Estimation methods for Levy based models of asset prices

José Enrique Figueroa-López

Financial mathematics seminar
Department of Statistics and Applied Probability
UCSB
October, 2006

Abstract

Stock prices driven by *Levy processes* or other related jump processes have received a great deal of attention in recent years. The scope of these models goes from simple *exponential Levy models* to *SDE with Poisson jumps both on the volatility and on the returns.* Several calibration methods have been proposed in the literature to deal with these models. In this talk we will review some classical methods, such as approximated maximum likelihood using FFT, and some recent methods using asymptotic limits of "power" variations. In particular, I will discuss in some detail nonparametric procedures.

Outline of key points

- The problem of statistical inference of continuous-time stochastic models:
  - Brief description: Parameters, Estimation, and Testing of Hypothesis.
  - Sampling schemes: continuous, discrete, high-frequency asymptotics and long-term horizon.
  - Parametric vs. non-parametric methods
  - Latent "variables": stochastic volatility, jump times and jumps sizes.
- Standard parametric methods:
  - Maximum likelihood estimation
  - Methods of moments
  - Bayesian methods: MCMC methods.
- "Nonparametric" methods based on high-frequency sampling:
  - Quadratic variation
  - Bipower variation

- Testing for jumps
- Threshold quadratic variations: Disentangling jumps
- General variations: An application to the non-parametric estimation of the Lévy density
- Discussion about robustness and practicality: Microstructure of asset prices.

Remarks on statistical inference methods for continuous-time models

Brief description: Parameters, Estimation, and Testing of Hypothesis

As the goal of modeling is to construct a **stochastic process** that accounts for the known statistical features of stock prices, the goal of statistical methods is to make inferences about the **parameters** that control the stochastic model based on **observations** of the process. In general there are two types of statistical inferences: **estimation and hypothesis testing**.

Infinite-dimensional and finite-dimensional specifications of the model: Parametric and non-parametric models.

The choice of the model usually may involve different levels of specification. When the model involves "functional" parameters, one can specify an explicit form determined by finitely-many parameters (*parametric model*) or specify only some qualitative properties (*non-parametric model*); e.g. monotonicity:

Example 1. *In a Lévy model where the Lévy measure is assumed to be determined by a Lévy density, one can consider parametric forms for the Lévy density (e.g.* a CGMY model*), non-parametric (e.g. bounded continuous monotonic density), or semiparametric (e.g. a tempered stable model).*

Parametric vs. non-parametric inference methods

Parametric methods enjoy better qualities (e.g. smaller standard error or greater power) when the the model is adequate to describe the sample, but they are model dependent and thus exposed to *mis-specification errors* (that is, errors arising when the model does not accurately account for the statistical features of the data; e.g. when some conditions are violated by the data). Also, they are generally more computational expensive.

Example 2. *Compare a maximum likelihood estimation where many times an optimization numerical scheme is needed vs. a histogram method.*

Pearson pointed out that the price we have to pay for pure parametric fitting is the possibility of gross mis-specification resulting in too *high a model bias.* On the other hand, Fisher pointed out that the nonparametric approach gave generally poor efficiency and estimates with high variability, especially for small sample size.

Sampling schemes: continuous vs discrete. High-frequency and long-horizon qualities of the methods.

While in principle continuous-time sampling is unrealistic and unfeasible, inference analysis based on this scheme is useful since the results obtained are usually more powerful and can serve as benchmarks of what can be accomplish by procedures based on discrete-time sampling. Sometimes continuous-time based statistics function as devises to construct feasible statistics through an approximation idea. This last paradigm has sometimes the bonus of allowing us to compute the total "error" by decomposing it into the error of the continuous-time based procedure plus the error of the approximation scheme.

Example 3. *Suppose that some procedure is based on the statistics* $\int_0^T g(t)\, dX_t$*, where* $\{X_t\}_{t\geq 0}$ *is the underlying process and* g *is a deterministic function. Then,* $\sum_{i=1}^n g(t_i)(X_{t_i} - X_{t_{i-1}})$ *could be used to construct feasible procedures. Similarly, imagine a procedure involves the statistics* $\sum_{t\leq T} g(\Delta X_t)$*, where* ΔX_t *is the jump of* X *at time* t*, then* $\sum_{i=1}^n g(X_{t_i} - X_{t_{i-1}})$ *could be used as an approximation.*

Another important point is to determine the qualities (e.g. convergence and robustness) when the frequency and the time-horizon of the sampling increase. Some procedures require high-frequency data (guaranteing convergence as the frequency increases) and other require long-horizon sampling (guaranteing convergence in the long run).

Latent variables

Most of modern financial models involve *latent variables* that are not directly observable.

Example 4. *All the following examples have latent variables that cannot be computed from discrete-time sampling (though some can be computed from continuous-time observations):*

1. *In a stochastic volatility model, the volatility is a latent variable.*
2. *In a Lévy process, the jump sizes and times are latent.*
3. *In a time-changed Lévy process, the random clock is not available.*

Sometimes is useful to devise statistical methods assuming that these latent variables are available and then, use the data to "approximate" them.

Standard parametric methods

Maximum likelihood estimation for Lévy processes

Description of method.

1. General idea: The maximum likelihood method postulates that the most sensible values of the parameters are those that maximize the likelihood of the observations; that is, the values that maximize the probability that the random variables take on values "close" those observed in the sample.

2. Suppose that $R_\delta, \ldots, R_{n\delta}$ are n consecutive log returns every δ (a chosen time span). In the geometric Lévy model[1], $R_\delta, \ldots, R_{n\delta}$ is a *random sample* (that is, i.i.d. observations), and the likelihood function is

$$L(\theta) = f_\delta(r_1; \theta) \ldots f_\delta(r_n; \theta),$$

where $r_1, \ldots, r_n$ are the values the returns take on the sample, $f_\delta(\cdot; \theta)$ is the density function of $R_\delta = X_\delta$, and θ are the parameters of this density. The *Maximum Likelihood Estimator* is defined by

$$\hat{\theta} = \mathrm{argmax}_\theta \; L(\theta).$$

Drawbacks and solutions

1. Lévy-based models are described in terms of the Lévy density. We usually specify a parametric form for the Lévy measure, and not the density f_δ. As a consequence, the characteristic function of X_δ is known in closed form, but the density f_δ is unknown. Even if f_δ enjoys a closed form, this might be intractable.

Example 5. *In the variance Gamma Lévy process the Lévy density is given by*

$$p(x) = \begin{cases} \alpha |x|^{-1} e^{-\frac{|x|}{\beta_-}}, & \text{if } x < 0, \\ \alpha x^{-1} e^{-\frac{x}{\beta_+}}, & \text{if } x > 0, \end{cases} \Longrightarrow f_\delta(x) = c_1 |x|^{\alpha\delta - \frac{1}{2}} e^{-c_2 x} K_{\alpha\delta - \frac{1}{2}}(c_3 |x|)$$

The CGMY Model has Lévy density

$$p(x) = \begin{cases} \alpha |x|^{-\nu} e^{-\frac{|x|}{\beta_-}}, & \text{if } x < 0, \\ \alpha x^{-\nu} e^{-\frac{x}{\beta_+}}, & \text{if } x > 0, \end{cases} \Longrightarrow \textit{No known closed expression}$$

One way to overcome the above problem is to evaluate numerically the density via

$$f_\delta(x\,;\theta) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-izx} \hat{f}_\delta(z\,;\theta)\, dz$$

(here, $\hat{f}$ stands for the characteristic function). However, the integral needs to be computed for each observation x_i (before called r_i), each time that the value of the likelihood $L(\theta)$ is needed in a gradient search optimization algorithm. *Carr et. al. [4]* suggests a method to approximate more efficiently the likelihood function using Fast Fourier Transform. The idea is compute $f_\delta(x\,;\theta)$ in a regular fine grid $-M = y_{-M}, \ldots, y_0 = 0, \ldots, y_M = M$ of the interval $[-M, M]$ (containing all observations) with mesh $\Delta := 1/M$. Then, we simply replace $f_\delta(x_i\,;\theta)$ in the likelihood function by $f_\delta(y_j\,;\theta)$, where y_j is the point of the grid closest to x_i.

[1] Recall in this model the price of the stock is given by $S_t := S_0 \exp X_t$, where X_t is a Lévy process and the log returns are just $R_{i\delta} = \log(S_{i\delta}/S_{(i-1)\delta}) = X_{i\delta} - X_{(i-1)\delta}$

Example 6. *The FFT and inverse FFT are implemented in MatLab. According to the manual, the vectors* $X = \text{fft}(x)$ *and* $x = \text{ifft}(X)$ *are related through the formulas*

$$X(k) = \sum_{j=1}^{N} x(j) w_N^{(j-1)(k-1)}$$
$$x(j) = \frac{1}{N} \sum_{k=1}^{N} X(k) w_N^{(j-1)(k-1)},$$

where $j, k = 1, \ldots, N$ *and* $w_N := e^{(-2\pi i)/N}$*. Using this functions, one can check that*

$$f\left(2\pi \frac{k-1}{N}\right) = \frac{1}{2\pi}\left(X_1(k) + N x_2(k)\right),$$

where $X_1 := \text{fft}(x_1)$*,* $x_2 := \text{ifft}(X_2)$*,* $x_1(j) := \hat{f}(j\Delta)$ *and* $X_2(k) := \hat{f}(-k\Delta)$*.*

2. Another procedure to approximate the likelihood function in the case of high-frequency data is to use the asymptotic result:
$$\lim_{\delta \to 0} \frac{1}{\delta} f_\delta(x) = p(x),$$
where p is the Lévy density of the process (that is, the Lévy measure is assumed to be of the form $p(x)dx$). The asymptotic behavior can be improved. (see *Cont and Tankov [7]* for more details and references).

3. It is believed (see e.g. *Cont [6]*) that four parameters[2] are needed to appropriately describe asset prices. The resulting likelihood functions are in general flat and non-concave, and hence it is hard and computational expensive to attain convergence to the ML estimators. The numerical procedure tends to be instable.

Example 7. Madan, Carr, and Chang *[5] calibrate a variance Gamma model to daily returns of the S&P500 index from Jan. 92 to Sept. 94 (691 observations). Let us remark that in this case, a closed form for the density function is available and thus, we don't need to invert the characteristic function. They reported ML estimates for the annualized mean return (m),* volatility *(s), and* kurtosis parameter *(*ν*) of* 5.91%, 11.72% *and* 0.002 *(then, the resulting kurtosis is* 5.19 *vs. the value* 3 *of the normal distribution). The estimates above corresponds to the annualized estimates* $(\hat{\theta}, \hat{\sigma}, \hat{\nu}, \hat{b}) = (-0.00056, 0.1171, 0.002, 0.05)$*. We simulate 691 daily observations with the above presumed estimates and perform MLE using the standard optimization functions provided by MATLAB. Using as initial estimates* $(\hat{\theta} - 0.001, \hat{\sigma} - 0.001, \hat{\nu} + 0.001, \hat{b} + 0.001)$*, MLE yields the estimates* $(-0.001, 0.1162, 0.003, 0.0538)$*. However, using as initial estimate the method of moments estimate* $(0, 0.11, 0.001, 0.01)$*, the algorithm did not converge, which usually happen.*

Other parametric methods

Methods of moments. *See Handbook of financial econometrics* available on the web.
Bayesian methods: MCMC methods. *See Handbook of financial econometrics* available on the web.

[2]accounting for location, scale, decay of tails, and skewness

Nonparametric procedures based on high-frequency sampling

In this part, we survey some non-parametric procedures to approximate or even extract latent variables so that the parameters involved in those latent variables can subsequently be estimated. Here, we assume that n observations of the prices process $\{S_t\}_t$ are taken during the time horizon $[0,t]$ at times $0 = t_0 < \cdots < t_n = t$. The mesh of the partition is defined as $\max_{0\le i\le n}\{t_i - t_{i-1}\}$. Below, X_t is the log-return process

$$X_t := \log \frac{S_t}{S_0},$$

and thus,

$$X_{t_i} - X_{t_{i-1}} \log \frac{S_{t_i}}{S_{t_{i-1}}},$$

represents the (log) return on the asset during the time interval $(t_{i-1}, t_i]$.

Realized volatility (quadratic variation): $\sum_{i=1}^{n}(X_{t_i} - X_{t_{i-1}})^2$.

Setup and Conditions: X is a semimartingale with continuous part M^c; that is, X has the decomposition

$$X = \underbrace{M}_{\text{Loc. Mart.}} + \underbrace{B}_{\text{B.V.}} = \underbrace{M^c}_{\text{Cont. Mart.}} + \underbrace{M^d}_{\text{Pur. Disc. Mart.}} + \underbrace{B^c}_{\text{Cont. B.V.}} + \sum_{s\le t} \Delta B_s.$$

In the first decomposition, B is a process of bounded variation and M is a local martingale, which can be decomposed into a continuous local martingale M^c, and a "purely discontinuous" local martingale M^d. If M is assumed predictable, the decomposition is unique and moreover, the M^d and B jumps at different times.

Asymptotic behavior: For a semimartingale X, the following limit in probability holds true:

$$RV_n(t) := \sum_{i=1}^{n}(X_{t_i} - X_{t_{i-1}})^2 \stackrel{\text{mesh}\to 0}{\longrightarrow} \langle M^c, M^c\rangle_t + \sum_{s\le t} (\Delta X_s)^2\,,$$

where M^c is the continuous local martingale part of X and ΔX_t is the jump of X at time t.

Applications to some models proposed in the literature:

Lévy process: Suppose that $X = \sigma W + Z$, where W is a standard Brownian motion and Z is the jump part of X (that is, Z is the limit of compensated compound Poisson processes plus a compound Poisson process):

$$RV_n(t) \stackrel{\text{mesh}\to 0}{\longrightarrow} \sigma^2 t + \sum_{s\le t} (\Delta Z_s)^2\,,$$

The expression above explains the effect that the jumps has on the realized (or empirical) volatility of an asset.

Stochastic volatility with Lévy jumps:

$$X_t = \int_0^t \mu_s ds + \int_0^t \sigma_s dW_s + Z_t,$$

where Z_t is a pure-jump Lévy process. Then, the realized volatility has the asymptotic behavior:

$$RV_n(t) \stackrel{\text{mesh}\to 0}{\longrightarrow} \int_0^t \sigma_s^2 ds + \sum_{s\le t} (\Delta Z_s)^2 .$$

Stochastic volatility with finite-activity jumps:

$$X_t = \int_0^t \mu_s ds + \int_0^t \sigma_s dW_s + \sum_{i=1}^{N_t} J_i,$$

where

- N_t is a counting process, finite for any t;
- $\{J_i\}_{i\ge 1}$ arbitrary r.v.'s representing the jumps of the process;
- σ is a stochastic process with càdlàg (right-continuous with left limits) paths such that $\int_0^t \sigma_s^2 ds < \infty$;
- μ is a locally bounded predictable process.

Then,

$$RV_n(t) \stackrel{\text{mesh}\to 0}{\longrightarrow} \int_0^t \sigma_s^2 ds + \sum_{i=1}^{N_t} J_i^2.$$

Stochastic volatility with jumps driven by a Lévy process:

$$\begin{aligned} X_t = &\int_0^t \mu_s ds + \int_0^t \sigma_s dW_s + \sum_{|\Delta Z_t|>1} h(s, \Delta Z_s) \\ &+ \lim_{\varepsilon\to 0} \left\{ \sum_{s:\varepsilon<|\Delta Z_t|\le 1} h(s,\Delta Z_s) - \int_0^t \int_{\varepsilon<|z|\le 1} h(s,z)F(dz)ds \right\}, \end{aligned}$$

where Z_t is a pure-jump Lévy process with Lévy measure F and h is an appropriate deterministic function with $h(0,\cdot) = 0$. Notice that $\Delta X_t = h(t, \Delta Z_t)$, and thus, the jumps of the process are driven by a Lévy process (for the details see the recent paper by *Ait-Sahalia and Jacod*).

Then,

$$RV_n(t) \stackrel{\text{mesh}\to 0}{\longrightarrow} \int_0^t \sigma_s^2 ds + \sum_{s\le t} (\Delta X_s)^2.$$

References: One of the oldest nonparametric methods based on high-frequency sampling (see e.g. Jacod and Shiryaev book [9]).

Bipower variations: $\sum_{i=2}^{n} \left|X_{t_i} - X_{t_{i-1}}\right| \left|X_{t_{i-1}} - X_{t_{i-2}}\right|$.

Setup and Conditions: We assume the setup of the *stochastic volatility model with finite-activity jumps* or *stochastic volatility model with jumps driven by Lévy processes* previously described.

Asymptotic behavior: The following limit in probability is satisfied for *regular*[3] partitions $0 = t_0 < \cdots < t_i := (t/n)i < \cdots < t_n = t$:

$$BPV_n(t) := \sum_{i=2}^{n} \left|X_{t_i} - X_{t_{i-1}}\right| \left|X_{t_{i-1}} - X_{t_{i-2}}\right| \overset{n\to\infty}{\longrightarrow} k^2 \int_0^t \sigma_s^2 ds,$$

where $k = \mathbb{E}|\mathcal{N}(0,1)|$ is a known constant.

References: This concept was introduced by *Barndorff-Nielsen and Shephard*, and the limit behaviors have been proved through joint work with *Podolskij, Winkel, and Jacod* (see [3] and [11] and the references therein). For the proof in the setup of the jumps driven by Lévy processes see the recent work by *Ait-Sahalia and Jacod* [1].

Testing for jumps.

Based on Bipower Variation: See *Barndorff-Nielsen and Shephard [3]* and also, the thesis *Podolskij [11]*.

1. One of the applications of Bipower variations was to develop statistical tests for the presence of jumps. The key idea is to observe that

$$D_n(t) := RV_n(t) - \frac{1}{k^2} BPV_n(t) \overset{n\to\infty}{\longrightarrow} \underbrace{\sum_{s\leq t} (\Delta X_t)^2}_{\text{Aggregated squared jumps}},$$

   under say the model of "stochastic volatility with finite-activity jumps" or "the stochastic volatility model with jumps driven by a Lévy process". Therefore, $D_n(t)$ "significantly" positive for large n is evidence of jumps. To make precise the meaning of significantly positive, it is necessary to determine the asymptotic distribution of the statistic.

2. Within the stochastic volatility model with finite-activity jumps, the following asymptotic limit follows under the "null-hypothesis" that the "process is continuous":

$$\sqrt{\frac{n}{c\int_0^t \sigma_s^4 ds}} \left(RV_n(t) - \frac{1}{k} BPV_n(t) \right) \overset{\mathcal{L}}{\longrightarrow} \mathcal{N}(0,1).$$

   Moreover, the statistic

$$I_n(t) := \frac{1}{k^4} \sum_{i=4}^{n} \left|X_{t_i} - X_{t_{i-1}}\right| \ldots \left|X_{t_{i-3}} - X_{t_{i-4}}\right|$$

[3] Equally spaced partition points.

converges in probability to $\int_0^t \sigma_s^4 ds$, and hence, the following feasible statistic test

$$Z_n(t) := \sqrt{\frac{n}{c\, I_n(t)}} \left(RV_n(t) - \frac{1}{k}\, BPV_n(t) \right),$$

is asymptotically standard normal as $n \to \infty$. Then, *we reject the null-hypothesis that X is continuous* if $Z_n(t) > z_\alpha$ where z_α is an appropriate normal quantile (namely, $\mathbb{P}[\mathcal{N}(0,1) \geq z_\alpha] = \alpha$).

Higher-power based method: See *Ait-Sahalia and Jacod [1].*

- Key results: Consider the $p-$power variation

  $$\widehat{B}_n^p(t) := \sum_{i=1}^n \left| X_{t_i} - X_{t_{i-1}} \right|^p$$

  with $p > 2$.

  1. $\widehat{B}_n^p(t) \stackrel{n\to\infty}{\longrightarrow} \sum_{s\leq t} |\Delta X_s|^p$, for any semimartingale X.
  2. $n^{p/2-1} \widehat{B}_n^p(t) \stackrel{n\to\infty}{\longrightarrow} m_p\, t^{p/2-1} \int_0^t |\sigma_s|^p\, ds$,
     if X is continuous and follows the *stochastic volatility model with jumps driven by a Lévy process.*

- Main statistics and asymptotics: For $p > 2$ and a positive integer k,

  $$\frac{\widehat{B}_n^p(t)}{\widehat{B}_{kn}^p(t)} \stackrel{n\to\infty}{\longrightarrow} \begin{cases} 1, & \text{if } X \text{is discontinuous,} \\ k^{1-p/2}, & \text{if } X \text{is continuous.} \end{cases}$$

- Rate of convergence:

  $$\sqrt{n} \left(\frac{\widehat{B}_n^p(t)}{\widehat{B}_{kn}^p(t)} - 1 \right) \stackrel{\mathcal{L}}{\longrightarrow} Z$$

  where Z is a centered non-degenerate r.v. (actually, with conditionally normal r.v.), under further structural conditions on σ (essentially requiring that σ is measurable with respect to the Brownian motion, the Lévy process, and possibly other independent Brownian motions).

Truncated variations: $\sum_i (X_{t_i} - X_{t_{i-1}})^2 \mathbf{1}_{\left\{\left|X_{t_i} - X_{t_{i-1}}\right| \leq r(\text{mesh})\right\}}$.

References: See *Mancini [10].*

Setting and conditions:

- Assume *The stochastic volatility model with finite-activity jumps.*
- Equally-spaced returns ($t_i = hi$, where $h = t/n =$ Time-span).

Fundamental result: There exists a r.v. $H > 0$ such that for all $0 < h \leq H$:

$$\text{There was a jump on } (t_i, t_{i+1}] \iff \left|X_{t_i} - X_{t_{i-1}}\right| > c(h).$$

Here, $c(h)$ is a deterministic functions s.t.

$$c(h) \xrightarrow{h \to 0} 0 \quad \text{and} \quad \frac{\sqrt{h \log \frac{1}{h}}}{c(h)} \xrightarrow{h \to 0} 0,$$

(e.g. $c(h) = h^{\alpha/2}$ with $0 < \alpha < 1$).

Consequences:

Total volatility:

$$TRV_n(t) := \sum_{i=1}^{n} (X_{t_i} - X_{t_{i-1}})^2 \, \mathbf{1}_{\{|X_{t_i} - X_{t_{i-1}}| \leq c(h)\}} \xrightarrow{n \to \infty} \int_0^t \sigma_s^2 ds.$$

Disentangling jumps:

$$\sum_{i=1}^{n} \varphi(X_{t_i} - X_{t_{i-1}}) \, \mathbf{1}_{\{|X_{t_i} - X_{t_{i-1}}| > c(h)\}} \xrightarrow{n \to \infty} \sum_{i=1}^{N_t} \varphi(J_i).$$

Central Limit Theorem and Confidence intervals:

$$\frac{1}{\sqrt{I_n(t)}} \left(TRV_n(t) - \int_0^t \sigma_s^2 ds \right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, 2/3).$$

where $I_n(t) := \sum_{i=1}^{n} (X_{t_i} - X_{t_{i-1}})^4 \, \mathbf{1}_{\{|X_{t_i} - X_{t_{i-1}}| \leq c(h)\}}$

Non-parametric estimation of the Lévy density based on general realized variations

References: *Figueroa-Lopez & Houdré [8]*

Setting: Consider an exponential Lévy model driven by a Lévy process $\{X_t\}_{t \geq 0}$ with Lévy measure of the form $F(dx) = p(x)dx$.

Goal: Devise methods to estimate *directly* the Lévy density p.

General ideas:

- ☐ The function p can approximately be recovered from integrals of the form

$$\int \varphi(x) p(x) dx,$$

by taking test functions φ on certain classes of bases; e.g. indicator functions, splines, wavelet basis, etc.

☐ p controls the jump behavior of the process:

$$\mathbb{E}\,\frac{1}{t}\sum_{s\le t}\mathbf{1}_{\{a\le \Delta X_s\le b\}} = \int \mathbf{1}_{[a,b]}(x)\,p(x)dx$$
$$\Downarrow$$
$$\mathbb{E}\,\frac{1}{t}\sum_{s\le t}\varphi\left(\Delta X_s\right) = \int \varphi(x)p(x)dx$$

☐ p determines also the "short-term behavior" of X:

$$\frac{1}{h}\mathbb{E}\left[\varphi\left(X_h\right)\right] \xrightarrow{h\downarrow 0} \int \varphi(x)p(x)dx.$$

Statistics:

$$I_{n,\varphi}\left(t\right) := \sum_{k=1}^{n}\varphi\left(X_{t_k} - X_{t_{k-1}}\right), \quad \text{and} \quad I_{\varphi}(t) := \sum_{s\le t}\varphi(\Delta X_s)$$

Asymptotics:

Convergence in law: $I_{n,\varphi}\left(t\right) \xrightarrow{\mathfrak{D}} I_{\varphi}(t)$

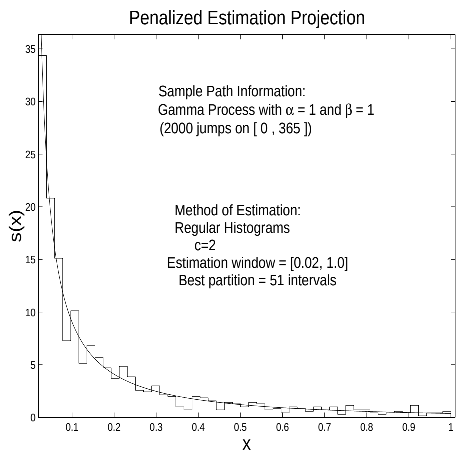
Asymptotic unbiasedness: $\mathbb{E}\left[\frac{1}{t}I_{n,\varphi}\left(t\right)\right] \xrightarrow{n\to\infty} \int \varphi(x)p(x)dx$

Asymptotic variance: $\mathrm{Var}\left[\frac{1}{t}I^{d}_{n,\varphi}\left(t\right)\right] \xrightarrow{n\to\infty} \frac{1}{t}\int \varphi^2(x)p(x)dx \xrightarrow{t\to\infty} 0.$

Example: The Gamma Lévy process

Model: Pure-jump Lévy process with Lévy density $p(x) = \frac{\alpha}{x}e^{-x/\beta}\mathbf{1}_{\{x>0\}}$.

Histogram like estimators: Outside the origin.



Performance:

- Least-square fit:
  Fit the model $\frac{\alpha}{x}e^{-x/\beta}$ (using least-squares) to the histogram estimator: $\hat{\alpha}_{LSE} = 0.93$ and $\hat{\beta}_{LSE} = 1.055$ (vs. $\hat{\alpha}_{MLE} = 1.01$ and $\hat{\beta}_{MLE} = 0.94$)
- Sampling distribution
  Means and standard errors of $\hat{\alpha}_{LSE}$ and $\hat{\beta}_{LSE}$ based on 1000 repetitions

Δt	PPE-LSF		MLE	
.1	0.81 (0.06)	1.40 (0.50)	1.001 (0.01)	0.99 (0.05)
.01	0.92 (0.08)	1.12 (0.31)	1.007 (0.07)	0.99 (0.08)
.001	0.93 (0.08)	1.13 (0.34)	1.007 (0.07)	0.99 (0.08)

Example: One-sided Tempered stable distribution

Model: Pure-jump Lévy process with Lévy density $p(x) = \frac{a}{x^{\alpha+1}} e^{-x/b} \mathbf{1}_{\{x>0\}}$.

Histogram type estimators:

- Paths generated from 36500 jumps on $[0, 365]$ with $a = b = 1$ and $\alpha = .1$.
- α estimated by Zolotarev method for stable distributions.
- Least-square fit to quantify the quality of the histogram estimator

Δt	Penalized Projection - Least-Squares Fit			Misspecified Gamma MLE	
.01	1.03 (0.15)	0.97 (0.14)	0.09 (0.0002)	1.2 (0.08)	0.89 (0.079)

Table 1: Sampling mean and standard errors (sample size=100 paths)

Example: One-sided Tempered stable distribution

Model: Pure-jump Lévy process with Lévy density $p(x) = \frac{a}{x^{\alpha+1}} e^{-x/b} \mathbf{1}_{\{x>0\}}$.

Histogram type estimators:

- Paths generated from 36500 jumps on $[0, 365]$ with $a = b = 1$ and $\alpha = .1$.
- α estimated by Zolotarev method for stable distributions.
- Least-square fit to quantify the quality of the histogram estimator

Δt	Penalized Projection - Least-Squares Fit			Misspecified Gamma MLE	
.01	1.03 (0.15)	0.97 (0.14)	0.09 (0.0002)	1.2 (0.08)	0.89 (0.079)

Table 2: Sampling mean and standard errors (sample size=100 paths)

Discussion about robustness and feasibility

- One of the main drawbacks of high-frequency methods is the so-called microstructure noise (e.g. the stock prices do not take arbitrary values).
- Therefore, it is imperative to analyze robustness of the methods towards "microstructure noise" and towards departures from the semimartingale assumption (which has been shown to be violated at the tick-by-tick level).

- There are several models for tick-by-tick data, but the "bridge" between these models and semimartingale type model is not well-understood yet.
- How frequent to sample? There is a tradeoff: The higher frequency, the smaller the error of the non-parametric methods (under absence of noise), but the higher the microstructure noise. (see e.g. *Ait-Sahalia, Mykland, Zhang [2]* for some partial answers).

References

[1] Y. Aït-Sahalia and J. Jacod. Testing for jumps in a discretely observed process. Technical report, Princeton University and Université de Paris VI, September 2006. Availlable on web.

[2] Y. Aït-Sahalia, P.A. Mykland, and L. Zhang. How often to sample a continuous-time process in the presence of market microstructure noise. *The review of financial studies*, 18(2), February 2005.

[3] O.E. Barndorff-Nielsen and N. Shephard. Econometrics of testing for jumps in financial economics using bipower variation. *Journal of Financial Econometrics*, 4(1):1–30, 2006.

[4] P. Carr, H. Geman, D. Madan, and M. Yor. The fine structure of asset returns: An empirical investigation. *Journal of Business*, pages 305–332, April 2002.

[5] P. Carr, D. Madan, and E. Chang. The variance Gamma process and option pricing. *European Finance Review*, 2:79–105, 1998.

[6] R. Cont. Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, 1:223–236, 2001.

[7] R. Cont and P. Tankov. *Financial modelling with Jump Processes.* Chapman & Hall, 2003.

[8] J.E. Figueroa-Lopez and C. Houdré. Nonparametric estimation of Lévy processes with a view towards mathematical finance. November 2004. Available at [ArXiv math.ST/0412351].

[9] J. Jacod and A.N. Shiryaev. *Limit Theorems for Stochastic Processes.* Springer, 2003.

[10] C. Mancini. Non parametric threshold estimation for models with stochastic diffusion coefficient and jumps. July 2006. Available at [ArXiv math.ST/0607378].

[11] M. Podolskij. *New Theory on estimation of integrated volatility with applications.* PhD thesis, Ruhr-Universität Bochum, April 2006. Available on the web.