



Title	確率密度推計法を用いた裾指数推計による株式のリスク度測定
Author(s)	高田, 輝子
Citation	国際公共政策研究. 2002, 6(2), p. 209-224
Version Type	VoR
URL	https://hdl.handle.net/11094/10452
rights	
Note	

The University of Osaka Institutional Knowledge Archive : OUKA

<https://ir.library.osaka-u.ac.jp/>

The University of Osaka

確率密度推計法を用いた裾指数推計 による株式のリスク度測定*

A Nonparametric Approach to Tail Index Estimation*

高田 輝子**

Teruko TAKADA**

Abstract

The Measurement of the risk of financial assets is an important problem for financial management. Although there are various approaches to measuring risk, obtaining detailed risk information is generally offset by the accuracy of the estimates, due to the constraint of the number of extreme observations available. This article proposes a new efficient risk evaluation method that estimates tail heaviness via nonparametric density estimation. The method is applied to evaluate and compare the tail heaviness of the densities of several Japanese and U.S. stock price change.

キーワード：裾指数、確率密度推計、株価変化率、リスク

Keywords : Tail index, Density estimation, Stock price change, Risk

* Colin McKenzie 教授には多くの有益なコメントを頂いた。ここに感謝申し上げます。

**大阪大学大学院国際公共政策研究科 研究生

I はじめに

金融資産価値の変動によってもたらされる損失の危険性（リスク）には様々なものがある。中でも、金融資産価格の突然の大幅な下落は、市場関係者のみならず、国の経済全体にも大きな影響をもたらしうる。従って、大幅な価格下落の危険性を正確に測定することは、学術研究のみならず、実務家にとっても重要な問題であるといえる。しかし、このような事象は、測定に利用できるデータの数が限られているのに加え、異常値を多く含むため、その可能性を正確かつ詳細に測定することは一般的に困難である。本稿の目的は、大幅な価格下落の危険性について、金融資産間の横比較や異時点比較ができるような普遍的な統計量を効率的に測定する、新しい方法を提案することである。

価格変動リスクの指標として一般的に良く用いられている統計量は大きく三つに分けられるが、それらが示すリスクは同じものではない。まず最も簡便な方法として、分散や尖度など、価格変化データのモーメントを利用するものがある。分散は2次のモーメントで、価格の変化しやすさ（ボラティリティー）を捉えるものである。また、尖度は4次のモーメントで、確率密度関数の形が（正規分布などの場合と比較して）平たい丘状になっているのか、または中心のモード部分が高いスリムな山状になっているか、を捉えるものである。尖度は、確率密度関数の裾部分にデータが集まっている度合（裾長度）を測る簡便な指標として、よく用いられる。ただ、モーメントをリスク度の指標として用いるのには多くの問題がある。モーメントは概して、わずかな異常値にも大きな影響を受けるなど、頑健性（robustness）に欠ける統計量である。またその算出においては、裾周辺だけでなく、データの全域において統計量を集計するので、大きな価格変化等、裾の振舞にだけ興味がある場合には、最適な指標であるとはいえない。価格上昇リスクと価格下落リスクが一緒に扱われてしまうという不都合もある。これらの統計量は、リスクについての大まかな傾向をつかむのには使えても、突然の大きな価格下落リスクなどを正確に判定するには不適といえる。

第二の測定法としては、VAR (Value at Risk) がある。これは最大限でどれくらいの損失を覚悟していればよいのかを表す指標で、所与の信頼区間において所与の期間内に当該保有資産が損失を被る確率として定義される。すなわち、資産価格変化率の確率密度関数の端に位置する X_p よりも大きい価格下落が発生する確率を求めるものである。実際の推計の手順としてはいくつかの異ったアプローチがある。例えば、近年注目されるものとして、regression quantiles を使った Engle and Manganelli (1999) の CAViaR (Conditional Autoregressive Value at Risk) がある。VAR の測定法のサーベイに関しては、Linsmeier and Pearson (1996) や Duffie and Pan (1997) を参照されたい。VAR は近年急速に普及してきた方法で

あり、多くの金融機関が自己保有資産のリスク度測定のために採用している。この方法はモーメントと異なり上昇リスクと下落リスクが区別でき、解釈もわかりやすい。しかし、 X_p の取り方に応じたリスクの水準を表しているため、リスクについての普遍的な統計量ではないという点が問題である。例えば、スケールの影響を受けるため、分布が裾長な場合の振舞とボラティリティーが大きい場合の振舞とを区別することは難しい。

第三の測定法は、金融データの確率密度関数の裾長度 (tail heaviness) をリスク度の指標とするものである。裾長度とは、極値理論を基に比較的起こりにくいけれど確率密度関数の両端の事象が、どれくらいの頻度で発生するかを測るものである。具体的には、確率密度関数の両端の漸近的な振舞を表すパラメータである裾指指数 (tail index) を求める。極値理論については、例えば Leadbetter, Lindgren and Rootzén (1983) を参照されたい。代表的な方法としては Hill 推計量 (Hill (1975)) がある。この方法は、VAR のようにリスク水準を求めるのではなく、リスク度の関数形を求めるものなので、より普遍的な統計量であるといえる。

以下本稿では、リスク度を裾長指数によって測定する。裾長指数の測定においては、裾周辺のデータ数が限られているため、正確な推計値を効率的に求める方法を選択することが非常に重要である。既存の裾指指数推計法には様々なアプローチがあるが、まだ決定的な推計方法の確立には至っていない。本稿では、まずデータの確率密度をノンパラメトリックに推計し、それを基に裾指指数を推計する方法を提案する。ノンパラメトリックな確率密度推計を用いることで、限りある裾周辺データの有効利用が可能になり、比較的少ないデータ数で正確なリスク度を測定できるようになる。特に裾部分の推計精度の高い確率密度推計法を選べばより詳細なリスク分析が可能になり、投資行動や政策策定に有用であると考えられる。この推計法の応用例として、日米の日次株価データを用い、リスク度の測定や比較、時間変化に伴うリスク度の変化、に関する分析を行った。

本論文の構成は以下の通りである。まず 2 節で裾長度や裾指指数が何を表すかを説明した後、既存の裾指指数推計法を振り返り、確率密度推計法を用いた新しい裾長指指数の推計方法を提案する。3 節では、2 節で提案した裾指指数推計法の精度を既知の確率分布を使って評価する。4 節では実証研究で用いるデータの属性とその処理方法を記述し、5 節で裾指指数の推計結果とその解釈を論じる。6 節は本論文をまとめ、今後の課題を展望する。

2 捷指数とその推計方法

2.1 捷指数

大きな価格変化の可能性など、発生頻度の少ない事象に関する情報は、確率密度関数の両端である、捷（tail）周辺に含まれる。金融資産価格データの例でいえば、大きな価格変化が比較的頻繁に起きる場合、その確率密度関数の両端は正規分布より厚くなり、なかなか頻度ゼロに収束していかない。このような分布は捷長分布（heavy-tailed distribution）と呼ばれる。実際、金融資産価格変化が捷長分布を持つことは Fama (1965) をはじめとする報告により、広く知られた事実である。そして確率分布の捷が長い程、リスク度が高いことになる。

確率分布の捷は、指数的捷と代数的捷の二つに大きく分けることができる。ある固定値 x の確率分布を $F(x)$ とするとき、指数的捷を持つ確率分布は、ある $\beta > 0$ 、 $\gamma > 0$ に対して

$$\lim_{x \rightarrow \infty} -\frac{\log(1-F(x))}{\beta x^\gamma} = 1 \quad (1)$$

という関係を満たすものであり、 γ が捷指数となる。代数的捷を持つ確率分布は、ある $\alpha > 0$ に対して

$$\lim_{x \rightarrow \infty} -\frac{\log(1-F(x))}{\alpha \log(x)} = 1 \quad (2)$$

満たすものであり、 $\alpha > 0$ が捷指数となる

より直観的に理解するために、(1)–(2)式を頻度の少ない事象 x が発生する確率、 $1-F(x)$ として次のように変形することができる。指数的捷を表す(1)式は、ある $\beta > 0$ に対して

$$1-F(x) \sim e^{-\beta x^\gamma}, \quad x \rightarrow \infty \quad (3)$$

と変形され、また、代数的捷を表す(2)式は

$$1-F(x) \sim x^{-\alpha}, \quad x \rightarrow \infty. \quad (4)$$

と表現することができる。

正規分布は指数的捷を持っており、 $\gamma=2$ に対応する。ARCHなどの分布は正規分布よりは捷長であるが、正規分布の組み合わせであるため指数的捷を持つ。 γ が 2 より小さくなる程分布は捷長となり、 $\gamma=0$ のところで代数的捷に変わる。代数的捷は、全ての指数的捷よりも捷長であり、 α が小さい程捷長となる。本稿では、捷長分布とは代数的捷を持つ分布のことを指し、その捷長度を測るのに捷指数 α を用いる。 t 分布の場合、自由度が α と一致する。金融データは捷長分布を持つことがあらかじめわかっているので (Fama (1965))、以下本稿では、

代数的裾指数、 α に焦点を絞って論じていく。

2.2 補指数推計法

2.2.1 既存の補指数推計法

補指数 α の推計法には色々なものがあるが、大きく二つのアプローチがある。ひとつのアプローチは t 分布、Stable 分布、ARCH など、特定の確率密度関数を仮定し、それらのパラメータを推計し、その確率密度推計値を使って補指数を求めようとするものである（例えば、McCulloch (1997)）。これは、推計に用いるデータの確率分布があらかじめわかっている場合は、効率的な方法である。しかし金融データへの応用にあたっては、その確率分布がどのような形を持つかについてはいまだコンセンサスが得られていない。このアプローチでは、データの確率密度関数が特に補周辺において仮定した形とは大きく異なる場合、得られた推計値の有意性が疑わしくなってしまう点が、問題である。

補指数推計のもう一つのアプローチは、特定の確率密度関数を仮定せずに、補データの漸近的振舞について(1)式や(2)式のような定義に基づいて、補周辺のデータから直接、補長度を表す補指数を推計するものである（例えば、Hols and de Vries (1991)）。よく用いられるのは最大順序統計 (largest-order statistics) を使った推計法であり、中でも代表的なのは Hill 推計量 (Hill (1975)) である¹⁾。ただ、補指数推計というものが、頻度の少ない補周辺のデータを要することから、少ない標本サイズの下で正確な推計量を得ることは一般的に難しい問題である。このアプローチでも精度向上のため、Hill 推計量の改良や、最大順序統計を用いた新たな推計量の考案、sub-sample bootstrap 法の併用など、様々な方法が提案されてきている。しかし、Kearns and Pagan (1997) や Resnick (1997) に論じられているように、補推計に使用するデータとしてどの範囲を選ぶかが定まっていない点や、推計法の選択にあたってはバイアスと効率性のトレードオフがある点などは、解決すべき問題として残されている。

2.2.2 本稿の補指数推計法

本稿では、補周辺のデータの少なさを補う効率的な補指数推計法として、ノンパラメトリック確率密度推計法を用いたアプローチを提案する。これは、まず補周辺の推計精度の高い方法で確率密度関数 $f(x)$ の推計値 $\hat{f}(x)$ を求め、(2)式の関係を使って $\hat{f}(x)$ の補部分の推

1) Hill 推計量は、標本 X_1, \dots, X_n の順序統計を $X_{(1)} \geq X_{(2)} \geq \dots \geq X_{(n)}$ とするとき、任意の $m < n$ に対して $\frac{1}{\hat{\alpha}} = \frac{1}{m-1} \sum_{i=1}^{m-1} \log X_{(i)} - \log X_{(m)}$ として定義される。 m の値によって推計結果に大きな違いが出てくるが、その選択法についての統一的な見解はない。

計値から裾指数を推計しようとするものである。

確率分布 $F(x)$ に対応する確率密度関数を $f(x)$ とするとき、(4)式は

$$f(x) \sim x^{-(\alpha+1)} \quad (5)$$

と変形される。(5)式の両辺の対数をとると、

$$\log f(x) = c - (\alpha + 1) \log(x) \quad (6)$$

と線形の関係式として表現できるため、 α の推計値は両対数プロットにおける単純な線形最小二乗法（OLS 法）から求められる。

x の裾領域のある点を x_i とし、そのノンパラメトリック確率密度推計値を $\hat{f}(x_i)$ とするととき、本稿で用いる裾指数推計式は

$$\log \hat{f}(x_i) = \tilde{c} - (\tilde{\alpha} + 1) \log(x_i) + \tilde{\varepsilon}_i \quad (7)$$

と表される。 $\tilde{\varepsilon}_i$ は確率密度推計値を基に裾指数を計算する回帰計算の誤差項であり、標本の大きさが n であるときの最小二乗推定量の標準誤差、 $\tilde{\sigma}_{ols}$ は

$$\tilde{\sigma}_{ols} = \sqrt{\frac{\sum \tilde{\varepsilon}_i^2}{n-2}} \quad (8)$$

で計算される。ここで、(6)式の c, α と (7)式の $\tilde{c}, \tilde{\alpha}$ との違いは、 $f(x)$ と $\hat{f}(x)$ の違いと、回帰計算の誤差からくるものである。

実際の応用において、 $\hat{f}(x)$ は与えられた標本を基に推計されるため、 $\hat{f}(x)$ には、 $f(x)$ から発生させた標本のばらつきや、確率密度推計の誤差が含まれてしまう。特に、ここでは裾の振舞に注目しているので、標本ごとのばらつきの影響を非常に強く受けることになる。ひとつの大標本に含まれるデータの数が十分でない場合、その影響はより大きくなる。なお、頻度の少ない事象 X_p の発生確率を算出するには裾指数の推計結果、 $\tilde{c}, \tilde{\alpha}$ により、 $1 - \hat{F}(X_p) = (\exp\{\tilde{c}\}/\tilde{\alpha}) X_p^{-\tilde{\alpha}}$ のように求められる。

2.3 Logspine 確率密度推計法

金融資産の価格変化率の確率密度関数は裾長であることがあらかじめわかっているため、本稿では Kooperberg and Stone (1991) の logspine 確率密度推計法を用いる。Logspine 法は他の代表的な確率密度推計法に比べて裾長分布の推計精度が高いことが、Takada (2001a 1 章、2001b) の比較分析により示されている。

Logspine 法は確率密度関数の対数値を 3 次のスプラインとして次式のようにモデル化する。

$$f(x, \theta) = \exp \left\{ \sum_{i=1}^p \theta_i B_i(x) - c(\theta) \right\}, \quad x \in \mathcal{R}. \quad (9)$$

ここで、 $c(\theta)$ は正規化のための項であり、関数 $B_i(x)$ は標準的な B-スプライン基底 (B-

spline basis) である。その係数の θ は最尤法により推計され、節点 (knot) と呼ばれるスプラインの接続点の配置は、ベイズ情報基準 (BIC) によって決定される。

3 裾指数推計精度評価

本稿で用いる(7)式による裾推計においても、 x の値域の取り方によって、推計値のバイアスと推計結果のぶれの大きさ（分散）のトレードオフの問題が出てくる。すなわち、(7)式の推計の値域を選択するに当たって、偏りを小さくするためには、なるべく端の方を値域とすべきであるが、推計のぶれを小さくするためには、より多くのデータを使えるよう、内側の値域を使うのが良い。実際の応用の前に、「推計値のバイアス」と「推計のぶれの大きさ」について予想される程度や傾向をあらかじめ知っておくことは、間違った解析を避ける上で極めて重要である。

本節では既知の確率分布を使って、本稿で用いる裾指数推計法の精度をモンテカルロ法により査定する。まず、標本を与えられたときに自動的に裾推計の値域を決めるアルゴリズムを説明する。次に裾推計の値域の変化による推計精度への影響についてモンテカルロ分析を行い、後の実証分析における推計値域をどのように決定するかを論じる。

本稿では異なる企業の株価データも扱うため、確率密度関数 $f(x)$ の裾推計における x の値域の上限と下限は、標本分位 (sample quantile) に従って決定する。例えば左裾指数推計の場合大きさ n の標本 X_1, X_2, \dots, X_n の順序統計を $X_{(1)} \leq X_{(2)} \dots \leq X_{(m)} \dots \leq X_{(n)}$ とするとき、 m 番目に小さい統計、 $X_{(m)}$ は、 m/n 標本分位 (sample quantile)、または $100m/n\%$ 標本分位になる。なお 50% 標本分位は標本メディアンである。左裾指数推計の際、 x の値域の下限を $j\%$ 標本分位と指定した場合、 $j \approx 100m/n$ とする m を探し、 $X_{(m)}$ を値域の実際の下限とする。本稿の裾推計で用いる x の値域は、0.1–0.2%、0.1–0.25%、0.1–0.3%、0.1–0.35%、0.2–0.4%、0.2–0.6%、0.2–0.8%、0.2–1.0% 標本分位である。右裾指数推計の場合のアルゴリズムも同様である。

精度評価のためのモンテカルロ分析のデザインは以下の通りである。既知の確率分布としては、本稿で用いる株式データの裾長度に近い、 $t(4)$ 分布を用いた。 $t(4)$ 分布の代数的裾指數 α は 4 である。標本の大きさは、本稿の実証研究で用いる期間である、14年分に対応する TOPIX のデータの大きさに合わせて、3652とした。本稿の実証分析で扱う標本の大きさは全て 3539–3660 の範囲内である²⁾。まず、大きさ 3652 の $t(4)$ 分布の標本を計算機実験により 500 回発生させる。次に、その 500 の標本についてそれぞれ(7)式に基づく裾指数推計を行う。表 1 は、その結果から得られた 500 の $\hat{\alpha}$ や $\hat{\sigma}_{ols}$ の平均や標準偏差を、いくつかの推計値域について

2) 実証分析に用いるデータの詳細については 4 節を参照されたい。

表1:t(4)分布の裾指指数推計精度のモンテカルロ分析結果

推計値域(%分位)	0.1%~0.2%	0.1%~0.25%	0.1%~0.3%	0.1%~0.35%
$\tilde{\alpha}$ 平均	3.944	3.877	3.830	3.791
($\tilde{\alpha}$ 標準偏差)	(0.957)	(0.840)	(0.773)	(0.736)
$\hat{\sigma}_{ols}$ 平均	0.009	0.010	0.011	0.011
推計値域(%分位)	0.2%~0.4%	0.2%~0.6%	0.2%~0.8%	0.2%~1%
$\tilde{\alpha}$ 平均	3.519	3.479	3.418	3.365
($\tilde{\alpha}$ 標準偏差)	(0.630)	(0.588)	(0.513)	(0.474)
$\hat{\sigma}_{ols}$ 平均	0.006	0.008	0.009	0.010

注:統計値はシミュレートされた500の標本についての平均値。 $\tilde{\alpha}$ 、 $\hat{\sigma}_{ols}$ の定義は(7)、(8)式に基づく。

て比較した結果をまとめたものである。 $\tilde{\alpha}$ の値が $t(4)$ 分布の真の裾指指数 $\alpha=4$ に近い程、推計誤差は小さいことになる。なお、 $t(4)$ 分布は対称なので、表には左裾指指数推計の精度分析のみを示した。

表1を見て気付くのは、推計量のバイアス $\alpha - \tilde{\alpha}$ に比べて、 $\hat{\sigma}_{ols}$ が格段に小さいことである。例えば、 α のバイアス $\alpha - \tilde{\alpha}$ が 0.056~0.635 であるのに対し、最小二乗推定量の標準誤差は 0.009~0.011 と一桁以上小さい。これはすなわち、推計誤差 $\alpha - \tilde{\alpha}$ のほとんどの部分が標本のばらつきからきており、確率密度関数推計値の裾部分を(7)式で回帰する際の誤差はほぼ無視できるレベルだということを示す。なお、推計値域が内側に拡大していくのに伴い、わずかずつ $\hat{\sigma}_{ols}$ が大きくなっていくのが観察されるが、これは両対数プロットにおける漸近的な線形性が、値域を内側に取る程弱まることによる。この点に関しての詳しい説明については補遺を参照されたい。

表1によると、推計値域を最も左端に最も短くとったものについては、裾指指数の推計値平均が 3.944 と最も真の値に近くなっているものの、標準偏差が 0.957 と最大である。そして、推計値域を左端から内側へとシフトさせる程、そして推計値域の幅を大きくする程、裾指指数の推計値は下方にバイアスがかかるようになるが、推計結果のぶれである標準偏差は小さくなっていく。このように、推計値域の選択に当たっては、推定量のバイアスと推定の分散の大きさのトレードオフがあり、適当なバランスのものを選ぶしかない。本稿では、表1の中では最も標準偏差が小さい、0.2~1%標本分位を値域として選択する。これは、推計結果の標準偏差があまり大きいと、いくつもの異なる企業の株価データを比較がしづらくなる点を考慮しての選択であるが、実証分析に当たっては、推計結果の下方へのバイアスに留意しなければならない。

4 データ

実証分析に用いるデータは、東京証券取引所上場の日本の自動車企業の株価として、トヨタ自動車、本田技研工業、日産自動車、ニューヨーク証券取引所 (NYSE) 上場の米国自動車企業としてフォードの株価、そして、TOPIX、NYSE の株価指数である。出所は、日本の個別企業株価が野村総合研究所、フォードの株価は CRSP、NYSE 株価指数が NYSE のウェブサイト、TOPIX 株価指数がデータストリームである。データは全て日次であり、期間は個別企業の株価が1983年から1996年までで、株価指数が1969年から1996年までである。図1に、本稿で扱った全ての株価の原データの時系列が示されている。

実証分析にあたり、これらの株価データについては、以下に述べる手順で日次変化率に変換し、季節性を除去した。以下、データ調整法を述べた後、実証分析に用いるデータの基本的な統計量について分析する。

時点 t の株価を p_t とするとき、日次株価変化率 Δp_t を

$$\Delta p_t = \log p_t - \log p_{t-1} \quad (10)$$

として求める。この日次株価変化率から、1) 1月-12月効果、2) 曜日効果、3) 取引日の間が空くことによる効果、などを除去する季節調整を行った³⁾。

季節調整の方法としては、以下に述べる Jorgenson (1964) の方法をとった。 y を季節調整前の長さ n のベクトル、 x を季節調整のダミー変数ベクトルとし、まず以下の回帰計算を行う。

$$y = x'\beta + e. \quad (11)$$

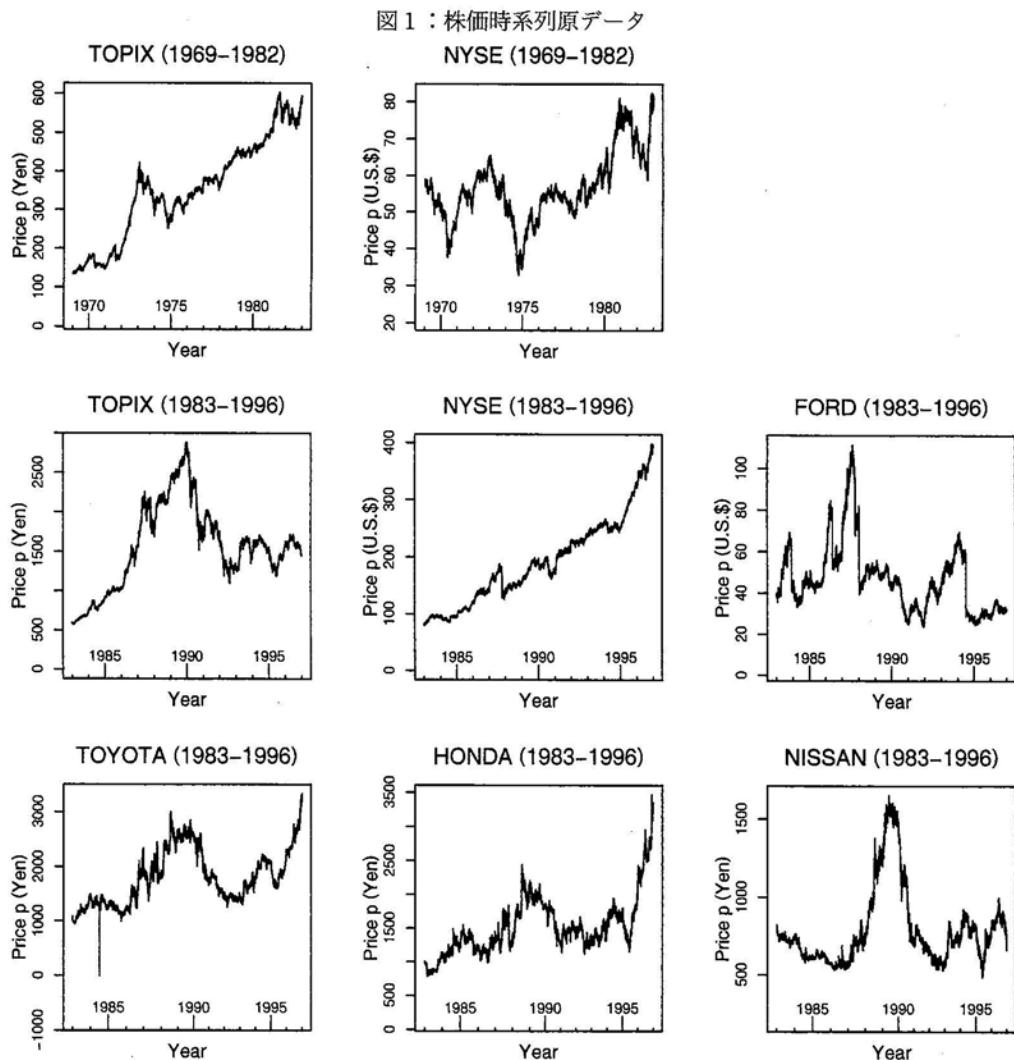
ここで、 β は最小二乗法によって得られた係数ベクトルであり、 e はその誤差ベクトルである。(11)式には通常、トレンドや循環的な要素を取り除くための調整項を入れるが、 Δp_t はほとんどトレンドのない系列なので、本稿では省いた。季節調整後の系列、 y_{adj} は、

$$y_{adj} = y - x'\beta \quad (12)$$

として求められる。

表2は後節の実証分析で用いる、季節調整後の株価日次変化率について、モーメントなどの基礎統計をまとめたものである。正規分布の場合、歪度係数、過尖度とともに0になる。リスク度の解析においては、分散はボラティリティーを示す指標として、尖度は裾の振舞を示す指標として、良く利用されるモーメントである。ただ尖度は必ずしも裾長度を表すのではない点に留意されたい。過尖度については、どのデータも有意に0よりも大きく、正規

3) これらの効果を除去するために用いたダミー変数は、Gallant, Rossi, and Tauchen (1992) とほぼ同じだが、本稿で使用する時系列データには不要と考えられる、土曜効果、戦争効果などの項は除いた。



分布の場合よりも尖度が高いことが示唆されている。また、分散の大小と過尖度が必ずしも対応していないことが観察される。例えば、トヨタ自動車、本田技研工業、日産自動車についてみると、三つの中で分散の最も小さいトヨタ自動車の過尖度は最も大きくなっている。表2の全系列について Jarque-Bera の正規性検定を行ったが、 χ^2 値は全て $2.2e-16$ 以下となり、正規性が棄却された。なお、表2に含まれる全系列に共通した季節性としては、年末年始効果が観察された。

表2：日次価格変化率データの基礎統計

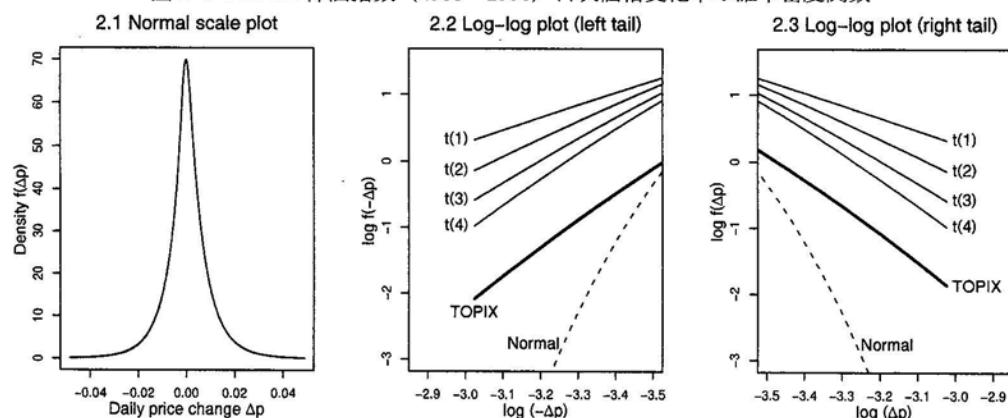
	データ数	平均	分散	歪度係数	過尖度	最小値	最大値
1969-1982							
TOPIX 株価指数	3652	0.000406	0.000054	-1.061	11.869	-0.077	0.045
NYSE 株価指数	3534	0.000090	0.000074	0.234	2.024	-0.037	0.051
1983-1996							
トヨタ自動車	3659	0.000305	0.000302	0.227	11.641	-0.210	0.162
本田技研工業	3660	0.000327	0.000421	0.297	6.733	-0.209	0.151
日産自動車	3660	-0.000055	0.000334	0.528	5.757	-0.134	0.140
TOPIX 株価指数	3652	0.000248	0.000115	-0.508	19.892	-0.157	0.092
NYSE 株価指数	3539	0.000450	0.000076	-4.385	106.754	-0.212	0.087
フォード	3540	-0.000046	0.000700	-12.182	294.499	-0.692	0.094

5 実証分析結果

本節では、まず TOPIX 株価指数を例にして裾指数推計の手順を述べ、そこから引き出される事実について論じる。次に、扱った全株価データについての裾指数推計の結果を解析する。なお、以下の推計は、確率密度関数やリスク度が、用いたデータの期間内については不变であるという仮定に基づいている。

図2.1は logspline 法による1983年から1996年までの TOPIX の日次株価変化率の確率密度関数推計結果である。モードが尖っているのが特徴的である。図2.2及び2.3は、図2.1の左裾、右裾部分を、両軸とも対数をとった後に拡大したものである。太い実線が TOPIX 株価指数である。裾長度を比較参照するために、正規分布を細い破線で、 $t(1)$ 、 $t(2)$ 、 $t(3)$ 、 $t(4)$

図2：TOPIX 株価指数（1983-1996）日次価格変化率の確率密度関数



分布を細い実線で表した。図2.2、2.3では縦軸、横軸とも対数をとっているので、代数的裾を持つ t 分布は直線に漸近し、指数的裾を持つ正規分布は下向きの曲線になる。代数的裾を持つ場合、直線の傾きが小さい程、裾長度が高くなる。

TOPIX の確率分布の裾は漸近的に、ほぼ直線を描いているのが観察される。従って、裾指数の推計にあたり、代数的裾指数を仮定した(7)式を用いるのは適当と考えられる。図2.2、2.3によると、TOPIX の確率密度関数の両対数プロットにおける傾きは、両裾とも $t(3)$ と $t(4)$ の間位であるとみられる。すなわち、TOPIX の確率分布は正規分布よりもはるかに裾長であることが明らかに観察され、正規分布を仮定した分析が非常に間違った結論を導きかねないことを示唆している。

一般に良く知られているように、ランダムな変数の和は中心極限定理によって正規分布に従う。全ての投資家が各自いつも独立な確率分布に基づいて投資行動を起こしているのだとすれば、その総和としての価格変動は正規分布に近いものになると推測できる。従って、株価変化の確率分布が裾長であることは、投資家の行動を司る確率変数が独立でないことを暗示しているのかもしれない。例えば、チャーティスト(chartist)のような相場の動きに連動した投資行動が、裾長分布の一因になっていると考えられる。実際、裾長という性質は、長期系列相関(long range dependence) や volatility clustering と強い関係があることが近年明らかになってきている(Resnick (1997))。

上記と同様の解析を表2に含まれる全ての株価日次変化率データについて行い、(7)式により代数的裾指数 $\tilde{\alpha}$ を推計した。ここでの裾長度は13年間一定と仮定している。裾推計に用いた日次株価変化率 Δp の値域は、0.2%–1.0%標本分位の範囲とした。表1の裾推計精度評価の結果を考慮し、本節の裾推計結果は0.635下方にバイアスがあり、0.474の標準偏差を持つことに留意すると、真の裾指数の95%信頼区間は $(\tilde{\alpha} + 0.635) \pm 2 * 0.474$ の範囲となる。

この裾指数推計結果をまとめたのが、表3である。まず図2に対応する1983–1996年のTOPIXの裾指数推計値を見ると、左裾は $\tilde{\alpha} = 3.01$ 、右裾は $\tilde{\alpha} = 2.64$ となっている。0.635の下方バイアス等を考慮すると、TOPIXの真の裾指数は、左が $3.65 \pm 2 * 0.474$ 、右が $3.27 \pm 2 * 0.474$ となり、図2における観察と整合的であることが確認できる。

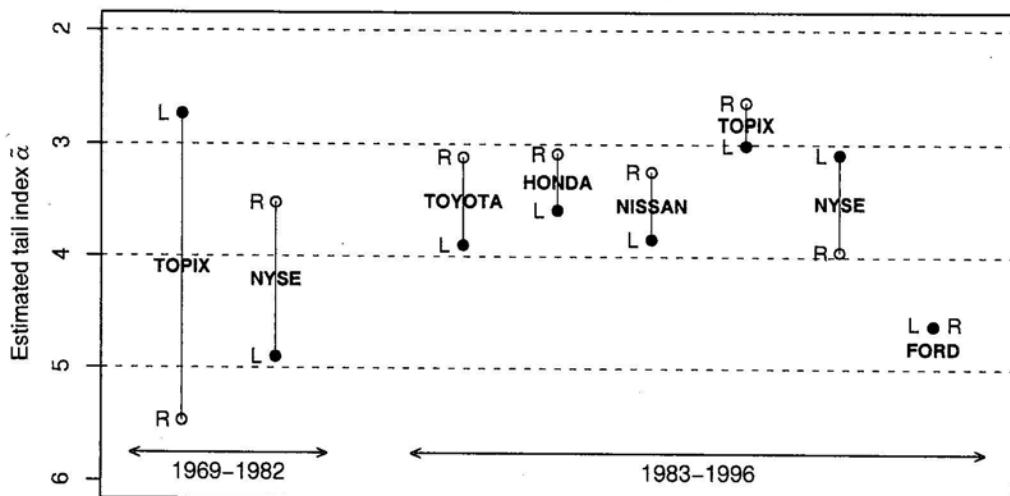
直観的に理解しやすいように、表3の $\tilde{\alpha}$ を図3にグラフ化した。横に L と表記された黒点は左裾の $\tilde{\alpha}$ を、横に R と表記された白点は右裾の $\tilde{\alpha}$ を示している。図の左側のNYSEとTOPIXは、1969–1982年の期間の $\tilde{\alpha}$ であり、そこから右側の6系列は全て1983–1996年の期間の $\tilde{\alpha}$ である。この図から発見できる事実として、次の二点を挙げることができる。

一つめの事実は、扱った全ての株価データについて、株価変化の分布が裾長である点である。中でも1983–1996年の期間のTOPIX及び日本の3企業の株価変化については、 $\tilde{\alpha}$ が概ね3から5の範囲にあり、正規分布に比べはるかに裾長である。裾長度の時間的な変化をみ

表3：裾指数推計結果

	左裾			右裾		
	$\tilde{\alpha}$	$\alpha \approx \tilde{\alpha} + b$	$\alpha \pm 2\sigma$	$\tilde{\alpha}$	$\alpha \approx \tilde{\alpha} + b$	$\alpha \pm 2\sigma$
1969-1982						
TOPIX 株価指数	2.74	3.37	2.4-4.3	5.47	6.10	5.2-7.0
NYSE 株価指数	4.90	5.54	4.6-6.5	3.52	4.15	3.2-5.1
1983-1996						
トヨタ自動車	3.89	4.53	3.6-5.5	3.12	3.75	2.8-4.7
本田技研工業	3.58	4.22	3.3-5.2	3.08	3.72	2.8-4.7
日産自動車	3.84	4.48	3.5-5.4	3.24	3.88	2.9-4.8
TOPIX 株価指数	3.01	3.65	2.7-4.6	2.64	3.27	2.3-4.2
NYSE 株価指数	3.09	3.73	2.8-4.7	3.95	4.59	3.6-5.5
フォード	4.62	5.25	4.3-6.2	4.62	5.25	4.3-6.2

注: $b=0.635$, $\sigma=0.474$ はそれぞれ、0.2%~1.0%標本分位を値域としたときに $t(4)$ 分布の α を推計するときの平均的なバイアス及び標準偏差である。

図3：裾指数 $\tilde{\alpha}$ の推定信頼区間

るために、二つの異なった期間における TOPIX と NYSE の $\tilde{\alpha}$ を比較してみたが、全体的にみて、近年の方がより裾長になってきているとは、いえないようである。

二つめの事実として挙げられるのは、1983-1996年の期間の日本の自動車企業3社の日次株価変化は右裾の方が長い点である。この3社の右裾と左裾の裾長指数の差は、0.50~0.77と、辛くも1標準偏差 ($1\sigma=0.47$) を上回っている程度なので解釈には注意を要するが、三企業の株価変化が全て同じ傾向を示しているため、右裾の方は左裾よりも長いという主張は、有意であると考えられる。ただこれは、1983-1996年の期間の NYSE やフォードのデータに

については当てはまらないので、右裾の方が長いのが当時の世界的な傾向だったとはいえないようである。また、1969-1982年のTOPIXについては、むしろ左裾の方が右裾よりも長くなってしまっており、日本の株価データも扱う時期によって、右裾と左裾の相対的な裾長度が変化することがわかる。

1983-1996年の期間に起こった世界情勢の大きな変化としては、1987年のブラック・マンデー、1992年の湾岸戦争などが挙げられる。日本市場について特筆すべきことは、1986年から1989年までの4年間で株価が3倍近くも上昇するバブル時代を含んでいることである。右裾の方が左裾よりも長いということは、1日に大きな価格上昇が起こる可能性の方が、大きな価格下落の可能性よりも高い、ということである。すなわち、この時期の日本の投資家心理はどちらかといえば強気で、相場が下がったから売るというよりは、相場が上がったから買う、という行動をとることが多かったと推測される。

6 おわりに

金融資産のリスク度を詳細かつ正確に測定することは、重要かつ難しい問題である。本稿は、裾長分布に強いノンパラメトリック確率密度推計を用いた裾指数推計を提案し、それを用いて株式のリスク度についてのより詳細かつ正確な情報を分析した。同推計法を用いた実証分析では、次の3点が示された。1) 本稿で扱った全ての株価データの日次変化率の分布は裾長であることが示された。特に1983-1996年における日本の株価変化の分布の裾長度に関しては、概ね $t(3)$ 分布から $t(5)$ 分布の間に位置し、正規分布に比べはるかに裾長である。2) 1983-1996年における日本の株式市場は、大きな価格下落のリスクよりも大きな価格上昇のリスクの方がやや高い傾向にある。

本研究の今後の方向として興味深いのは、リスクの時間変化の測定である。リスクを正確に測定するにはある程度の標本の大きさが必要であるため、長い時系列データを確保することが重要になってくる。例えば、逐次データを利用することは、時間変化をより詳細に論じることが可能になるための一つの方向である。

補 遺

両対数プロットにおける t 分布の確率密度関数の漸近的な傾きについて

$t(\nu)$ 分布の確率密度関数 $f(x)$ の右裾を、縦軸、横軸とも対数を取った図を描いたとき、裾は漸近的に線形となり、その傾きは $-(\nu+1)$ に近付いていく。なお、 t 分布の場合、裾指数 α は自由度 ν に一致する。しかし、サンプルデータをもとにこの傾きを測ろうとするとき、値

域を横軸の内側に取る程、傾きはその漸近値から増加し、確率密度関数の両対数プロットの直線性が弱まっていく。以下、 $t(4)$ 分布の右裾の場合を取り、具体的にこの事象を解説する。

$t(4)$ 分布の確率密度関数は

$$f(x) = 12 \left(\frac{1}{4+x^2} \right)^{5/2} \quad (13)$$

である。(13)式の両辺の対数をとると

$$\log f(x) = \log 12 - \frac{5}{2} \log(4+x^2) \quad (14)$$

となる。

$t(4)$ 密度関数の両対数プロットの傾きと、裾長指数 $\alpha=4$ の関係をより明確に示すため、(14)式を変形する。 $Y=\log f(x)$ 、 $X=\log x$ とすると、(14)式は $Y=\log 12 - \frac{5}{2} \log(4+e^{2X})$ となり、その傾きは

$$\frac{dY}{dX} = -\frac{5}{2} \frac{2e^{2X}}{4+e^{2X}} = -5 \frac{x^2}{4+x^2} \quad (15)$$

で表される。ここで、

$$\lim_{x \rightarrow \infty} \frac{dY}{dX} = -5 \quad (16)$$

であり、両対数プロットの傾き、 $-(\alpha+1)$ 、が漸近的に -5 になることがわかる。また、 x が漸近領域から離れる程、傾きは -5 から増加し、その増加する速さ $\left(\frac{d^2Y}{dX^2}\right)$ も速くなる。左裾の場合についても、傾きの符号が逆になることを除き、同じことがいえる。

参考文献

1. Duffie, D. and J. Pan (1997). "An Overview of Value at Risk," *Journal of Derivatives*, 4 7-49.
2. Engle, R. F. and S. Manganelli (1999). "CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles," working paper, University of California, San Diego.
3. Fama, E. F. (1965). "The Behavior of Stock Market Prices," *Journal of Business*, 38, 34-105.
4. Gallant, A. R., E. Rossi and G. Tauchen (1992). "Stock Prices and Volume," *Review of Financial Studies*, 5, (2), 199-242.
5. Hill, B. M. (1975). "A Simple General Approach to Inference about the Tail of a Distribution," *Annals of Statistics*, 3, 1163-1174.

6. Hols M. C. A. B. and C. G. de Vries (1991). "The Limiting Distribution of Extremal Exchange Rate Returns," *Journal of Applied Econometrics*, **6**, 287-302.
7. Jorgenson, D. W. (1964), "Minimum Variance, Linear, Unbiased Seasonal Adjustment in Economic Time Series," *Journal of the American Statistical Association*, **59**, 681-725.
8. Kearns, P. and A. Pagan. (1997). "Estimating the Density Tail Index for Financial Time Series," *Review of Economics and Statistics*, **79**, 171-175.
9. Kooperberg, C. and C. J. Stone (1991). "A Study of Logspline Density Estimation," *Computational Statistics & Data Analysis*, **12**, 327-347.
10. Kooperberg, C. and C. J. Stone (1992). "Logspline Density Estimation for Censored Data," *Journal of Computational and Graphical Statistics*, **1**, 301-328.
11. Leadbetter, M. R., G. Lindgren and H. Rootzén (1983). *Extremes and Related Properties of Random Sequences and Processes*, Springer-Verlag, New York.
12. Linsmeier, T. J., and N. D. Pearson (1996). "Risk Measurement: An Introduction to Value at Risk," working paper, University of Illinois at Urbana-Champaign.
13. McCulloch, J. H. (1997) "Measuring Tail Thickness to Estimate the Stable Index α : A Critique," *Journal of Business and Economic Statistics*, **15**, 74-81.
14. Resnick, S. I. (1997). "Heavy Tail Modeling and Teletraffic Data," *Annals of Statistics*, **25**, 1805-1869.
15. Takada, T. (2001a) *Density Estimation for Robust Financial Econometrics*, PhD thesis, University of Illinois at Urbana-Champaign.
16. Takada, T. (2001b). "Nonparametric Density Estimation: A Comparative Study," *Economics Bulletin*, **3**, (16), 1-10.