Ultra high frequency data and statistical inference: Back to the continuous-time paradigm

Nakahiro Yoshida (University of Tokyo and CREST JST, Japan)
nakahiro@ms.u-tokyo.ac.jp

Abstract

High-frequency financial data is one of the latest objects to be challenged by the most advanced statistics. Ad hoc methods often mislead and fail data analysis. For two stochastic processes observed high-frequently and asynchronously, the estimated covariance between them cannot correctly be captured by the realized covariance with a "natural" interpolation method. It is recognized that the non-synchronicity of sampling schemes, though it is inevitable for real financial data, causes such phenomena generically called the Epps effect. Theory of non-synchronous estimation has been developing in the last decade and successfully applied to actual data analyses.

Market microstructure is another main factor that causes the Epps effect. Remarkable progresses were recently made in volatility estimation problems by proposing effective filters that remove microstructure noises and at the same time treat non-synchronous sampling schemes. Non-synchronicity and microstructure are now the point where theoretical statistics, probability theory and real data analysis are confluent.

The latest issue is modeling of ultra high frequency phenomena by point processes. This enables us to model microstructure itself rather than eliminating it as noise. In ultra high frequency sampling, the central limit theorem does not work and there is no longer Brownian motion as the driving factor of asset prices, differently from the standard mathematical finance. The world of real data is already beyond a standard theory based on Brownian or other drivers, but this is the reality statisticians are confronted with. Developments in measurement and storage technologies are enhancing feasibility of continuous-observation paradigm in finance, and statistical inference for stochastic processes is advancing back to there.

In this talk, we introduce a point process regression model that enables us to express asynchronicity of observations, lead-lag relation and microstructure. Our model can describe self-exciting/self-correcting effects of the point processes as well as exogenous effects. Non-ergodic statistics is constructed in the QLA (quasi likelihood analysis) framework. The point process regression model has applications to price models and limit order book models. This is a joint work with Teppei Ogihara.