Distribution of maximal deviation for Lévy density estimators

Vladimir Panov

Higher School of Economics (Moscow, Russia)

Joint work with Valentin Konakov

LSA Winter meeting - 2017
Snegiri, 4. December

Jump-type dynamics

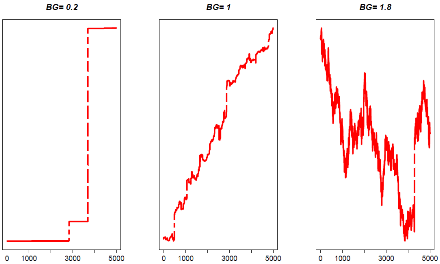
How to describe jumps $\Delta X_t = X_t - X_{t-} \neq 0$ of a Lévy process $X = (X_t)_{t \geq 0}$:

Lévy measure of X_t is defined by

$$\nu(B) := \mathbb{E}\left[\sharp \left\{ t \in [0,1] : \Delta X_t \in B, \ \Delta X_t \neq 0 \right\} \right], \qquad B \subset \mathbb{R}.$$

Blumenthal-Getoor index is equal to

$$\mathrm{BG}(X) := \inf \left\{ r > 0 : \int_{|x| \leq 1} |x|^r \nu(dx) < \infty \right\} \in [0,2].$$



Setup

Let $(X_t)_{t\geq 0}$ be a Lévy process with Lévy triplet (μ, σ, ν), and assume that ν has a density s, that is,

$$\nu(B) = \int_B s(u)du, \quad B \in \mathcal{B}(\mathbb{R}).$$

Data: assume that some discrete equidistant observations $X_0, X_\Delta, ..., X_{n\Delta}$ of the process X_t are available.

Aim: $$(X_\Delta, ..., X_{n\Delta}) \Longrightarrow \{s(x),\ x \in D\}, \qquad D = [a, b] \subset \mathbb{R}/\{0\}.$$

High-frequency data: $\Delta = \Delta_n \to 0, \quad T = n\Delta_n \to \infty$ as $n \to \infty$.
Comte & Genon-Catalot (2013), Figueroa-López (2011), Figueroa-Lopéz & Houdré (2006)

Low-frequency data: $\Delta -$ fixed , $\quad T = n\Delta$.
Nickl & Reiss (2013), Gugushvili (2012), Belomestny (2011), Comte & Genon-Catalot (2010), Chen, Delaigle & Hall (2010), Neumann & Reiss (2009), van Es, Gugushvili & Spreij (2007)

Aim of the research

Bounds for quadratic risk: for an estimate $\hat{s}_n^{\circ}(x)$ and a collection of Lévy processes $\mathcal{T}$,

$$\sup_{\mathcal{T}} \mathbb{E}\left(\hat{s}_n^{\circ}(x) - s(x)\right)^2 \leq f(n) \to 0, \qquad \forall x \in D,$$

$$\inf_{\{\hat{s}_n(x)\}} \sup_{\mathcal{T}} \mathbb{E}\left(\hat{s}_n(x) - s(x)\right)^2 \geq g(n) \to 0, \qquad \forall x \in D,$$

where by $\{\hat{s}_n(x)\}$ we denote the set of all estimates of the Lévy density $s(x)$.

Our aim: find the distribution of maximal deviation of $\hat{s}_n^{\circ}(x)$, that is, the cdf of

$$MD_n := \sup_{x \in D} \left(\frac{|\hat{s}_n^{\circ}(x) - s(x)|}{\sqrt{s(x)}} \right).$$

Source of inspiration: V.Konakov and V.Piterbarg' 84
- study the kernel estimator of the regression function;
- prove that the rate of convergence to the asymptotic distribution given in P.Bickel and M.Rosenblatt' 73 is very slow (of logarithmic order);
- obtain a sequence of distribution laws which approximate the MD distribution with power rate of convergence.

Collections of basis functions

Consider $\{\varphi_r(x) : D \to \mathbb{R},\ r = 1..d\}$ - an orthonormal collection in $\mathcal{L}^2(D)$.
Project $s(x)$ on the space $S := \text{span} < \varphi_1(x), ..., \varphi_d(x) >$:

$$\tilde{s}(x) := \sum_{r=1}^{d} \beta_r \varphi_r(x), \qquad \text{where} \qquad \beta_r = \beta(\varphi_r) = \int_D \varphi_r(u) s(u) du.$$

For any $m \in \mathbb{N}$ there exists a set of normalized bounded functions $\left\{\psi_j^m : D \to \mathbb{R}\right\}_{j=0}^{J}$ supported on $[a, a+\delta)$, where $\delta = (b-a)/m$, such that

$$\left\{\varphi_r(x),\ r = 1..d\right\} \quad = \quad \left\{\psi_j^m\left(x - \delta(p-1)\right) \mathbb{I}\left\{x \in I_p\right\}, j = 0..J,\ p = 1..m\right\},$$
$$\text{where} \qquad I_p := [a + \delta(p-1), a + \delta p).$$

<u>**Main construction:**</u> basis on $[a, a+\delta)$ is constructed from a basis $\{\widetilde{\psi}_j\}_{j=0..J}$ on some "standard" interval $[\tilde{a}, \tilde{b}]$ by changing the variables:

$$\psi_j^m(x) = \sqrt{\frac{\tilde{b} - \tilde{a}}{\delta}} \cdot \widetilde{\psi}_j\left(\frac{(\tilde{b} - \tilde{a})(x-a)}{\delta} + \tilde{a}\right).$$

Examples

(i) Trigonometric basis on $[0, 2\pi]$:

$$\left\{\widetilde{\psi}_j(x)\right\} = \left\{ \frac{1}{\sqrt{2\pi}}, \quad \sqrt{2}\cos(jx), \quad \sqrt{2}\sin(jx), \quad j = 0..J\right\}.$$

(ii) Legendre polynomials on $[-1, 1]$:

$$\left\{\widetilde{\psi}_j(x)\right\} = \left\{\sqrt{(2j+1)/2} \cdot P_j(x), \quad j = 0..J\right\}.$$

where $P_j(x) = (j!2^j)^{-1}\left[\left(x^2-1\right)^j\right]^{(j)}$, $j = 0..J$.

(iii) Wavelets, for instance Haar wavelets on $[0, 1]$:

$$\left\{\widetilde{\psi}_j(x)\right\} = \left\{1, \quad \mathbb{I}\{x \in [1/2, 1]\} - \mathbb{I}\{x \in [0, 1/2]\}\right\}.$$

Assumption: $\psi_j^m(x)$ depends on m such that

$$\sup_{x \in I_1} |\psi_j^m(x)| \leq C_1\sqrt{m}, \qquad \lim_{\substack{a = x_0 < \ldots < x_n = a+\delta \\ \max_i |x_i - x_{i-1}| \to 0}} \sum_{i=1}^{n} \left|\psi_j^m(x_i) - \psi_j^m(x_{i-1})\right| \leq C_2\sqrt{m}.$$

Estimation of the coefficients β_r

Recall: $\tilde{s}(x) := \sum_{r=1}^{d} \beta_r \varphi_r(x)$, where $\beta_r = \beta(\varphi_r) = \int_D \varphi_r(x) s(x) dx$.

High-frequency setup: Wörner' 03, Figueroa-López' 04:

$$\hat{\beta}(\varphi_r) := \frac{1}{T} \sum_{k=1}^{n} \varphi_r \left(X_{k\Delta} - X_{(k-1)\Delta} \right).$$

Finally we get the estimate:

$$\boxed{\hat{s}_n(x) := \frac{1}{T} \sum_{r=1}^{d} \left[\sum_{k=1}^{n} \varphi_r \left(X_{k\Delta} - X_{(k-1)\Delta} \right) \right] \varphi_r(x).}$$

In low-frequency setup, this idea fails because

$$\frac{1}{\Delta} \cdot \frac{1}{n} \sum_{k=1}^{n} \varphi_r \left(X_{k\Delta} - X_{(k-1)\Delta} \right) \to \frac{1}{\Delta} \mathbb{E} \left[\varphi_r \left(X_\Delta \right) \right]$$

$$= \frac{1}{\Delta} \int_D \varphi_r(x) F_\Delta(dx) \neq \int_D \varphi_r(x) s(x) dx.$$

Assumptions

1. **Small-time asymptotics:** Generally, Corollary 8.9 from Sato "Lévy processes and infinitely divisible distibutions" yields that

$$\hat{\beta}(\varphi_r) \xrightarrow[n\to\infty]{\mathbb{P}} \beta(\varphi_r).$$

We assume that there exist positive k and Δ^0 such that

$$\sup_{x\in D}\left|\frac{1}{\Delta}\mathbb{P}\left\{X_\Delta \geq x\right\} - \nu\left([x,+\infty)\right)\right| < k\Delta, \qquad \forall\ \Delta < \Delta^0.$$

For instance, this property is fulfilled when s is Lipshitz in an open set containing D and uniformly bounded on $|x| > q$ for any positive q (see Figueroa-López' 11).

2. **Relation between $n, m, T \to \infty$:** in what follows, we assume that

   - $T = n^{\varkappa}$ for some $\varkappa \in (0,1)$.
   - $\Lambda_n := m\frac{\sqrt{\log n}}{n^{\varkappa/2}} \to 0$ as $n, m \to \infty$.

Theorem 1. Reformulation in terms of Gaussian processes

Denote

$$Z_n := \sup_{x \in D} \left(\frac{|\hat{s}_n(x) - \mathbb{E}\hat{s}_n(x)|}{\sqrt{s(x)}} \right), \qquad F_{Z_n}(u) := \mathbb{P}\left\{ Z_n \leq u \right\}.$$

Then there exist positive constants $c_1, c_2, \lambda_1, \lambda_2$ *such that*

$$\begin{aligned} F_{Z_n}\left(\sqrt{\frac{m}{T}}u\right) &\leq \left[F_\zeta\left(\sqrt{m}u + c_1\sqrt{m}n^{-\lambda_1}\right)\right]^m + c_2 n^{-\lambda_2}, \\ F_{Z_n}\left(\sqrt{\frac{m}{T}}u\right) &\geq \left[F_\zeta\left(\sqrt{m}u - c_1\sqrt{m}n^{-\lambda_1}\right)\right]^m - c_2 n^{-\lambda_2}, \end{aligned}$$

where by $F_\zeta(\cdot)$ *we denote the distribution function of the r.v.*

$$\zeta = \zeta^{J,m} := \sup_{x \in [a, a+\delta)} \left| \Upsilon^{J,m}(x) \right|, \qquad \Upsilon^{J,m}(x) := \sum_{j=0}^{J} \xi_j \psi_j^m(x)$$

with i.i.d. standard normal r.v.'s ξ_j, $j = 0..J$.

Supremum of Gaussian process: Markus & Shepp' 71

For any centered Gaussian process G_t with a.s. bounded sample paths,

$$\lim_{u\to\infty} \frac{\log \mathbb{P}\left\{\sup_{t\in K} G_t \geq u\right\}}{u^2} = -\frac{1}{2\sigma_K^2}, \qquad K \subset \mathbb{R},$$

where $\sigma_K^2 = \sup_{t\in K} \mathbb{E} G_t^2$. For a single Gaussian variable $\xi \sim N\left(0, \sigma^2\right)$:

$$\lim_{u\to\infty} \frac{\log \mathbb{P}\left\{\xi \geq u\right\}}{u^2} = -\frac{1}{2\sigma^2}.$$

Main term in the asymptotics of $\mathbb{P}\{\xi \geq u\}$:

$$\mathbb{P}\left\{\xi \geq u\right\} = \frac{\sigma}{u\sqrt{2\pi}} \exp\left(-\frac{u^2}{2\sigma^2}\right)\left(1 + o(1)\right), \quad u \to +\infty,$$

which does not necessary coincide with the asymptotics of $\mathbb{P}\{\sup G_t \geq u\}$.

Some related results: *Borell' 75, Cirelson, Ibragimov & Sudakov' 76, Samorodnitsky '87, Talagrand '94.*

Theorem 2: Asymptotic behaviour of $\zeta^{J,m}$

Let u grow with m so that $u/\sqrt{m} \to \infty$. Then it holds

$$\mathbb{P}\left\{\zeta^{J,m} \geq u\right\} = 2\frac{h_1 \; m^{k/2}}{u^k} \exp\left\{-h_2 \; u^2/m\right\} \left(1 + \tau\left(u/\sqrt{m}\right)\right),$$ *where*

- $\tau(x) \to 0$ *as* $x \to \infty$;
- $k = 0$ *for trigonometric basis,* $k = 1$ *for Legendre pol. and wavelets;*
- $h_1, h_2, \tau(x)$ *depend on the basis, but do not depend on m.*

Proof crucially depends on the properties of the covariance function:

(i) trigonometric basis: the process is stationary, and moreover the covariance function of $X(t)$ has the asymptotics $r(t) = 1 - \frac{1}{2}t^2 + \alpha t^4 + o(t^4), t \to 0$ $\Longrightarrow$ Pickands theorem or related techniques;

(ii) Legendre polynomials: variance attains its maximum only in finite number of points $\Longrightarrow$ double sum method;

(iii) wavelets: direct calculation.

Theoretical background: *V.Piterbarg' 96. Asymptotic methods in the theory of Gaussian processes and fields.*

Non-stationary case (Legendre polynomials)

Main assumption: let the variance function attain its maximum in finite amount of points $x_{\max}^{(1)}, \ldots, x_{\max}^{(q)}$:

$$\sigma^2(x) = \sum_{j=0}^{J} \frac{2j+1}{2} P_j^2(x), \quad x_{\max}^{(1)} = 1,\ x_{\max}^{(2)} = -1.$$

Further assumptions.

1. With some $A_j > 0, \beta_j > 0$, it holds
$\sigma(x) = 1 - \left| A_j (x - x_{\max}^{(j)}) \right|^{\beta_j} (1 + o(1)) \quad \text{as} \quad x \to x_{\max}^{(j)}, \quad j = 1..q.$
2. *Local homogeneity:*
correlation function $\rho(x, y)$ satisfies with some $C_j \neq 0, \alpha_j \in (0, 2]$,
$\rho(x, y) = 1 - C_j |x - y|^{\alpha_j} (1 + o(1)), \quad \text{as} \quad x \to x_{\max}^{(j)},\ y \to x_{\max}^{(j)}, j = 1..q$
3. *Global Hölder condition:*
there exist some $g > 0,\ G > 0$, such that
$\mathbb{E} \left(\Upsilon^{J,m}(x) - \Upsilon^{J,m}(y) \right)^2 \leq G \left| x - y \right|^g, \quad \forall x, y.$
4. for any $j_1, j_2 = 1..q,\ j_1 \neq j_2, \qquad \rho(x_{\max}^{(j_1)}, x_{\max}^{(j_2)}) < 1.$

Theorem 3: Asymptotic behaviour of the maximal deviation

For any $y \in R$,

$$\mathbb{P}\left\{\sqrt{\frac{T}{m}} \sup_{x \in D}\left(\frac{|\hat{s}_n(x) - s(x)|}{\sqrt{s(x)}}\right) \leq u_m(y)\right\} = e^{-2e^{-y}}\left(1 - 2e^{-y}R(m)\right),$$

where

- *the sequence $u_m(y)$ is equal to*

$$u_m(y) := \frac{y}{a_m} + \left(b_m - \frac{c_m}{b_m}\right),$$

with $\quad a_m := 2h_2 b_m, \quad b_m := \sqrt{h_2^{-1}\ln(h_1 m)}, \quad c_m := \frac{k}{2h_2}\ln b_m;$

- *the residual term is equal to*

$$R(m) := \tau(u_m) - \frac{k^2}{16}\frac{(\log\log m)^2}{\log m}(1 + o(1)), \quad \text{as} \quad m \to \infty.$$

Main conclusion: the rates of covergence are typically logarithmic, e.g., for trigonometric basis $R(m) \asymp C \log(m)^{-1/2}$, for wavelets $R(m) \asymp C\,(\log\log m)^2 / \log m$.

Theorem 4: Sequence of accomponaying laws

Consider the case of trigonometric basis. Define the sequence of distribution functions

$$A_m(y) := \begin{cases} \exp\left\{-2\exp\left\{-y - \frac{y^2}{4\ln(h_1 m)}\right\} - 2m\left(1 - \Phi\left(u_m\sqrt{\frac{b-a}{J}}\right)\right)\right\}, & \\ & \text{if } y \geq -b_m^{3/2}, \\ 0, & \text{if } y < -b_m^{3/2}, \end{cases}$$

where $u_m = y/(2h_2 b_m) + b_m$ *and* $b_m = \sqrt{h_2^{-1}\ln(h_1 m)}$ *, and* $\Phi(\cdot)$ *is the distribution function of the standard normal r.v. Then there exist some positive constants* $\bar{c}, \beta$, *such that for sufficiently large* n *and for any* $y \in \mathbb{R}$,

$$\boxed{\sup_{y\in\mathbb{R}} \left| \mathbb{P}\left\{ \sqrt{\frac{T}{m}} \sup_{x\in D} \left(\frac{|\hat{s}_n(x) - s(x)|}{\sqrt{s(x)}} \right) \leq u_m(y) \right\} - A_m(y) \right| \leq \bar{c}\, n^{-\beta}.}$$

Application: Construction of asymptotic confidence bands

From previous theorem it follows that

$$I_{\alpha,m} \quad := \quad \left(-\frac{k_{\alpha,m}^{(+)}}{2} + \sqrt{\frac{\left(k_{\alpha,m}^{(-)}\right)^2}{4} + \hat{s}_n(x)}, \quad \frac{k_{\alpha,m}^{(+)}}{2} + \sqrt{\frac{\left(k_{\alpha,m}^{(+)}\right)^2}{4} + \hat{s}_n(x)} \right),$$

where

$$k_{\alpha,m}^{(\pm)} := \sqrt{m/T} \left(q_{\alpha,m} a_m^{-1} + b_m \pm C_\alpha n^{-\beta} a_m^{-1} \right),$$

$C_\alpha > 0$, $q_{\alpha,m}$ - $(1-\alpha)$- quantile of the distribution function $A_m(\cdot)$, is a $(1-\alpha)-$confidence band, that is, for m large enough,

$$\boxed{\mathbb{P}\left\{ s(x) \in I_{\alpha,m}, \ \forall x \in D \right\} = 1 - \alpha.}$$

Summary

- Object of this study: projection estimates for Lévy densities in high-frequency setup.

$$\hat{s}_n(x) = \frac{1}{n\Delta} \sum_{r=1}^{d} \left[\sum_{k=1}^{n} \varphi_r \left(X_{k\Delta} - X_{(k-1)\Delta} \right) \right] \varphi_r(x), \qquad x \in D.$$

- Focus on the asymptotic properties of the distribution of max. deviation:

$$MD_n := \sup_{x \in D} \left(\frac{|\hat{s}_n(x) - s(x)|}{\sqrt{s(x)}} \right).$$

- Main idea: reformulate the problem in terms of Gaussian processes of some special type.
- We show that the exact rates of convergence are typically logarithmic, and construct the sequence of accompanying laws, which approximate the deviation distribution with polynomial rate.
- V.Konakov and V.Panov
  Sup-norm convergence rates for Lévy density estimation.
  Extremes. 2016. No. 19 (3), 371-403.

Summary

- Object of this study: projection estimates for Lévy densities in high-frequency setup.

$$\hat{s}_n(x) = \frac{1}{n\Delta} \sum_{r=1}^{d} \left[\sum_{k=1}^{n} \varphi_r \left(X_{k\Delta} - X_{(k-1)\Delta} \right) \right] \varphi_r(x), \qquad x \in D.$$

- Focus on the asymptotic properties of the distribution of max. deviation:

$$MD_n := \sup_{x \in D} \left(\frac{|\hat{s}_n(x) - s(x)|}{\sqrt{s(x)}} \right).$$

- Main idea: reformulate the problem in terms of Gaussian processes of some special type.
- We show that the exact rates of convergence are typically logarithmic, and construct the sequence of accompanying laws, which approximate the deviation distribution with polynomial rate.
- V.Konakov and V.Panov
  Sup-norm convergence rates for Lévy density estimation.
  Extremes. 2016. No. 19 (3), 371-403.

Thank you for your attention.