論文審査の報告

論文題目：Goodness-of-Fit-Tests for Heavy Tailed Distributions
氏名：松井宗也

論文の内容：

経済学では以前から金融時系列データが裾の厚い分布によってよく説明されることが指摘されている。また保険の分野でも分布の裾の厚さが重要な話題である。その他、裾が厚いとされるデータが観察される分野は、統計物理、信号処理、通信など多岐にわたり重要である。本論文はこの裾の厚い分布の適合度検定に関して詳細に検討している。分布の適合度検定とはデータを特定の確率分布によってモデル化する際、その当てはまり具合を調べるための統計的に重要な分野の一つである。モデルにおいて仮定される分布とデータから得られる経験分布との距離を何らかの方法で測り、その離れ具合を見て検定を行う。本論文で詳しく扱っている分布のクラスは安定分布という裾の厚い分布のクラスであり、正規分布以外は２次のモーメントを持たないことで知られている。

第１章では研究に必要な基本的な事項が述べられている。第1.1節では今後の応用が期待されている様々な裾の厚い分布族の性質を整理している。一般化ハイパボリック分布、一般化t分布、安定分布の3つである。いずれも正規分布より裾の厚い分布を実現でき、さらに左右の非対称性を実現できることから様々なデータに柔軟に対応できる。特に後の章で取り扱う安定分布に関しては第1.2節で Feller (1971) に沿って理論的な性質についてコンパクトにまとめている。第1.3節では適合度検定の基本的なものとしてクラメール・フォンミーゼス統計量とアンダーソン・ダーリング統計量を Anderson and Darling (1952) に基づいてやや詳しく解説している。両統計量は実際に最もよく使われているノンパラメトリックな適合度検定である。これらは仮定される分布とデータから得られる経験的な分布の2乗距離に重み関数をかけて積分し導出される。経験過程を2乗して重みをつけて積分したものと理解でき、結果として両統計量は漸近的にガウス過程の2乗を積分したものに収束する。クラメール・フォンミーゼス統計量の重みは定数、アンダーソン・ダーリング統計量の重みはガウス過程の共分散関数の逆数を用いる。第4、5章で導入される経験特性関数を用いた適合度検定も2乗距離を測るタイプの検定であり共通する部分が多い。

第2章では安定分布の密度関数の数値計算が詳しく論じられている。多くの安定分布は、その特性関数は解析的に表現できるのに対し、密度関数は初等的な関数で表すことができない。そのため統計的な取り扱いに工夫が必要である。密度関数を計算する方法としては特性関数を数値的に直接反転する方法、Zolotarev による積分表現を用いる方法、密度関数の無限級数展開を用いる方法がある。最

初の方法は一般的ではあるが、振動する関数の無限区間の積分が必要となり数値的な困難がある。そこで本論文では密度関数の Zolotarev による積分表現と密度関数の無限級数展開を組み合わせ、対称安定分布の場合に密度の正確な数値計算を行っている。具体的には、積分表現のみを用いるとパラメータや標本の範囲によっては密度関数の値が不正確になる場合があり、その不正確な範囲を確定し無限級数表現で代替している。さらに密度のパラメータに関する微分を数値的に求めパラメータによる分布の変化を詳しく分析している。その結果を用いてフィッシャー情報量を計算しまた最尤法のモンテカルロ実験を行っている。両者が整合的であったことから数値計算の正確性が保証される。モンテカルロ実験による観測情報量も、尤度関数の 2 回微分を使うものと使わないものとに分けて分析されている。前者は後者と比べ分散が小さいかわりにバイアスが大きいことが明らかにされた。

第 3 章では正規分布の近傍での安定分布のフィッシャー情報量行列の挙動が詳しく分析されている。安定分布の中で全てのモーメントを有する正規分布 $(\alpha = 2)$ は特異な存在である。他の安定分布と異なり裾が多項式ではなく指数的に減少する。それにもかかわらず一次元の分布に関する限り安定分布は連続的に正規分布に近づく。ただし推定に関しては、そのパラメータ $(\alpha, \beta, \mu, \sigma)$ を最尤推定する際に正規分布 $(\alpha = 2)$ のところで α のフィッシャー情報量が無限大に発散することが知られている。これは標本が大きい場合、正規分布からの標本ならば α の最尤推定量は $\alpha = 2$ の値を高い確率で返すという興味深い現象である。このような場合には無限大に発散する際のオーダーを求めることが興味深い問題である。先行研究では Nagaev and Shkol'nik (1988) は対称安定分布の正規分布近くでの挙動を明らかにし、α の情報量の発散のオーダーを求めている。密度関数の Zolotarev による積分表現を詳しく解析することで正規分布近くの安定分布の密度に漸近的に迫るアプローチである。安定分布の無限級数展開のアイディアも用いられている。第 3 章ではこれを拡張して一般安定分布の正規分布近くでの挙動を明らかにしている。パラメータ β が 0 でない場合、密度関数は正負の領域で非対称的に正規分布へ近づく。更に密度の微分を詳しく解析し全てのパラメータの情報量の挙動を明らかにしている。正規分布のところでは情報量行列の β の要素が 0 に退化したり α と σ の交差する要素が無限大に発散するなどパラメータ毎に異なる。正規分布近くの安定分布は応用も広いので今後数値的にこれらの結果を検証することも重要である。

第 4 章では安定分布のひとつであるコーシー分布の適合度検定が論じられている。コーシー分布は t 分布にも属し特にその密度関数が簡単に表現できるが、安定分布への応用を考慮に入れて経験特性関数を用いた適合度検定が行われている。この検定統計量は経験特性関数とパラメータに推定量を代入した特性関数との 2 乗誤差を考え、重み関数をかけて積分したもので与えられる。推定量として分位点を用いたものは Gürtler and Henze (2000) に既に与えられており、本論文では

それを拡張してパラメータ推定を経験特性関数と MLE によるもので与え比較している。モンテカルロ実験によると先行研究と競合的な検定統計量を構成することが分かる。ここでは対立仮説や重みによって検出力が変化するので重み関数の選択が重要となる。先行研究と大きく異なる点は、検定統計量の漸近分布をモンテカルロ実験によるのではなく漸近理論に基づく近似により求めた点である。具体的な手順は以下のようである。本論文で提案されている検定統計量は経験特性関数過程に重みをかけて 2 乗して積分したものと考えられ、従って検定統計量は漸近的にガウス過程に重みをかけて 2 乗積分したもので表現できる。このガウス過程の共分散構造の固有値展開を用いれば、検定統計量は漸近的にカイ 2 乗分布に固有値をかけたものの無限和で表現できる。よってその特性関数が無限積の形で求まり、それを反転することにより検定統計量の漸近分布が求まる。本論文ではガウス過程の共分散構造が具体的に計算され固有値を数値的に近似している。特に共分散構造においては、ガウス過程の正の時間と負の時間で相関が無いことから、漸近統計量は自由度 2 のカイ 2 乗分布の無限和に従い、特性関数の反転の際に留数の定理を用いることができる。数値計算で近似された検定統計量の漸近分布と、標本の大きさ 200 程度のモンテカルロ実験によるそれとが一致いてしてることから、両者の整合性が確認されている。

第 5 章では対称安定分布の適合度検定が論じられている。対象とされている検定は対称安定分布全体の分布系に対する検定 (H_1) と指数が既知 (α 固定) の対称安定分布に従うという検定 (H_2) の 2 つである. 第 4 章のコーシー分布の研究を安定分布に拡張したものと考えられるが、密度関数が陽に表せないため数値計算が多く用いられている。また 4 章の結果を一般化し、経験特性関数を用いた適合度検定で推定量が漸近有効性を持つ場合に検定統計量の漸近的なガウス過程を導いている。すなわち、ガウス過程の共分散構造が特性関数と特性関数のパラメータに関する微分、情報量行列で表現できることを示している。固有値近似を経て検定統計量の漸近分布の特性関数が近似され、それを反転すれば分布が求まる点はコーシー分布の場合と同様である。ただし留数の定理が一般的には利用できないため、Slepian(1957) の手法を用い特性関数を複素平面に拡張して経路を工夫することで反転公式を計算している。特性関数は 2 価関数の無限積となっており分岐点を 2 個ずつ組み合わせその間を切断し、分岐点のペアをそれぞれまわる周回積分の手法を用いている。検定統計量の漸近分布の密度関数はこの周回積分の無限和になっている。無限和は交代級数になっており収束は速い。H_1 と H_2 の両仮説とも推定量に経験特性関数によるものと最尤法によるものを用いた 2 つの適合度検定が考えられている。両推定量とも漸近的な分布が計算でき、漸近検定統計量のガウス過程の共分散関数が計算されている。ただ対象とする範囲が分布系全体と広いため数値的な研究は最尤法によるものに限られている。ここでもコーシー分布の場合と同様に数値計算で近似された検定統計量の漸近分布と、標本の大きさ 200 程度のモンテカルロ実験によるそれとが一致していることから両者の整合

性が確かめられている。経験特性過程は裾が重い場合に原点近くでの収束が速いことも併せて確かめられている。分布系に関する (H_1) 適合度検定をどの様に構成するかに関してはいくつかのアイディアが述べらているがまだ検討が必要である。なぜならこの場合は α の値によって検定統計量の分布が変化するため、α の範囲を制限したり上限をとるなどの検定の手続きが必要となる。従って対立仮説の分布を与えたもとでのモンテカルロ実験などはなされていない。パラメータ α 固定の安定分布に従うという検定 (H_2) では対立仮説には t 分布を用いてモンテカルロ実験をしている。統計モデルでしばしばその確率要素が t 分布か安定分布かということが問題になるからである。安定分布の帰無仮説の裾が厚い場合は良く検出できているが、裾が短く正規分布に近い場合は検出力は小さい。これは正規分布近くの安定分布と t 分布の形が非常によく似ているためである。裾の挙動は両者でかなり異なるのが、分布全体の形状は似ているので実際に検定を用いる際には注意が必要であることがわかる。

講評：第 1 章における裾の厚い分布の様々なモデルに関しては、それぞれのモデルの意味づけが必ずしも十分ではないという印象があった。第 2 章および第 3 章の安定分布の密度関数やフィッシャー情報量に関する結果は、既存の文献結果をかなり改善しており有用であると判断される。第 4 章および第 5 章における適合度検定については、提案されている統計量が一般的な検定統計量であり、t 分布などの特定の対立仮説に関してかならずしも良い検出力を持たない点が指摘された。

論文審査の結論：松井氏は、大学院課程を通じて、裾の重い分布族として応用上も重要な安定分布の研究に取り組み、複素積分上の工夫や数値計算上の工夫を駆使して、安定分布の扱いに関する多くの理論的結果を導いた。これらの結果は、安定分布を用いた統計的モデリングを支える基礎的な技術として重要なものであり、今後の応用も期待できる。このような意味で、本論文は本研究科が要求する課程博士の基準を十分に満たしていると考えられる。したがって、この審査委員会は、本論文を博士（経済学）の学位を授与するにふさわしいと全員一致で判断した。

2006 年 1 月

審査委員：竹村彰通（主査）
国友直人
矢島美寛
久保川達也
大森裕浩