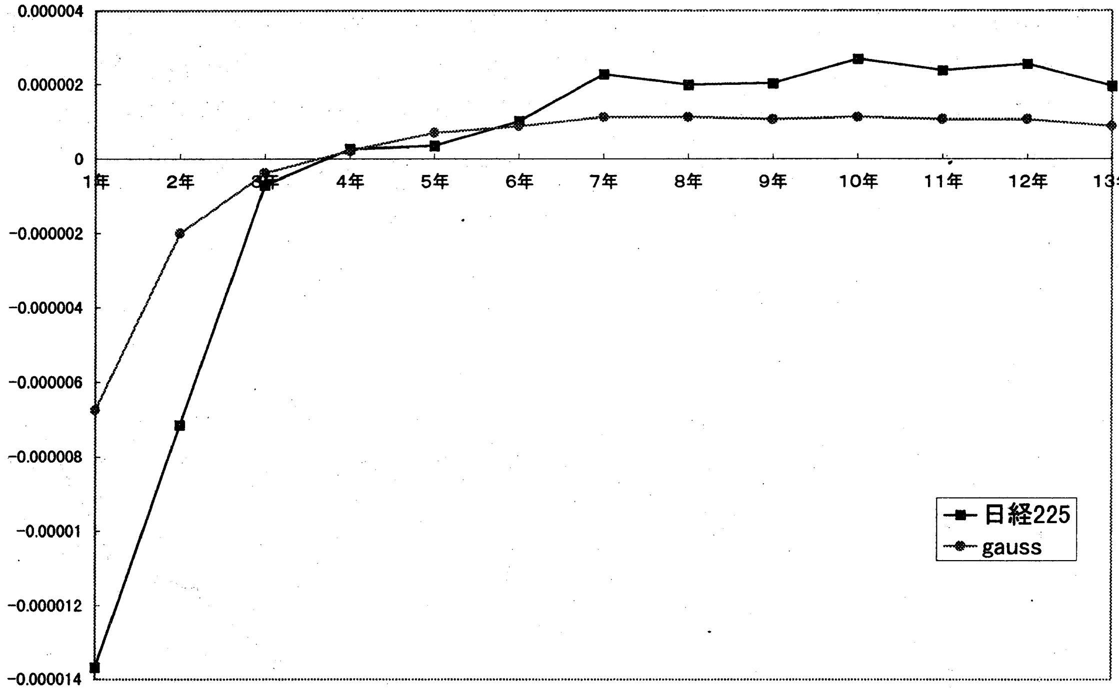


図 3.3　スクランブル後の日経 225 の情報次元の推移

日経 225 の系列が i.i.d.であれば、観測値の順序に相関がないからデータをスクランブルしてもその得られる分布の特徴に変化は見られないはずである。しかし、図 3.3 の結果は、スクランブルの前とスクランブルの後では情報次元の推移が大きく異なることを示している。また、スクランブル後の日経 225 の情報次元の推移は、独立な正規乱数の変化と同様の単調増加な変化とみなせる。以上の結果は、株価の系列が生起する順序には何等かの時系列相関が存在しており、その系列相関が観測値をスクランブルすることによって壊され独立な系列に変化したと考えれば説明できる。

次に、情報次元が株価の収益率分布によく見られる「high peak」という分布の持つ局所

的な特徴をよく評価する指標であることを示そう。図 3.4、3.5、3.6 は、NEC、TOYOTA、SONY の月次の収益率に対して、その収益率分布の標本数増加に伴う、high peak、エントロピー、情報次元の変化の様子を比較したものである。

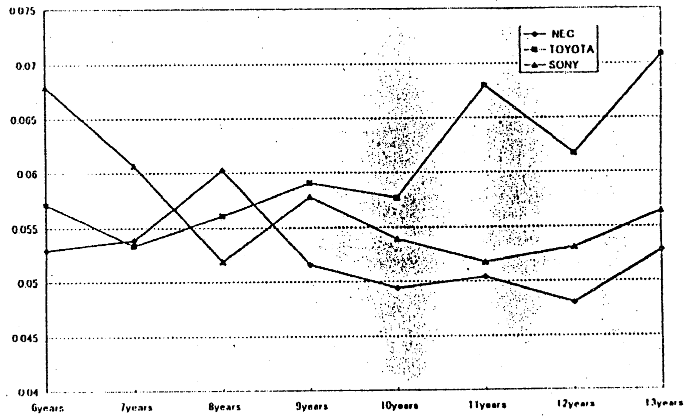


図 3.4 月次収益率分布の標本数変化に対する high peak の変化

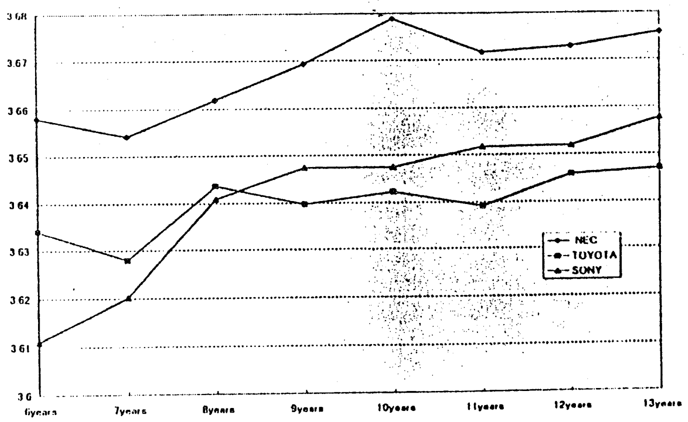


図 3.5 月次収益率の標本数変化に対するエントロピーの変化

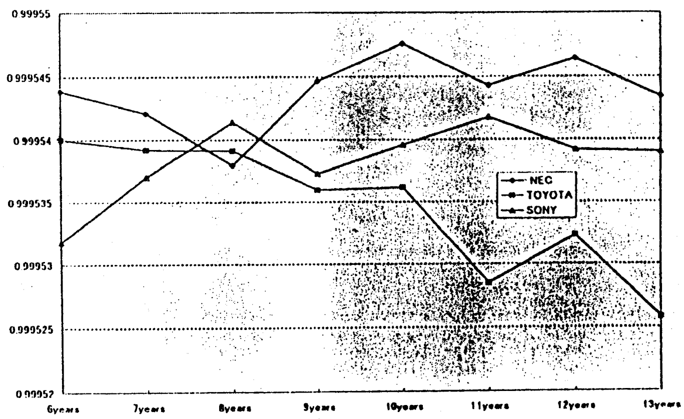


図 3.6 月次収益率の標本数変化に対する情報次元のに変化

high peak のオーダーとエントロピー、情報次元のオーダーがそれぞれ常に一致しているわけではないが、8 年間の分布を中心に一度 3 銘柄の high peak の値が接近し、その後離れていくという変化の様子は、明らかに情報次元の方がエントロピーに比べ同様の変化

をしていることが分かる。すなわち、収益率の分布に現れる high peak という特徴を情報量という視点から解析するには、情報次元の方がエントロピーより適していると考えることができる。これは、確率分布の high peak という局所的な性質を情報次元は良く評価するということであり、情報次元はエントロピーとは異なる側面から確率分布の複雑さを計量するのである。

いま、株価の系列に時系列相関が存在するということを前提にすれば、high peak という特徴は平均近くの事象が独立な系列よりもより多く生起するという意味で、過去との相関の度合いを示す典型的な特徴とも考えられる。この前提に立てば、情報次元が分布の high peak をうまく評価することから、その系列相関の度合いをよく計量しうるとの見方も可能であろう。

次に、分布の計測期間を変化させて情報次元の推移の様子を考察する。図 3.7 は、日経 225 の収益率分布の計測期間を、1 日、5 日、10 日、15 日、20 日、４０日と変化させたときの情報次元の推移の様子を示したグラフである。ここで、各計測期間の情報次元の値は、分布の標本数が 6 年間から１３年間分のデータでそれぞれ計算した平均値である。

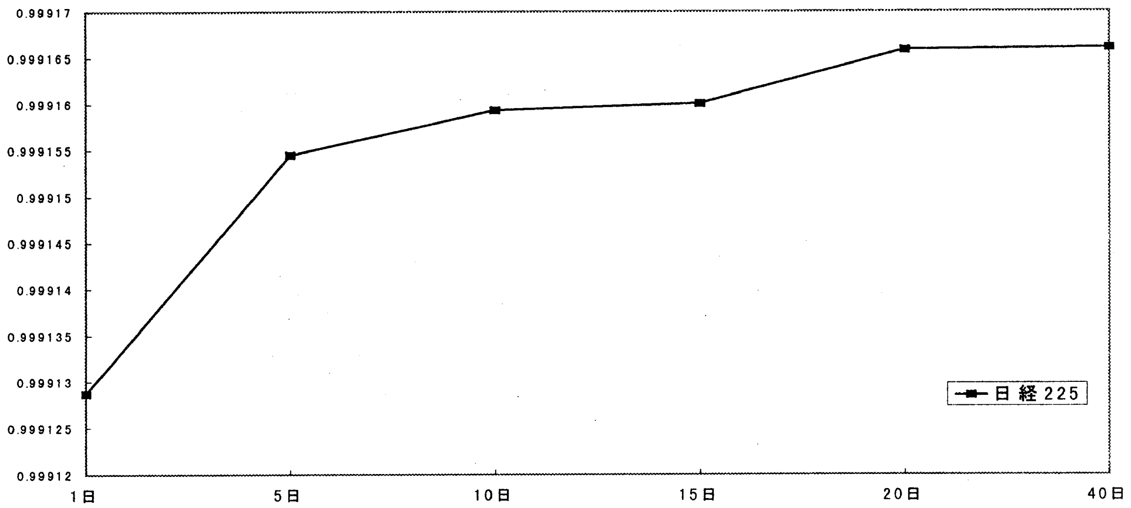


図 3.7　計測期間の変化に伴う情報次元の推移

情報次元の値は、分布の計測期間を長くするにつれて増大する傾向が見られる。つまり、系列相関の度合いが次第に弱まり、株価の変動がランダムになっていく様子を示している。これは、時間の経過とともに過去の株価が現在の株価に及ぼす影響が次第に弱くなっていくということで、ある意味で自然な結果である。収益率分布の計測期間の変換は、一種の時間に対するスケール変換と見ることもできるから、そのスケールの違いによって生じる情報次元の変化は、株価変動の階層構造的な時系列相関の有り様を示すものと考える事ができ、その株価の系列が従うダイナミクスの特徴を示すものと言える。

図 3.8 は、バブル崩壊の時期が、日経 225 のダイナミクスの構造変化によって特徴付けられることを示したグラフである。

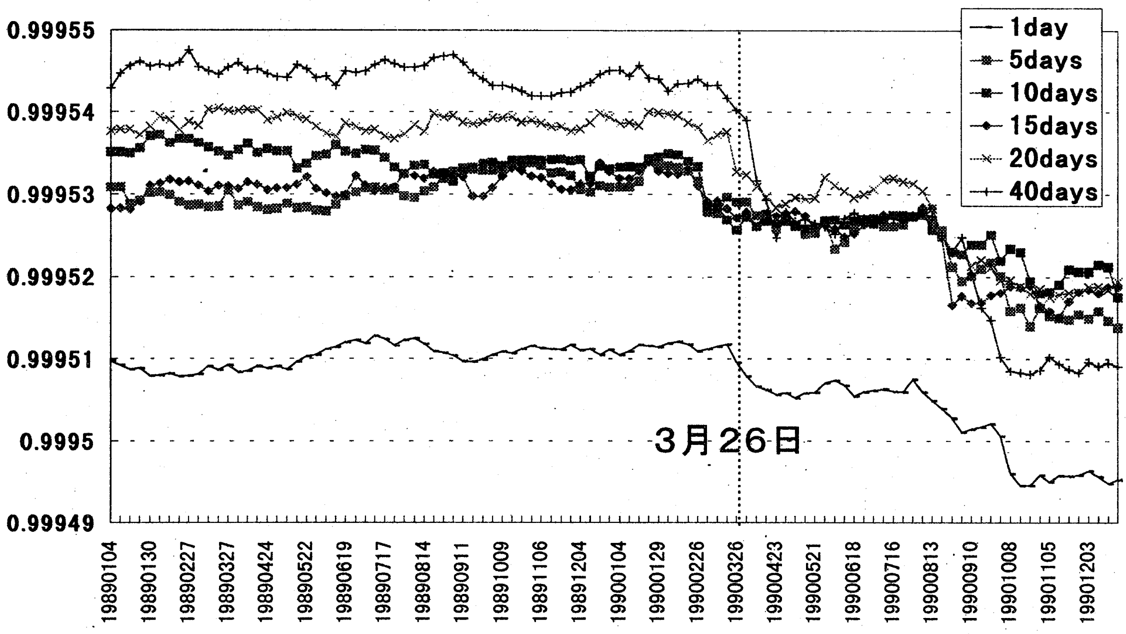


図 3.8　情報次元を用いたバブル崩壊の特徴付け

図 3.8 では、バブル前後の情報次元の変化の様子を詳しく解析するために、計測期間が 1 日、5 日、10 日、15 日、20 日、40 日の分布それぞれに対し、標本数を 1 週間分づつ増やした時の分布変化に対する情報次元の推移を示している。1990 年 3 月中旬から下旬を境に、時系列相関の階層構造の有り様が大きく変化しているのが分かる。3 月下旬後の情報次元の推移においては、計測期間が長くなるほどその情報次元の値も大きくなるという自然な構造が壊れているように見える。バブル崩壊という現象が、株価のダイナミクスそのものに変化をもたらしていると考えることもできるし、逆にダイナミクスに何らかの変化が生じたことによってバブル崩壊という現象がもたらせたとみることもできよう。

以上、本節では、状態のフラクタル次元が、株価変動に存在する時系列相関を計量しうる指標となることを示し、株価変動のダイナミクスの階層構造的な側面を定量的に特徴付けられることを解説した。次節では、個々の銘柄株の系列相関の階層構造の類似性（すなわち、個々の銘柄株のダイナミクスの類似性）という視点から銘柄株の分類を行った結果を紹介する[9,10]。

4. 情報次元を用いた株価変動の分類とその評価[9,10]

本節で対象とするのは、以下に示す４０銘柄である。

銘柄名	**鹿島、清水建設、青木建設、アサヒビール、** **鐘紡、帝人、旭化成工業、日本通運、** **信越化学工業、協和発酵工業、塩野義製薬、富士写真フィルム、** **日本石油、ブリヂストン、新日本製鐵、川崎製鉄、**

	神戸製鋼所、三菱マテリアル、コマツ、東芝、 NEC、ソニー、日産自動車、トヨタ自動車、 本田技研工業、キャノン、凸版印刷、三菱商事、 三越、富士銀行、三菱信託銀行、東京海上火災保険、 野村證券、三菱地所、東武鉄道、日本通運、 日本郵船、全日本空輸、東京電力、東京ガス

4.1 銘柄分類

我々は、情報次元によって捉えられる株価変動の特徴を反映させた分類を行いたい。そこで、各対象銘柄に対して、87 年 4 月 1 日から 95 年 3 月 31 日までのデータを用いて収益率分布の計測期間を１日から２００日まで１日ごとに変化させた分布の情報次元を計算し、その情報次元の系列の類似性をクラスター分析により分類する。ここで、グループ間の距離の与え方は最長距離法、クラスター間の距離には相関係数を用いる。また、情報次元を用いた分類の比較として、従来の手法、すなわち、任意の計測期間の株価変動系列（ここでは、２０日の株価変動系列）の相関係数を用いて同様にクラスター分析で分類する。

図 4.1 は、SONY の株価系列に対して、1 日から 200 日までの計測期間の変化に伴う情報次元系列を計算した結果である。

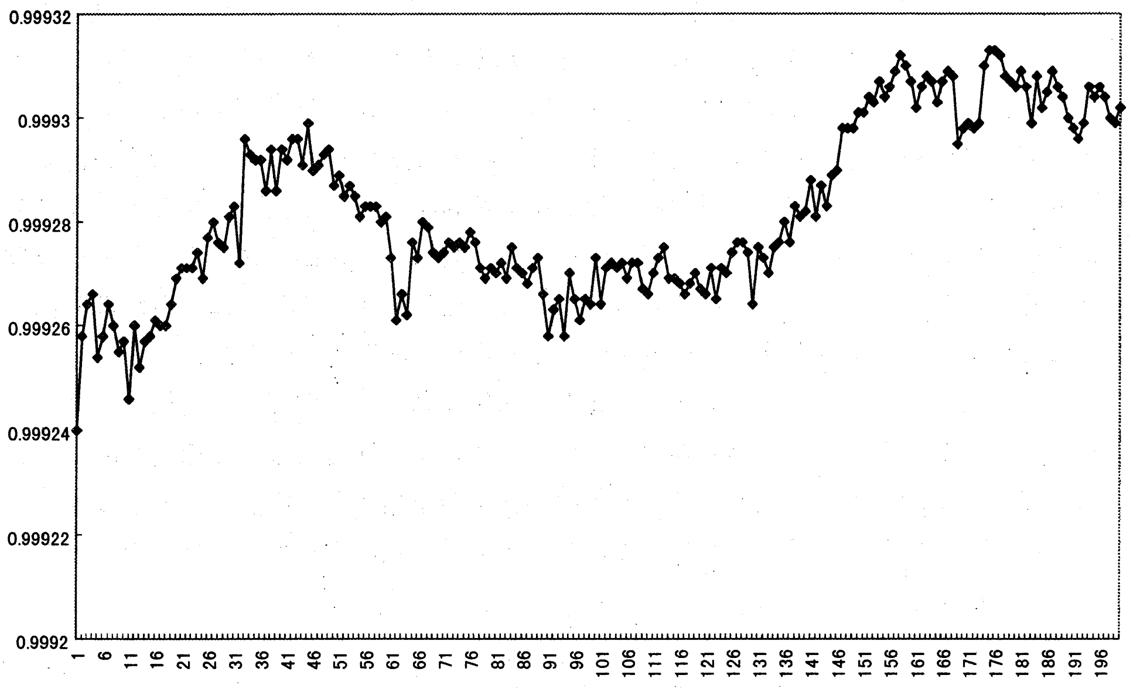


図 4.1　計測期間の変化に対する SONY の情報次元の推移

SONY の情報次元の値は、計測期間が長くなるにつれて単調に増加するというよりは、周期的な変動をしているように見える。これは、ある程度時間が経過すると、再び過去の株価の影響が強くなるといった長期の系列相関の存在を示唆するものであり、系列相関の有り様に二重の意味での階層構造が存在しているとも考えれらる。以上のような系列相関

の階層構造は、個々の銘柄によって微妙にあるいは大きく異なっており、情報次元によるクラスター分析はこのような系列相関の類似性を反映したものと考えることができる。

情報次元および従来の手法を用いた分類は、それぞれ次のように大きく3つのグループに分類される。

＜情報次元による分類＞

1st Group	鹿島、日本通運、東京海上火災保険、三菱信託銀行、凸版印刷、三菱地所、川崎製鉄、神戸製鋼所
2nd Group	京成電鉄、清水建設、富士写真フィルム
3rd Group	塩野義製薬、本田技研工業、ブリジストン、全日本空輸、アサヒビール、三越、日産自動車、キャノン、NEC、トヨタ自動車、青木建設、東芝、富士銀行、野村證券、新日本製鐵、鐘紡、東武鉄道、帝人、三菱マテリアル、旭化成工業、日本石油、三菱商事、協和醗酵工業、東京ガス、ソニー、コマツ、日本郵船、信越化学工業、東京電力

＜従来の手法による分類＞

1st Group	青木建設、清水建設、鹿島、旭化成工業、コマツ 京成電鉄、東武鉄道、日本通運、全日本空輸、日本石油 日本郵船、帝人、三菱マテリアル、鐘紡、三菱商事 東京ガス、三菱地所、野村証券、三菱信託銀行、東京海上火災保険 東京電力、富士銀行、アサヒビール、三越
2nd Group	東芝、日産自動車、新日本製鐵、川崎製鉄、神戸製鋼所、
3rd Group	凸版印刷、信越化学工業、ブリジストン、協和醗酵工業、塩野義製薬、NEC、トヨタ自動車、本田技研工業、キャノン、富士写真フィルム、ソニー

4.2 ポートフォリオリターンによる銘柄分類の評価

情報次元による分類と従来の手法による分類から、簡単なポートフォリオを作成し、そのリターンを考察することにより、両者の分類の比較を行う。

以下の手順でポートフォリオを構築する。

（1）各銘柄から1銘柄ずつ選択する。
（2）等ウェイトポートフォリオを構築する。

このようなポートフォリオのリターンを、95年4月から97年11月まで1ヶ月ことに月末時点で計測し、それぞれの分類の全ての組み合わせから作られるポートフォリオリター

ンの平均を比較する。

以下に示した表がその結果である。平均μは考えられる全てのポートフォリオのリターンの平均、標準偏差σはその平均からのばらつきを示す。

期間	i.i.d. を仮定した場合の分類				情報次元による分類				
	平均μ	標準偏差 σ	$\mu-\sigma$	$\mu+\sigma$	平均μ	標準偏差 σ	$\mu-\sigma$	$\mu+\sigma$	
9504-9505	-0.0653	0.0195	-0.0848	-0.0458	-0.0567	0.0201	-0.0768	-0.0366	○
9504-9506	-0.1117	0.04	-0.1517	-0.0717	-0.1058	0.0452	-0.151	-0.0606	○
9504-9507	0.0011	0.0435	-0.0424	0.0446	0.0096	0.0452	-0.0356	0.0548	○
9504-9508	0.1035	0.0529	0.0506	0.1564	0.1105	0.0409	0.0696	0.1514	○
9504-9509	0.1114	0.0522	0.0592	0.1636	0.1084	0.0419	0.0665	0.1503	○
9504-9510	0.0916	0.0575	0.0341	0.1491	0.0804	0.0496	0.0308	0.13	×
9504-9511	0.1391	0.0523	0.0868	0.1914	0.1348	0.0427	0.0921	0.1775	○
9504-9512	0.1976	0.0668	0.1308	0.2644	0.2065	0.057	0.1495	0.2635	○
9504-9601	0.2363	0.0697	0.1666	0.306	0.2462	0.0563	0.1899	0.3025	○
9504-9602	0.1989	0.0729	0.126	0.2718	0.2195	0.0559	0.1636	0.2754	○
9504-9603	0.2478	0.0725	0.1753	0.3203	0.2917	0.0689	0.2228	0.3606	○
9504-9604	0.2827	0.0702	0.2125	0.3529	0.3184	0.0617	0.2567	0.3801	○
9504-9605	0.2697	0.0761	0.1936	0.3458	0.3237	0.0661	0.2576	0.3898	○
9504-9606	0.2984	0.0882	0.2102	0.3866	0.3424	0.0777	0.2647	0.4201	○
9504-9607	0.218	0.0827	0.1353	0.3007	0.2518	0.0768	0.175	0.3286	○
9504-9608	0.1872	0.0816	0.1056	0.2688	0.2142	0.0792	0.135	0.2934	○
9504-9609	0.2498	0.0898	0.16	0.3396	0.2797	0.0847	0.195	0.3644	○
9504-9610	0.1933	0.0949	0.0984	0.2882	0.2278	0.0917	0.1361	0.3195	○
9504-9611	0.2193	0.1056	0.1137	0.3249	0.2236	0.1144	0.1092	0.338	×
9504-9612	0.1634	0.1159	0.0475	0.2793	0.1446	0.139	0.0056	0.2836	×
9504-9701	0.0876	0.1391	-0.0515	0.2267	0.0731	0.1638	-0.0907	0.2369	×
9504-9702	0.1015	0.1537	-0.0522	0.2552	0.0846	0.1848	-0.1002	0.2694	×
9504-9703	0.1055	0.1621	-0.0566	0.2676	0.0584	0.2033	-0.1449	0.2617	×
9504-9704	0.1644	0.1778	-0.0134	0.3422	0.1048	0.2446	-0.1398	0.3494	×
9504-9705	0.1658	0.1766	-0.0108	0.3424	0.1409	0.2184	-0.0775	0.3593	×
9504-9706	0.2137	0.1844	0.0293	0.3981	0.1751	0.2248	-0.0497	0.3999	×
9504-9707	0.1806	0.2358	-0.0552	0.4164	0.1288	0.2769	-0.1481	0.4057	×
9504-9708	0.1039	0.2402	-0.1363	0.3441	0.0688	0.2692	-0.2004	0.338	×
9504-9709	0.0541	0.2844	-0.2303	0.3385	0.0178	0.3222	-0.3044	0.34	×
9504-9710	-0.0288	0.2529	-0.2817	0.2241	-0.0686	0.2941	-0.3627	0.2255	×
9504-9711	-0.0682	0.3156	-0.3838	0.2474	-0.1316	0.3658	-0.4974	0.2342	×

(注) 最終列 ○：情報次元の ($\mu-\sigma$) の方が大きい。△：同じ。×：情報次元の ($\mu-\sigma$) の方が小さい。

この表から、標準偏差について比較すると、約1年半までは情報次元による分類の方が小さいケースが多い。すなわち、各ポートフォリオのリターンが平均近くに集まり同様なパーフォマンス特性を持つことを示している。また、マイナスのばらつきを考慮した上でのリターン$(\mu-\sigma)$を比較すると、約1年半までは情報次元が高いケースが多い。これらの結果は、ポートフォリオの銘柄選択をする場合、情報次元による分類が従来の手法を用いた分類よりも、約1年から1年半ぐらいのスパンでポートフォリオの銘柄選定に依らず、リターンに関して類似したポートフォリオを構築できることを示唆している。すなわち、情報次元による分類が従来の分類よりも、選択する際の各グループの銘柄がより類似した時系列相関構造（類似したダイナミクス）を持つので、各グループから任意の銘柄を取り出して構築されるポートフォリオもまた同様のパーフォマンスが実現したみることができる。ただし、このような傾向は、約1年半くらいのスパンに対するものであり、それ以降のリターンの評価は、逆に従来の手法の方がよいという結果が得られている。このことに

関しては、個々の銘柄の株価変動を特徴付けるダイナミクスそのものが変動していると考えれば、短期の未来では情報次元を用いた分類によってダイナミクスの類似性がある程度は捉えきれるものの、ダイナミクスが変化するので、次第に過去の時点での分類では銘柄の類似性をうまく捉えきれなくなるためという見方ができる。よって、（今回の結果によれば）1 年から 1 年半くらいの期間でそのつど新しい過去のデータを含めて情報次元による分類を更新していけば、従来の手法よりも恒常的にリターンの評価を上げることも可能であると考えられよう。

参考文献

[1]S. Figlewski, W. L. Silber & M.G. Subrahmanyam (Editor) “Financial Options: From Theory to Practice”, IRWIN, New York (1990).

[2]R. Gibbons ; “Game Theory for Applied Economists”, Princeton University Press (1992).

[3]E. E. Peters ; “Chaos and Order in the Capital Markets”, John Wiley & Sons, New York (1991).

[4]大矢雅則、今井秀樹、小嶋 泉、中村八束、廣田正義 編集“数理情報科学辞典”、朝倉書店（1995）.

[5]R. S. Ingarden, A. Kossakowski & M. Ohya ; “Information Dynamics and Open Systems”, Kluwer (1997).

[6]安藤隆宏、大矢雅則、松岡隆志；“状態のフラクタル次元を用いた株価変動解析－資本市場解析における新しいアプローチー”、信学技法 96、pp.37-42 (1996).

[7]児澤秀幸、松岡隆志、大矢雅則；“状態のフラクタル次元による株価変動解析－効率的市場仮説に対する情報力学的検証の試みー”、「第 20 回情報理論とその応用シンポジウム」Proceedings、pp.(1997).

[8]T. Matsuoka & M. Ohya, “Measurement of time serial correlation in stock price movement by using fractal dimension of a state”, in preparation.

[9]磯貝明文、松岡隆志；“状態のフラクタル次元による銘柄分類とその分類の利用可能性について”、MTEC10 周年記念論文集に掲載予定。

[10]A. Isogai, T. Matsuoka, M. Ohya; “New fractal method for classification of stock returns and its application to portfolio selection problem”, in preparation.

[11]大矢雅則、松岡隆志；“複占市場における情報獲得と相互エントロピー”、京都大学数理解析研究所講究録、1013、pp.28-40 (1997).

[12]H. Scarf ; “The approximation of fixed points of a continuous mapping “,SIAM J. Applied Mathematics 15, pp.1328- (1967).

[13]T. Matsuoka & M. Ohya ; “Simulated annealing and its application to Cobb-Douglas economic model”, Open Systems & Information Dynamics 3, pp. 357-368 (1995).

[14]奥村隆洋、松岡隆志、大矢雅則；“Scarf アルゴリズムの改良による多重均衡解の探索”、「第 18 回情報理論とその応用シンポジウム」Proceedings、pp.871-874 (1995).

[15]H. Markowitz ; “Portfolio selection”, Journal of Finance (1952).

[16]M. Ohya “Fractal dimension of states”, in Quantum Probability and Related Topics VI, pp.359-369 (1991).

[17]T. Matsuoka & M. Ohya; “Fractal dimensions of states and their application to Ising model”, Rep. Math. Phys. 36, pp.365-379 (1995).

[18]B. B. Mandelbrot; “The Fractal Geometry of Nature”, W.H.Freemann and Company, San

Francisco (1982).

[19]A. N. Kolomogorov, "Theory of transmission of information", Amer. Math. Soc. Translation, Ser. 2, no.33, pp.291-329 (1963).

[20]A. N. Kolomogorov, V. M. Tihomirov, " ε -Entropy and ε -capacity of sets in function space", Amer. Math. Soc. Translation, Ser. 2, no.17, pp.277-364 (1961).

[21]T. Matsuoka & M. Ohya; "Fractal dimension of a state and its application to shape analysis problem", to appear in Journal of Polymath.

[22]K. Inoue, T. Matsuoka, M. Ohya, "Calculation for ε -entropy of a state in discrete system", in prepalation.

[23]L. Bachelier;"Theory of speculatin", in P. Cootner, ed.,"The random character of stock market price", Cambridge, MA: M.I.T. Press 1964. (Orginally published in 1900.)

[24]P. Cootner; ed.,"The random character of stock market price", Cambridge, MA: M.I.T. Press 1964.

[25]B. B. Mandelbrot; "The Variation of certain speculative prices", J. of Business ,36, pp. 394 (1963).

[26]B. B. Mandelbrot; "Statistical methodology for nonperiodic cycles: From the covariance to R/S analysis", Annales of Economic and Social Measurement, 1, pp. 259 (1972).

[27]H. J. Fama; "The behavior of stock market prices", J. of Business,38, pp.34 (1965).

[28]U. A. Muller, M. M. Dacorogna, R. B. Olsen, O. V. Pictet, M. Schwarz, C. Morgenegg; "Statistical study of foreign exchange rates, empirical evidence of a price change scaling law, and intraday analysis", J. Banking and Finance,14, pp.1189 (1990).