Adaptive wavelet estimation of a compound Poisson process

C. Duval

Paris Est-CREST

Dixième Colloque "Jeunes Probabilistes et Statisticiens"
CIRM - Marseille, 16-20 Avril 2012

Statistical setting

Definition

- Let

$$X_t = X_0 + \sum_{i=1}^{R_t} \xi_i, \quad t \geq 0$$

where the ξ_i are *i.i.d* with density f with respect to the Lebesgue measure and independent of the standard homogeneous Poisson process R of intensity $\vartheta \in (0, \infty)$.

- Suppose we have discrete data

$$X_0, X_\Delta, X_{2\Delta}, \ldots \text{ over } [0, T].$$

Objective

Estimate the density f on a compact interval $\mathscr{D} \subset \mathbb{R}$ in the microscopic regime, namely

$$\Delta = \Delta_T \to 0 \qquad \text{as } T \to \infty.$$

Heuristic

Microscopic approximation

When $\Delta_T \to 0$, most of the jumps are recovered.
Number of observed nonzero increments

$$N_T = \sum_{i=1}^{\lfloor T\Delta^{-1} \rfloor} \mathbb{1}_{\{X_{i\Delta} - X_{(i-1)\Delta} \neq 0\}} \approx R_T \approx \vartheta T.$$

Achievable rates of convergence

If X is continuously observed over $[0, T]$: $R_T \approx \vartheta T$ *i.i.d* realisations of f.
The minimax rates of convergence in –up to constants and logarithmic factors–

$$\boxed{T^{-\alpha}.}$$

Usual approach

Lévy-Khintchine formula (for compound Poisson process)

$$\mathbb{E}\left[e^{iw(X_{i\Delta}-X_{(i-1)\Delta})}\right] = \exp\left(\Delta\vartheta(\hat{f}(w)-1)\right).$$

Results (Comte and Genon-Catalot (09) or Figueroa-López (09))

Estimator attains minimax rate of convergence for the L_2 loss under the constraint

$$\boxed{T\Delta_T \leq 1 \quad \text{or} \quad T\Delta_T^2 \leq 1.}$$

Motivation

Goal

Construct an adaptive wavelet estimator that

- achieves minimax rates of convergence for the L_p loss ($p \geq 1$)
- with no constraint on Δ_T.

This model is central in many application fields

Statistical physics, queuing theory, financial series, mathematical insurance...

Plan

Statistical setting

Notation

Define

- $\mathbf{D}^{\Delta} X_i = X_{i\Delta} - X_{(i-1)\Delta}$.
- $S_i = \inf\left\{j > S_{i-1}, \mathbf{D}^{\Delta} X_j \neq 0\right\} \wedge T$.

Data

Null increments bring no information on f

$$\left(\mathbf{D}^{\Delta} X_{S_1}, \ldots, \mathbf{D}^{\Delta} X_{S_{N_T}}\right).$$

Statistical setting

Main difficulties for the estimation

- Random number of data N_T
- The law of the $\left(\mathbf{D}^{\Delta} X_{S_i}\right)$ is

$$\mathbf{P}_{\Delta}[f](x) = \sum_{m=1}^{\infty} \mathbb{P}\left(R_{\Delta} = m \middle| R_{\Delta} \neq 0\right) f^{\star m}(x) \neq f(x), \qquad \text{for } x \in \mathbb{R}.$$

For Δ small enough we have

$$1 - \Delta \leq \mathbb{P}\left(R_{\Delta} = 1 \middle| R_{\Delta} \neq 0\right) \leq 1.$$

Estimation procedure in 2 steps

Step 1 : inversion of the operator $\mathbf{P}_\Delta$

We take advantage of, for all $K \in \mathbb{N}$

$$f = \mathbf{P}_\Delta^{-1}\left[\mathbf{P}_\Delta[f]\right]$$
$$= \frac{1}{\vartheta\Delta}\sum_{m=1}^{K+1}\frac{(-1)^{m+1}}{m}(e^{\vartheta\Delta}-1)^m\mathbf{P}_\Delta[f]^{\star m} + O\left(\Delta^{K+1}\right).$$

Step 2 : Estimation of $\mathbf{P}_\Delta[f]^{\star m}$ for $m = 1, \ldots, K+1$

- Data : $\mathbf{D}_m^\Delta X_{S_i} = \mathbf{D}^\Delta X_{S_i} + \mathbf{D}^\Delta X_{S_{N_{T,m}+i}} + \cdots + \mathbf{D}^\Delta X_{S_{(m-1)N_{T,m}+i}}$, where $N_{T,m} = \lfloor N_T/m \rfloor$.
- Density estimator : wavelet threshold estimators.

Wavelet threshold density estimators

Let (φ, ψ) be a pair of scaling function and mother wavelet that generate a "suitable" basis. We have

$$f = \sum_{k \in \Lambda_0} \alpha_{0k} \varphi_{0k} + \sum_{j \geq 1} \sum_{k \in \Lambda_j} \beta_{jk} \psi_{jk},$$

where $\varphi_{0k}(\bullet) = \varphi(\bullet - k)$ and $\psi_{jk}(\bullet) = 2^{j/2} \psi(2^j \bullet - k)$ and

$$\alpha_{0k} = \int \varphi_{0k} f \qquad \beta_{jk} = \int \psi_{jk} f.$$

Consider wavelet threshold estimator of f of the form

$$\widehat{f}(x) = \sum_{k \in \Lambda_0} \widehat{\alpha}_{0k} \varphi_{0k}(x) + \sum_{j \leq J} \sum_{k \in \Lambda_j} \widehat{\beta}_{jk} \mathbb{1}_{\left\{ |\widehat{\beta_{jk}}| \geq \eta \right\}} \psi_{jk}(x), \qquad x \in \mathscr{D}.$$

Estimation of $\mathbf{P}_\Delta[f]^{\star m}$ for $m = 1, \dots, K+1$

Estimator of $\mathbf{P}_\Delta[f]^{\star m}$

Let $\eta > 0$ and $J \in \mathbb{N}\setminus\{0\}$ and define

$$\widehat{\alpha}_{0k}^{(m)} = \frac{1}{N_{T,m}} \sum_{i=1}^{N_{T,m}} \varphi_{0k}\left(\mathbf{D}_m^\Delta X_{S_i}\right) \quad \text{and} \quad \widehat{\beta}_{jk}^{(m)} = \frac{1}{N_{T,m}} \sum_{i=1}^{N_{T,m}} \psi_{jk}\left(\mathbf{D}_m^\Delta X_{S_i}\right).$$

The estimator $\widehat{P_{\Delta,m}}$ of $\mathbf{P}_\Delta[f]^{\star m}$ over $\mathscr{D}$ is for x in $\mathscr{D}$

$$\boxed{\widehat{P_{\Delta,m}}(x) = \sum_k \widehat{\alpha}_{0k}^{(m)} \varphi_{0k}(x) + \sum_{j=0}^{J} \sum_k \widehat{\beta}_{jk}^{(m)} \mathbb{1}_{\left\{|\widehat{\beta}_{jk}^{(m)}| \geq \eta\right\}} \psi_{jk}(x).}$$

We plug then into the approximation of f.

Estimator corrected at order K

Definition

We define $\widetilde{f}^K_{T,\Delta}$ the estimator corrected at order K for K in $\mathbb{N}$ and x in $\mathscr{D}$ as

$$\widetilde{f}^K_{T,\Delta}(x) = \sum_{m=1}^{K+1} \frac{(-1)^{m+1}}{m} \frac{\left(e^{\widehat{\vartheta}_T \Delta} - 1\right)^m}{\widehat{\vartheta}_T \Delta} \widehat{P_{\Delta,m}}(x),$$

where

$$\widehat{\vartheta}_T = -\frac{1}{\Delta} \log\left(1 - \frac{N_T}{\lfloor T\Delta^{-1} \rfloor}\right).$$

Convergence rate

Theorem (Part I)

Work under suitable assumptions on the wavelet basis. Let $s > 1/\pi$, $p \geq 1 \wedge \pi$ and take J and η such that for sufficiently large $\kappa > 0$

$$\eta = \kappa N_T^{-1/2} \sqrt{\log\left(N_T^{1/2}\right)} \quad \textit{and} \quad 2^J N_T^{-1} \log\left(N_T^{1/2}\right) \leq 1.$$

Let

$$\alpha = \min\left\{ \frac{s}{2s+1}, \frac{s+1/p-1/\pi}{2\left(s+1/2-1/\pi\right)} \right\}.$$

Convergence rate

Theorem (Part II)

1. *The estimator* $\widehat{P_{\Delta_T,m}}$ *verifies for* T *large enough*

$$\sup_{\mathbf{P}_{\Delta_T}[f]^{\star m}\in\mathscr{F}(s,\pi,\mathfrak{M})} \mathbb{E}\left[\left\|\widehat{P_{\Delta_T,m}} - \mathbf{P}_{\Delta_T}[f]^{\star m}\right\|^p_{L_p(\mathscr{D})} \middle| N_T\right] \leq \mathfrak{C} N_T^{-\alpha p},$$

up to logarithmic factors.

2. *The estimator corrected at order* K *verifies for* T *large enough for and any positive constants* $\underline{\mathfrak{T}}$ *and* $\overline{\mathfrak{T}}$

$$\sup_{\vartheta\in[\underline{\mathfrak{T}},\overline{\mathfrak{T}}]} \sup_{f\in\mathscr{F}(s,\pi,\mathfrak{M})} \mathbb{E}\left[\left\|\widetilde{f}^K_{T,\Delta_T} - f\right\|^p_{L_p(\mathscr{D})}\right] \leq \mathfrak{C}\max\left\{T^{-\alpha p}, \Delta_T^{p(K+1)}\right\},$$

up to logarithmic factors.

Achievable rates

Trade-off

- Deterministic approximation of f.
- Estimation of the $\left(\mathbf{P}_\Delta[f]^{\star m}, m \in \{1, \dots, K+1\}\right)$.

Achievable rates

Since $\alpha \leq 1/2$ to attain the minimax rate $T^{-\alpha}$ it is sufficient to choose K such that

$$\boxed{T\Delta_T^{2K+2} \leq 1}.$$

Always possible if T and Δ are polynomially related.

Plan

Parameter choice

- Estimate a mixture between a Gaussian and a Laplace.
- Choose $T = 10000$ and $\Delta = 0.1$ (such that $T\Delta^4 = 1$ and $\widehat{p} = 0.9508$ where $p = \mathbb{P}(R_\Delta = 1 | R_\Delta \neq 0)$).
- Compare the result with an oracle : the estimator computed on the actual jumps.

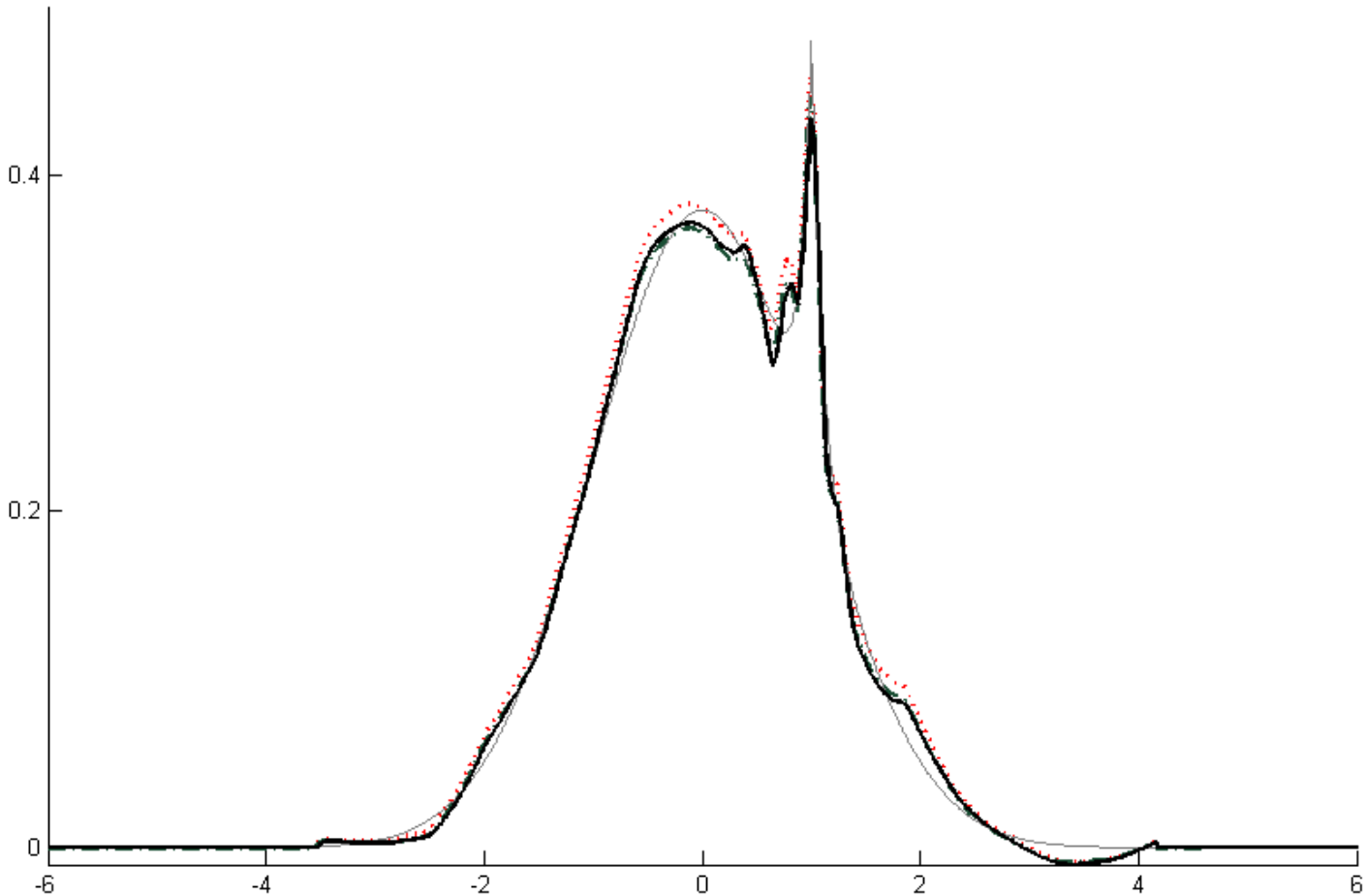

FIGURE: Estimators of the density f (plain grey) : the uncorrected (dotted red), the 1-corrected (dashed green) and the oracle (plain dark).

Comparison of the estimators for the L_2 loss

Estimator	Oracle	$K = 0$	$K = 1$	$K = 2$	$K = 3$
L_2 loss ($\times 10^{-4}$)	0.1117	0.1842	0.1353	0.1350	0.1350
Std devd ($\times 10^{-5}$)	(0.3495)	(0.4434)	(0.4363)	(0.4366)	(0.4366)

Discussion

- Those results partly generalise results of Comte and Genon-Catalot (09) and Figueroa-López (09) in the Poisson case.
- It is possible to extend the results (and the construction) to renewal reward processes. Computation of the inverse is no longer explicit.

References (selected)


Buchmann, B. and Grübel, R. (2003). Decompounding : an estimation problem for Poisson random sums. *Annals of Statistics* **31**, 1054–1074.


Comte, F. and Genon-Catalot, V. (2009). Nonparametric estimation for pure jump Lévy processes based on high frequency data. *SPA* **119**, 4088–4123.


Figueroa-López J.E. (2009). Nonparametric Estimation for Lévy Models Based on Discrete-Sampling. *IMS Lecture Notes-Monograph Series.* Vol. 57, 96–116.


Härdle, W., Kerkyacharian, G., Picard, D. and Tsybakov, A. (1998). *Wavelets, Approximation, and Statistical Applications.* Lecture Notes in Statistics, 129. Springer.


Kerkyacharian, G. and Picard, D. (2000). Thresholding algorithms, maxistes and well-concentrated bases. *Test*, Vol. 9, No. 2, 283–344.