Statistical methods for financial models

 driven by Lévy processes

José Enrique Figueroa-López

Department of Statistics, Purdue University

PASI

Centro de Investigación en Matemáticas (CIMAT)

Guanajuato, Gto. Mexico

May 31 - June 5, 2010
Part III: Traditional Parametric Methods for Geometric Lévy Models
Program

• Statistical inference of continuous-time models.
• Maximum likelihood estimation
• Numerical Examples:
  – Variance Gamma Model
  – Normal Inverse Gaussian Model
• Some preliminary empirical results
Statistical-methods for continuous-time models

Goal: To make inferences (predictions) about the parameters that control the statistical behavior of the stochastic model, based on observations of the stochastic process of interest.

Parametric and non-parametric models: The “parameters” of the model can be finite-dimensional or infinite-dimensional such as a function with some qualitative properties (e.g. monotonicity).

Parametric and non-parametric inference methods:

- A pure parametric fitting \( \Rightarrow \) \{ \begin{array}{l} \star \text{ Better performance} \\ \star \text{ High model bias} \\ \star \text{ Comp. demanding and instable} \end{array} \)
- Nonparametric approach \( \Rightarrow \star \text{ Methods with lower performance} \)
Observations:

- **Discrete** $\Rightarrow$ ★ The only feasible choice
- **Continuous** $\Rightarrow$ ★ More powerful results, providing benchmarks
  ★ Serve as devices to construct feasible methods

Latent variables: Hidden stochastic variables in the model that are not directly observable, requiring methods to “extract” or approximate their values.

**Model:**
- Stochastic volatility model: Volatility.
- Lévy process: Times and magnitudes of the jumps.
- Time-changed Lévy process: Random clock
Maximum Likelihood Estimation

- Geometric Lévy model:

\[ S_t = S_0 \, e^{X_t}, \]

\[ X_t \text{ is a Lévy process} \]

\[ \implies \text{Equally-spaced log-returns are i.i.d.} \]

Recall that

\[ R = \text{Log-return on } [t, t + \delta] = \log \frac{\text{Final price}}{\text{Initial price}}. \]

- Maximum Likelihood Principle:

Most sensible values of the parameters are those that maximize the likelihood of the sample observations.
• Implementation of the method:
  
  – $r_1, \ldots, r_n$ are the sample values of $n$ equally-spaced returns (with time-span $\delta$).

  – Suppose that $f_\delta(\cdot; \theta)$ is the probability density function of the log return $R_\delta$ on a time interval of $\delta$ length:

    $$
P[a < R_\delta < b] = \int_a^b f_\delta(r; \theta)dr.
    $$

  – $f_\delta(r; \theta)$ is a good proxy of the probability of observing a return of magnitude close to $r$;

  – The likelihood function is $L(\theta) = f_\delta(r_1; \theta) \ldots f_\delta(r_n; \theta)$. The Maximum Likelihood Estimator is defined by

    $$
    \hat{\theta} = \arg\max_\theta L(\theta).
    $$
Main issue: Lévy-based models are typically described in terms of their Lévy density. As a consequence, the characteristic function $\hat{f}_\delta$ is known in a closed form, but the density $f_\delta$ is many times unknown or intractable.

Possible Solutions:

– “Probabilist” solution: Inversion formula.

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

– Approximate the integral using Fast Fourier Transform.
A Numerical Example

- Consider a variance Gamma process with drift:

\[ X_t = \sigma W_{\tau_t} + \theta \tau_t + b t, \]

\( W \) standard Brownian motion and \( \tau_t \) Gamma Lévy process:

\[ \tau_1 \overset{D}{\sim} \text{Gamma} \left( \alpha = \frac{1}{\kappa}, \beta = \kappa \right). \]
Density has a close form in terms of “Bessel functions of second kind”:

\[ p_t(x) = \frac{2e^{\theta(x-bt)/\sigma^2}}{\sigma \sqrt{2\pi \kappa^t/\kappa} \Gamma\left(\frac{t}{\kappa}\right)} \left( \frac{|x - bt|}{\sqrt{\frac{2\sigma^2}{\kappa} + \theta^2}} \right)^{\frac{t}{\kappa} - \frac{1}{2}} \]

\[ \times K_{\frac{t}{\kappa} - \frac{1}{2}} \left( \frac{|x - bt| \sqrt{\frac{2\sigma^2}{\kappa} + \theta^2}}{\sigma^2} \right), \]
Moments are given in closed-form as follows:

\[ \mu_1(X_\delta) := \mathbb{E}(X_\delta) = (\theta + b)\delta, \]
\[ \mu_2(X_\delta) := \text{Var}(X_\delta) = (\sigma^2 + \theta^2 \kappa)\delta, \]
\[ \mu_3(X_\delta) := \mathbb{E}(X_\delta - \mathbb{E}X_\delta)^3 = (3\sigma^2 \theta \kappa + 2\theta^3 \kappa^2)\delta, \]
\[ \mu_4(X_\delta) := \mathbb{E}(X_\delta - \mathbb{E}X_\delta)^4 \]
\[ = (3\sigma^4 \kappa + 12\sigma^2 \theta^2 \kappa^2 + 6\theta^4 \kappa^3)\delta + 3\mu_2(X_\delta)^2. \]
Part IV: Non-parametric method based on high-frequency sampling observations
Program

- **Nonparametric methods** based on high-frequency sampling:
  - Realized volatility (quadratic variation)
  - Bipower variation
  - Testing for jumps
  - Threshold quadratic variations: Disentangling jumps
  - Application of general realized variations to the non-parametric estimation for Lévy processes
  - Discussion about robustness and feasibility: Microstructure effects.
Non-parametric methods based on high-frequency data

Set-up:

• \(S_t = \) Price of the asset at time \(t\) and \(X_t = \) (log) return on \([0, t]\):

\[
X_t := \log \frac{S_t}{S_0}.
\]

• We sample the price process during the time interval \([0, t]\) at times

\[0 = t_0 < \cdots < t_n = t\] (typically equally spaced)

• We consider the asymptotic behavior of certain statistics based on the

log-returns,

\[
\log \frac{S_{t_1}}{S_{t_0}} = X_{t_1} - X_{t_0}, \ldots, \log \frac{S_{t_n}}{S_{t_{n-1}}} = X_{t_n} - X_{t_{n-1}},
\]

as the mesh \(:= \max_i t_i - t_{i-1}\) goes to 0.

Application: To approximate (or extract) some latent variables, so that one
can subsequently construct a model and perform statistical inference.
Non-parametric Statistics

Realized volatility (quadratic variation): \( \sum_i (X_{t_i} - X_{t_{i-1}})^2 \).

Bipower variation: \( \sum_i \left| X_{t_i} - X_{t_{i-1}} \right| \left| X_{t_{i-1}} - X_{t_{i-2}} \right| \).

References: [Barndorff-Nielsen and Shephard: 2003-2006]

Threshold realized quadratic variation:
\[
\sum_i (X_{t_i} - X_{t_{i-1}})^2 \mathbf{1}\{ |X_{t_i} - X_{t_{i-1}}| \leq r(h) \}.
\]

References: [Mancini, 2003-2006]

Other realized variations: \( \sum_i \varphi \left( X_{t_i} - X_{t_{i-1}} \right) \).

References: [Jacod (2006), Figueroa-Lopez (2004)]
Realized Volatility (quadratic variation)

Conditions and notation:

\( X \) accept the following decomposition:

\[
X = M + B \Rightarrow X = M^c + M^d + B^c + \sum_{s \leq t} \Delta B
\]

\( X \) is called a **semimartingale**, \( M^c \) is its continuous part and \( \Delta X_t \) is the jump of \( X \) at time \( t \).

Asymptotic behavior:

\[
RV_n(t) := \sum_{i=1}^{n} (X_{t_i} - X_{t_{i-1}})^2 \xrightarrow{mesh \to 0} \langle M^c, M^c \rangle_T + \sum_{s \leq t} (\Delta X_s)^2.
\]
Applications:

Lévy process: \( X = \sigma W + Z \), where \( W \) is a standard Brownian motion and \( Z \) is the pure-jump part of \( X \):

\[
RV_n(t) \xrightarrow{\text{mesh} \to 0} \sigma^2 t + \sum_{s \leq t} (\Delta Z_s)^2.
\]

Notice that \( \frac{1}{t} \mathbb{E}[RV_n(t)] = \sigma^2 + \int z^2 F(dz) \).

Stochastic volatility with Lévy jumps:

\[
X_t = \int_0^t \mu_s ds + \int_0^t \sigma_s dW_s + \underbrace{Z_t}_{\text{PURE-JUMP LEVY}},
\]

\[
RV_n(t) \xrightarrow{\text{mesh} \to 0} \int_0^t \sigma_s^2 ds + \sum_{s \leq t} (\Delta Z_s)^2.
\]
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

- \( N_t = \# \) of jumps up to \( t \) (counting process), finite for any \( t \);
- \( \{J_i\}_{i \geq 1} \) arbitrary r.v.’s representing the jumps of the process;
- \( \sigma \) is a stochastic process with rcll paths such that \( \int_0^t \sigma_s^2 \, ds < \infty \);
- \( \mu \) is a locally bounded predictable process.

Then,

\[ RV_n(t) \xrightarrow{\text{mesh} \to 0} \int_0^t \sigma_s^2 \, ds + \sum_{i=1}^{N_t} J_i^2. \]
Stochastic volatility with jumps driven by a Lévy process:

\[ X_t = \int_0^t \mu_s ds + \int_0^t \sigma_s dW_s + \sum_{s \leq t: |\Delta Z_s| > 1} h(s, \Delta Z_s) \]

\[ + \lim_{\varepsilon \to 0} \left\{ \sum_{s \leq t: |\Delta Z_s| \leq 1} h(s, \Delta Z_s) - \int_0^t \int_{\varepsilon < |z| \leq 1} h(s, z) F(dz) ds \right\}, \]

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 \( \Delta X_t = h(t, \Delta Z_t) \).

\[ RV_n(t) \xrightarrow{\text{mesh} \to 0} \int_0^t \sigma_s^2 ds + \sum_{s \leq t} \{h(s, \Delta Z_s)\}^2. \]
Bipower variation

Possible settings and conditions:

- Assume *The stochastic volatility model with finite-activity jumps* or *The stochastic volatility model with jumps driven by Lévy processes*.
- Equally-spaced returns \( (t_i = \frac{t}{n} \ i, \text{ for } i = 1, \ldots, n) \)

Asymptotic behavior: The following limit in probability is satisfied:

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

where \( k = \mathbb{E}|N(0, 1)| \).

References: *Barndorff-Nielsen, Shephard, Podolskij and Winkel*. For the second setting see a very recent work by *Ait-Sahalia and Jacod*. 
Testing for jumps


- **Idea:**
  \[ D_n(t) := RV_n(t) - \frac{1}{k^2} BPV_n(t) \xrightarrow{n \to \infty} \sum_{s \leq t} (\Delta X_t)^2 \]

- **CLT:** Within the stochastic volatility model with finite-activity jumps,
  \[ \sqrt{n} \frac{1}{c \int_0^t \sigma_s^4 ds} \left( RV_n(t) - \frac{1}{k^2} BPV_n(t) \right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1). \]

under “null-hypothesis” that the process is continuous.
• Feasible test: The statistic

\[ I_n(t) := \frac{1}{k^4} \sum_{i=4}^{n} |X_{t_i} - X_{t_{i-1}}| \cdots |X_{t_{i-3}} - X_{t_{i-4}}| \]

converges in probability to \( \int_0^t \sigma_s^4 ds \), and thus,

\[ Z_n(t) := \sqrt{\frac{n}{c I_n(t)}} \left( RV_n(t) - \frac{1}{k^2} BPV_n(t) \right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1). \]

Then, we reject the null-hypothesis that \( X \) is continuous if \( Z_n(t) > z_\alpha \) where \( z_\alpha \) is such that \( \mathbb{P}[\mathcal{N}(0, 1) \geq z_\alpha] = \alpha \).
Higher-power based method: [Ait-Sahalia and Jacod, 2006]

• **Key results:** Consider the $p$—power variation

\[
\hat{B}_{n}^{p}(t) := \sum_{i=1}^{n} |X_{t_i} - X_{t_{i-1}}|^p
\]

with $p > 2$.

1. $\hat{B}_{n}^{p}(t) \xrightarrow{n \rightarrow \infty} \sum_{s \leq t} |\Delta X_s|^p$, for any semimartingale $X$.

2. $n^{p/2-1} \hat{B}_{n}^{p}(t) \xrightarrow{n \rightarrow \infty} m_p \ t^{p/2-1} \int_{0}^{t} |\sigma_s|^p \, ds$, 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 $k$,

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

• **Rate of convergence:**

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

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

Setting and conditions:

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

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

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 function 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:**

$$TRV_n(t) := \sum_{i=1}^{n} (X_{t_i} - X_{t_{i-1}})^2 \mathbf{1}_{\{|X_{t_i} - X_{t_{i-1}}| \leq c(h)\}} \xrightarrow{n \to \infty} \int_{0}^{t} \sigma^2_s ds.$$  

**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:**

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

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

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 \( \varphi \) on certain classes of bases; e.g. indicator functions, splines, wavelet basis, etc.
General ideas: Cont.

- $p$ controls the jump behavior of the process:

$$\mathbb{E} \frac{1}{t} \sum_{s \leq t} 1_{\{a \leq \Delta X_s \leq b\}} = \int 1_{[a,b]}(x) p(x) \, dx$$

$$\Downarrow$$

$$\mathbb{E} \frac{1}{t} \sum_{s \leq t} \varphi(\Delta X_s) = \int \varphi(x) p(x) \, dx$$

- $p$ determines also the “short-term behavior” of $X$:

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

\[ I_{n, \varphi}(t) := \sum_{k=1}^{n} \varphi(X_{t_k} - X_{t_{k-1}}), \quad \text{and} \quad I_{\varphi}(t) := \sum_{s \leq t} \varphi(\Delta X_s) \]

Asymptotics:

Convergence in law: \[ I_{n, \varphi}(t) \xrightarrow{D} I_{\varphi}(t) \]

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

Asymptotic variance:

\[ \text{Var} \left[ \frac{1}{t} I_{n, \varphi}^d(t) \right] \xrightarrow{n \to \infty} \frac{1}{t} \int \varphi^2(x)p(x)dx \xrightarrow{t \to \infty} 0. \]
Example: The Gamma Lévy process

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

**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

<table>
<thead>
<tr>
<th>$\Delta t$</th>
<th>PPE-LSF</th>
<th>MLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>.1</td>
<td>0.81 (0.06) 1.40 (0.50)</td>
<td>1.001 (0.01) 0.99 (0.05)</td>
</tr>
<tr>
<td>.01</td>
<td>0.92 (0.08) 1.12 (0.31)</td>
<td>1.007 (0.07) 0.99 (0.08)</td>
</tr>
<tr>
<td>.001</td>
<td>0.93 (0.08) 1.13 (0.34)</td>
<td>1.007 (0.07) 0.99 (0.08)</td>
</tr>
</tbody>
</table>
Example: One-sided Tempered stable distribution

Model: Pure-jump Lévy process with Lévy density

\[ p(x) = \frac{a}{x^{\alpha+1}} e^{-x/b} 1_{\{x>0\}}. \]

Histogram type estimators:

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

<table>
<thead>
<tr>
<th>(\Delta t)</th>
<th>Penalized Projection - Least-Squares Fit</th>
<th>Misspecified Gamma MLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>.01</td>
<td>1.03 (0.15)</td>
<td>0.97 (0.14)</td>
</tr>
</tbody>
</table>

Table 1: Sampling mean and standard errors (sample size=100 paths)
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] for some partial answers).