

【書 評】

林 高樹／佐藤彰洋

『金融市場の高頻度データ分析』

— データ処理・モデリング・実証分析 —

朝倉書店 2016.7 viii+196 ページ

高度な情報技術の発展は膨大な量の観測データを金融市場にもたらしている。そして本書序文の冒頭に記されているように、ビッグデータという「Volume(量)」、「Variety(多様性)」、「Velocity(速度)」の「3V」で特徴づけられるデータの分析は、金融市場における価格形成過程の解明、市場の取引ルールや制度変更が市場にもたらす影響などを明らかにすると期待されている。しかしそれらの解明には、統計学や計量経済学が提供してきた標準的な分析手法だけでなく、高頻度データ固有の問題に対応した適切な統計手法と、金融市場の制度やルールを理解した上での分析が必要となる。そのような価格形成過程に対する取引メカニズムの影響を分析する市場のマイクロストラクチャーとよばれる研究分野は理論研究にとどまることなく、高頻度データの統計的分析をともない急速な進展をとげている。ビッグデータの分析という名のもとで統計分析手法の紹介のみにとどまっている著作物が多い中で、本書はマイクロストラクチャー理論による価格形成過程のモデルに関する丁寧な記載も与えられており、この分野を代表する良書と言える。

はじめの2つの章では金融市場で観測される高頻度データとはどのようなものであり、それらを分析することの意義が記されている。実際に金融市場で観測されるデータは統計学の教科書に載っているように整然と規則正しく観測されるものではない。例えば、ある一つの銘柄に関しても、売り手と買い手がそれぞれの立場で提示する売り気配、買い気配、さらに約定値など複数の価格系列が存在するし、複数の銘柄に着目すれば、それらは同じタイミングで取引されることは稀で、観測される時点が異なる非同期データとなっている。これら最初の2章は、金融市場で観測されているデータがもっている特性を理解した上で統計分析を実践することの重要性を認

識させてくれる導入部となっている。さらに続く第3章では観測系列のプロットという基本的な手法をもちいて、分析対象の高頻度データがもつ特性を具体的に引き出して見せている。そして描き出された特性から、どのような統計モデルを構築するかといった着想を得るに至るプロセスを想起させる予備的分析が紹介されている。

第4章で説明される価格変動を記述するモデルはランダム・ウォークモデルである。この基本となるモデルをもとに、ランダム・ウォーク仮説の検証法と実証例が与えられている。実証例では観測時間間隔を15秒、30秒、1分、5分、10分、30分とした6種類の価格系列がもちいられている(分析は価格自体ではなく、収益率に関して行われる)。時間間隔が短くなるほど高頻度に観測される、このようなデータが高頻度データとよばれているものに該当する。そしてこの高頻度データではランダム・ウォーク仮説は棄却され、時間間隔を10分、30分とした系列では棄却されないという結果が示される。ランダム・ウォーク仮説が棄却される高頻度収益率系列では負の自己相関が観測されることがよく知られているが、その理由を説明するモデルとしてマイクロストラクチャー理論の基本的なモデルであるRollモデルが紹介されている。Rollモデルは収益率系列が負の自己相関を示す仕組みを単純明瞭に表現したモデルであり、金融市場の価格形成過程の解明には統計的な分析技術だけでなく、市場の取引ルールや制度に関する理解も重要であることを示す好例になっている。

第5章ではボラティリティの変動を記述するモデルとして、いわゆるARCH系のモデルを、そして第6章では取引間隔の変動を記述するモデルとして、ARCH系モデルと同じ流れをくむACDモデルを取り上げている。ARCH系のボラティリティ変動モデルは高頻度データをもちいなくても比較的推定が容易であるが、高頻度データを分析対象とした場合にはモデリング上、相応の拡張が必要となる。第5章ではそのような拡張に関する先行研究の例示と解説が与えられている。ボラティリティが直接観測されない量であるのに対して、取引時間間隔や取引頻度(件数)は観測可能な量であり、市場の活況の程度をとらえる上でも興味深いものである。そのような時間間隔や頻度の系列は点過程とよばれている。こ

の点過程のモデリングは標準的な統計学や計量経済学のテキストでは学ぶ機会が少ない主題であり、そのモデリングと分析方法を事例とともに学べるという点でも魅力的な章になっている。

高頻度データを利用したボラティリティや相関の計測に関しては第7章で取り扱われている。1日内のティックデータが利用可能な場合、第5章のARCH系のような特定のパラメトリックモデルを考えなくても、高頻度に観測される日内の収益率のデータを利用してボラティリティの推定は可能である。高頻度データを利用したボラティリティ推定に関しては実現ボラティリティ(Realized Volatility)をはじめ非常に多くの研究成果があるが、マイクロストラクチャー・ノイズの影響も含め、わかりやすい形で整理されている。

複数の金融資産の収益率間の相関は、実務上も研究上も注目される重要な量である。その計量には共分散の推定が必要だが、既に述べたように利用される高頻度データは非同期観測データであるため、適切な拡張を施した共分散の推定方法が必要となる。本書の著者の一人である林氏は代表的な推定量である Hayashi-Yoshida 推定量の考案者の一人でもある。余談ではあるが、評者がこの分野に興味を持ったきっかけは、Epps 効果(観測時間間隔を短くしていくと計測される相関係数の絶対値が小さくなる傾向)とその原因と考えられる非同期観測データに対する共分散推定量の構築についての林氏によるセミナー講演であった。

ボラティリティ、相関に続いて金融リスクの管理において重要な主題であるテール・リスクの評価が第8章で扱われている。統計学における極値理論の応用分野としても研究成果が期待される分野でもある。極値理論は発生確率が低い極端な事象の分布やそのような事象のモデリングに関するものである。文字どおり、極端な事象の発生頻度は低く、対応する統計モデルの理解には多くの理論的な予備知識を必要とするが、この章では分析目的を見失わせることのないように簡潔な説明とともに実証例が与えられる。さらに第8章では、自己相似性やスケールリング則という興味深い内容であるものの、応用例を伴う邦文による教科書的な解説は多く見かけない主題も取り扱っている。

第9章は高頻度にデータが観測される代表的な市場である外国為替市場の分析にあてられている。第3章で説明された探索的データ分析を実践するとこ

ろからはじまり、本書で扱われた分析手法をもちいた分析結果が示されている。さらにこの第9章では他章では扱われていないパワー・スペクトラムを使った分析が行われている。時系列分析は時間領域(time domain)における分析と周波数領域(frequency domain)における分析にわけることができるが、これまでの実証研究は時間領域における分析が主流であった。低・中程度の観測頻度のデータによる周波数領域での分析では興味深い分析結果を得られることが多くないことがその主な理由としてあげられる。一方、高頻度データを利用した周波数領域での分析では興味深い分析結果が期待されており、第9章での実証例はその一つである。

最後の第10章はエージェント・モデルに関するものである。エージェント間の相互作用の結果として集団現象が創発されるという発想は、マイクロストラクチャー理論が目指している個々の投資家の行動と市場との関連を解明することと同義といってもよい。しかし、マイクロストラクチャー理論が統計分析による理論の検証を分析の主眼に置いているのに対し、エージェント・モデルは理論の検証というよりも数値シミュレーションによる分析という印象が強い。いずれにせよ市場における投資家の行動原理の解明に資するべく研究が進展している分野であり、興味を持つ読者にはそれらを学べる格好の章であろう。

以上、全10章について見てきたが、本書にはさらにウェブサイトでのサポートページが提供されている。そこでは本書を読み進めていくうえで必要とされる知識を補うための補足説明やデータ分析にもちいられる R や Perl のコードが用意されている。これらのコードは入門的なものの例示だけだが、高頻度データ分析の実践の第一歩をふみだす方々には大きな手助けとなるものである。また本書で与えられた数多くの参考文献はこの分野に興味をもつ者たちへの最良の道標となっている。本書の読者らによって、日本におけるこの分野の研究がより進展していくことは間違いないであろう。

[大屋幸輔]