Preface

Statistics for stochastic processes is a topic in full development, driven by the needs of various applied fields, such as finance, bioscience or telecommunication. This volume of the series "Lévy Matters" is completely dedicated to this topic.

From an historical perspective, the topic started with the situation where the process under consideration is completely observed over some time interval $[0, T]$, and in the asymptotic theory the time horizon T goes to infinity. However, except for point or marked point processes, for which the times and sizes of the jumps are quite often observed, complete observation of the path over some time interval is possible in very rare cases only. Under almost all practical circumstances, a process can only be observed at discrete times, often equally spaced, or sometimes irregularly spaced. Recent mathematical advances allow us to deal with this situation of a discretely observed stochastic process, at least when this process has a nice structure, such as being a semimartingale or even an Itô semimartingale, and when the sampling scheme is regular. Irregular sampling schemes have also been considered, but they pose new challenges, especially when the sampling times are endogenous, that is, depend on the process itself. The Itô semimartingale assumption may appear, and is a serious mathematical restriction, but most models used by practitioners are of this type, because they are solutions of a stochastic differential equation driven by a Lévy process, or by a Brownian motion and a Poisson random measure.

A comprehensive statistical analysis of discretely observed Itô semimartingales is still far from being complete. However, the simplest semimartingales are Lévy processes, so a first step is to understand as well as possible the situation, when the underlying process is a Lévy processes observed at the times $i\Delta_n$ for $i = 0, 1, \ldots, n$, with a mesh size Δ_n, which can be a constant (we then speak of *low frequency* observations), or is small and eventually goes to 0 as $n \to \infty$ (the *high frequency* setting). This setting is simple enough to allow for the development of efficient statistical tools, which hopefully can be extended to more general semimartingales, and it also plays the role of a benchmark, since any statistical procedure which works for semimartingales should *a fortiori* work for Lévy processes.

These reasons motivate the editing and writing of this volume, whose aim is to provide a rather extensive account on the most recent developments in the field of statistics for discretely observed Lévy processes.

Let us now be more specific. A Lévy process has a rather simple structure, as its law is completely characterized by three ingredients: the variance σ^2 of the Gaussian part, the drift b, and the Lévy measure F which describes the structure of the jumps, so the statistical problems amount to getting some information on the triple (b, σ^2, F). (This is in deep contrast with the general semimartingale case, for which the characteristics are *a priori* random, thus inducing non-standard statistical problems, where the "parameters" to estimate may be random.) So the main question here is how to estimate, in one way or another, the parameters b and σ^2, and also a parameter which may describe the family of Lévy measures in the model, or the measure F itself in a non-parametric way.

The observed increments $X_{i\Delta_n} - X_{(i-1)\Delta_n}$ are indeed i.i.d. variables, whose law only depends on the triple (b, σ^2, F) and on the mesh Δ_n. So in the low frequency setting $\Delta_n = \Delta$ we theoretically are on the known ground of the observation of an i.i.d. sample of variables. However, even in this case the problem is not trivial since we are after (b, σ^2, F) which, although in one-to-one correspondence with the distribution function of X_Δ (or with its density when it exists), has almost never an explicit form given in terms of these. In the high-frequency case, the observed increments are i.i.d., but their laws depend on n and become degenerate as $n \to \infty$.

The variance σ^2 has the distinctive property that it can be consistently estimated, in principle with the rate $\sqrt{n}$, whatever the asymptotic behaviour of Δ_n. In contrast, consistently estimating b and F requires $T_n := n\Delta_n \to \infty$, and the rates depend on T_n rather than n itself (typically $\sqrt{T_n}$ for b, whereas for F the rates are more complex to describe and strongly depend on the assumed hypotheses on F such as being a parametric family or a non-parametric family of finite measures, or other types of assumptions).

The three chapters below consider the statistical problem under different view-points:

Chapter "Estimation and Calibration of Lévy Models via Fourier Methods": D. Belomestny and M. Reiß study the low frequency situation. The method is based on the empirical characteristic function; they show in particular that estimators based on the empirical characteristic function enjoy rate-optimality for the two parameters b and σ^2, and also the optimal (minimax) non-parametric rate for the Lévy measure, mostly (but not only) in the case when the Lévy measure is finite. They also study the estimation of the so-called Blumenthal-Getoor index, which is a number in $[0, 2]$ and measures the degree of concentration of F near the origin, and the estimation when the observed process is not a Lévy process *stricto sensu*, but a time-changed Lévy process, which of course allows for a much wider range of applications.

Chapter "Adaptive Estimation for Lévy Processes": F. Comte and V. Genon-Catalot consider the non-parametric estimation, mainly in the high-frequency case with a time horizon $n\Delta_n$ going to infinity. All three quantities b, σ^2 and F are studied, although the emphasis is rather on the Lévy measure, assuming that it has a density, and as a general rule they find non-parametric estimators with better rates than in chapter “Estimation and Calibration of Lévy Models via Fourier Methods”, when expressed in terms of $T_n = n\Delta_n$. This is plausible, because in this situation one clearly has a better handgrip on the real size of the jumps. A distinctive feature of this part is the construction of *adaptive* estimators, based on deconvolution or projection or kernel methods.

Chapter “Parametric Estimation of Lévy Processes”: H. Masuda, in contrast with the other authors, considers a completely parametric situation, when all three components c, σ^2 and F depend on a—possibly multidimensional—parameter θ. The emphasis is on maximum likelihood estimation and the Local Asymptotic Normality (LAN) property with a careful analysis of the various rates (for the drift, the diffusion and/or the Lévy measure), at which this property holds in the high-frequency case. He also proposes a method based on the median of suitably chosen functions of the observed increments, proving that (as for most methods of moments) it is rate-efficient. A large number of concrete example are treated in detail, showing how an actual implementation is possible.

Overall, these three chapters cover the main aspects of the estimation of discretely observed Lévy processes, when the observation scheme is regular, from an up-to-date viewpoint. We hope that the reader will find here a solid background on which statistical procedures for more general stochastic processes can be developed.

Aarhus, Denmark — Ole E. Barndorff-Nielsen
Zurich, Switzerland — Jean Bertoin
Paris, France — Jean Jacod
Munich, Germany — Claudia Küppelberg
August 2014

http://www.springer.com/978-3-319-12372-1

Lévy Matters IV
Estimation for Discretely Observed Lévy Processes
Belomestny, D.; Comte, F.; Genon-Catalot, V.; Masuda, H.; Reiß, M.
2015, XV, 286 p. 21 illus., 14 illus. in color., Softcover
ISBN: 978-3-319-12372-1