

## 「特集 高頻度金融データに基づく統計的推測と モデリング」について

川崎 能典<sup>†</sup>・萩原 哲平<sup>†</sup> (オーガナイザー)

「統計数理」における金融データの統計解析に関連した過去の特集としては、2002年12月号(第50巻第2号)の「ファイナンス統計学」、2011年6月号(第59巻第1号)の「金融リスクの統計解析」がある。これら間口の広いテーマ設定に比較して、今回の「高頻度金融データに基づく統計的推測とモデリング」は、かなり絞り込まれた感があるが、にもかかわらず8本の論文が集まっているところに、この分野の研究の活発さが反映されていると見ることができよう。

金融市場のマイクロ構造(microstructure)に関する仮説形成は、経済学や金融論においては古くから行われてきたわけだが、それらの仮説が検証されるには、データの利用可能性を待たなければならなかった。学説史的な観点から厳密に研究の源流を探る、というような作業は筆者の能力と専門性を越えているので、極めて私的な印象論を述べれば、1990年代の半ばにOlsen & Associates社が、1992年10月から1993年9月の1年分ではあったが外国為替市場の高頻度データを、少なくとも学術研究機関には極めて安価に提供し、合わせて国際会議をホストすることで、高頻度金融データの分析を促進したことは特筆に値する。こうした時代の研究がまとめられたものとして、Dunis and Zhou (1998)やDacorogna et al. (2001)といった論文集がある。また、金融市場のマイクロ構造に関する基本文献であるO'Hara (1995)も同時期の出版である。

さて、本特集は、高頻度金融データの利用を前提とした、確率過程の推測や時系列モデリングを扱っているが、分析の切り口を幾つかのキーワードを挙げながら取り上げておきたい。

高頻度金融データによってもたらされた最大のインパクトは、分単位、秒単位、あるいはそれ以下の観測間隔で得られるデータを集約することで、日次(に限らないが典型的には日次)のボラティリティが容易に得られるようになったことである。いわゆる実現ボラティリティ(realized volatility, RV)である。研究テーマとしてのRVは、市場のマイクロ構造に起因するノイズ(microstructure noise)を避けながらRVを得るにはどうすればよいか、という観点から、多数の問題を統計学者に提供してきたと言える。こうした流れの中で、日次データに基づく非線形・非ガウスフィルタリングといった手法は、かつてほどの重要性を持たなくなったと言えるだろう。

RV計算に絡むもう一つの現実的問題点は、価格プロセスに含まれるジャンプである。ジャンプの存在は、RVの計算に当然大きな影響を与えるわけだが、成り行き注文に起因するbid-askバウンスのようなノイズと、ジャンプの影響を排してボラティリティを推定する、という問題に対しては、ここまでさまざまな研究が行われている。

一方、観測頻度という観点からも、高頻度金融データは新たな切り口を必要としている。それは、データが原則不等間隔(irregularly spaced)にしか観測されないことである。モデリングの方向性としては少なくとも二通り考えられるが、ひとつは取引と取引の生起間隔(duration)に着目した時系列モデルを作成することである。ARCH/GARCH系のモデルからの類推で導入

---

<sup>†</sup>統計数理研究所：〒190-8562 東京都立川市緑町 10-3

された, 生起間隔に関する条件付き自己回帰モデル (Autoregressive Conditional Duration, ACD モデル) とその変種は, 計量ファイナンス (financial econometrics) の分野で盛んに研究された. Hautsch (2004) に詳しい. もうひとつは, 点過程 (point process) として扱い, その条件付き強度ないし生起度関数 (intensity function) のモデリングを行うというやり方である.

ひとつの資産に関する取引レコードが不等間隔データであるということは, 複数の資産の相関を考える際にはいわゆる非同期観測の問題が発生する. 複数資産のポートフォリオ管理においては, 相関の計算は基本的なことであるが, データの補間で便宜的に同期させて RV を求めると, 観測期間が小さくなるにつれてバイアスが生じる (Epps 効果). この点を克服する推測理論は, Hayashi and Yoshida (2005) を嚆矢に, 以後活発に提案されている.

以下では, 本特集に採録された論文の内容を概観しておこう. 最初の 4 編は, 高頻度金融データを扱うための基礎となる統計理論に関連する話題を取り扱っている.

荻原論文は, 連続的な path をもつ確率過程である拡散過程による日内株価のモデリングと分散・共分散等のリスク量の統計推測手法の理論研究を詳解している. 特に高頻度金融データ特有の問題をモデルに組み込み, それを考慮した統計理論を紹介している. 上原・増田論文では, 変動にジャンプを含む Lévy 過程を用いたモデリングを研究しており, Lévy 過程の係数であるスケール係数とドリフト係数を段階的に推定する手法を研究している. 内田論文では, 拡散過程モデルのパラメータ推定において, 確率過程のドリフト項の係数と拡散係数のパラメータを分けて推定し, また, 最適化関数を簡易なものから段階的に尤度関数に近づける方法により, 安定的で高速な推定手法を研究している. 深澤論文では, 拡散過程モデルにおいて拡散過程自体が高頻度観測されるのではなく, その時間積分値が高頻度観測される場合の拡散係数項の推定手法を論じている. これはファイナンスにおいては株価ボラティリティの推定に対応するが, 物理学における分子運動の Langevin モデルとも関係が深い.

後半の 4 編では日本株式市場における実証分析に関連する話題を取り扱っている. 田代・川口論文では, 東京証券取引所の個別株式の板情報を多次元 Hawkes 過程と呼ばれる点過程を用いてモデリングし, 高速な株式売買システム (arrowhead) の導入前後における注文タイプ別の発生頻度の傾向を分析している. 林論文では, 東京証券取引所と 2 つの私設証券取引所において取引される個別株式について, 取引所間の株価変動の先行遅延関係について調べている. 吉田論文は, 実現ボラティリティに価格のジャンプが与える影響を, 切断実現ボラティリティとの比較で実証的に分析した結果を報告している. 容易に想像される通り, 観測時間間隔の取り方と閾値の選択は, 分析結果を大きく左右する. 森本・川崎論文は, 日次実現ボラティリティに対する予測モデルとして, 経験類似度に基づくある種のモデル平均法を軸に, heterogeneous autoregression (HAR) モデルを主たる比較対象に, それらの予測能力を内挿・外挿の両面からシミュレーションで検討した結果を報告している.

この巻頭言前半に列挙した論点からすると, この 8 編の論文で取り上げられているテーマは必ずしも網羅的とは言えないが, 2017 年時点でのスナップショットとしての価値は確実に有するものと自負している. 寄稿して下さった研究者の多くは, 統計数理研究所の客員教授・准教授として, あるいは共同利用研究員としての登録のある方である. 共同利用・共同研究の成果を論文として結実できたことに, オーガナイザーとしてこの場を借りて感謝申し上げたい. 本特集が, 読者諸兄の興味を少しでも喚起できたなら, 編者・著者として望外の幸せである.

## 参 考 文 献

- Dacorogna, M. M., Gençay, R., Muller, U., Olsen, R. B. and Pictet, O. V. (2001). *An Introduction to High-frequency Finance*, Academic Press, San Diego.

- Dunis, C. and Zhou, B. (eds.) (1998). *Nonlinear Modeling of High Frequency Financial Time Series*, John Wiley & Sons, Chichester.
- Hautsch, N. (2004). *Modelling Irregularly Spaced Financial Data*, Springer-Verlag, Berlin.
- Hayashi, T. and Yoshida, N. (2005). On covariance estimation of nonsynchronously observed diffusion processes, *Bernoulli*, **11**, 359–379.
- O’Hara, M. (1995). *Market Microstructure Theory*, John Wiley & Sons, Cambridge, Massachusetts.