Statistics versus mean-field limit for Hawkes process

with Sylvain Delattre (P7)

The model

We have N individuals.

$Z_t^{i,N} :=$ number of actions of the i-th individual until time t.

$Z_t^{i,N}$ jumps (is increased by 1) at rate

$$\lambda_t^{i,N} = \mu + \frac{1}{N} \sum_{j=1}^{N} \theta_{ij} \int_0^t \varphi(t-s) dZ_s^{j,N}, \quad 1 \leq i \leq N$$

where

$$\mu \in (0, \infty), \quad \varphi : [0, \infty) \to [0, \infty),$$

$$(\theta_{ij})_{i,j=1,\dots,N} \quad \text{i.d.d. Bernoulli}(p)$$

Two types of actions: *autonomous* and *by mimetism*.

Example : $\varphi = \mathbf{1}_{[0,K]}$.

Goal

We observe the activity of the N individuals until time t that is

$$Z_s^{i,N}, \quad i = 1, \dots, N, \quad s \in [0, t]$$

and we want to estimate p (i.e. the main characteristic of the interaction graph), in the asymptotic

$$N \to \infty \quad \text{and} \quad t \to \infty$$

We consider μ and φ as nuisance parameters.

Intuitively: not very easy... how to know if a jump is autonomous or is due to excitation by another individual (and which one) ?

Hawkes 1971, Hawkes-Oakes 1974.
Finance, Neurons, Earthquake replicas, etc.
Estimation of (more general) μ and (nonparametric) φ_{ij} at N fixed as $t \to \infty$: Hansen, Reynaud, Rivoirard, Gaiffas, Hoffmann, Bacry, Muzzy, Rasmussen, etc.

Mean-field limit ($N \to \infty$, t fixed)

For each given $k \geq 1$ and $t > 0$, the process

$$Z_s^{i,N}, \quad i = 1, \dots, k, \quad s \in [0, t]$$

goes in law as $N \to \infty$ to

$$Y_s^i, \quad i = 1, \dots, k, \quad s \in [0, t]$$

a family of i.i.d. inhomogeneous Poisson processes with intensity $(\lambda_s)_{s \in [0,t]}$ satisfying

$$\lambda_s = \mu + p \int_0^s \varphi(s - u)\lambda_u du, \quad s \in [0, t].$$

The limit depends only on μ and $p\varphi$ thus it is not identifiable.

Main result

Set $\Lambda := \int_0^\infty \varphi(t)dt$.

- Subcritical case: $\Lambda p < 1$. Then roughly, $\bar{Z}_t^N \simeq t$ (on an event where (θ_{ij}) behaves reasonably). We put $m_t = t$
- Supercritical case : $\Lambda p > 1$. Then roughly, $\bar{Z}_t^N \simeq e^{\alpha_0 t}$, with α_0 defined by $p \int_0^\infty e^{-\alpha_0 t}\varphi(t)dt = 1$. We put $m_t = e^{\alpha_0 t}$.
- Critical case: zoology, we do not treat.

Theorem

Under some (reasonnable) technical assumptions on φ and if $\Lambda p \neq 1$, there exists an (explicit) estimator $\hat{p}_t^N$ such that

$$\Pr\Big(|\hat{p}_t^N - p| \geq \varepsilon\Big) \leq \frac{C}{\varepsilon}\Big(\frac{1}{\sqrt{N}} + \frac{\sqrt{N}}{m_t}\Big)^{1-}$$

where C depends only on p, μ, φ.

The precision $\frac{1}{\sqrt{N}} + \frac{\sqrt{N}}{m_t}$ seems to be optimal (Gaussian toy model).
We need t large if N large because the MF limit is not identifiable.

Subcritical case 1

Very pleasant point: we will not have to estimate the non-parametric nuisance parameter φ (this would of course lead to a much less precise estimation).

We will build 3 estimators

$$\mathcal{E}_t^N \simeq \frac{\mu}{1-\Lambda p}, \quad \mathcal{V}_t^N \simeq \frac{\mu^2\Lambda^2}{(1-\Lambda p)^2} p(1-p), \quad \mathcal{W}_t^N \simeq \frac{\mu}{(1-\Lambda p)^3}.$$

One then easily find Φ such that $\hat{p}_t^N = \Phi(\mathcal{E}_t^N, \mathcal{V}_t^N, \mathcal{W}_t^N) \simeq p$ (as well as estimators of Λ and μ).

Subcritical case 2

It is easily seen that

$$\mathsf{E}_\theta[Z_t^{i,N}] = \mu t + \frac{1}{N}\sum_{j=1}^{N}\theta_{ij}\int_0^t \varphi(t-s)\mathsf{E}_\theta[Z_s^{j,N}]ds,$$

Assuming that $Z_t^{i,N} \approx \mathsf{E}_\theta[Z_t^{i,N}] \simeq \gamma_N(i)t$ for large t (N fixed).

$$\gamma_N(i) \simeq \mu + \frac{1}{N}\sum_{j=1}^{N}\theta_{ij}\Lambda\gamma_N(j).$$

Thus, with $A_N = (\frac{1}{N}\theta_{ij})_{1\le i,j\le N}$ and $Q_N = (I - \Lambda A_N)^{-1}$.

$$\gamma_N(i) = \mu\sum_{j=1}^{N}Q_N(i,j) = \mu\sum_{j=1}^{N}\sum_{k\ge 0}\Lambda^k A_N^k(i,j).$$

Conclusion: $Z_t^N(i) \simeq \mu\ell_N(i)t$ with $\ell_N(i) = \sum_{j=1}^N Q_N(i,j)$.

Remark: Q_N exists with high probability in the subcritical case.

Subcritical case 3

We study now ℓ_N. True (and quantified) that

$$\ell_N(i) = \sum_{\ell \geq 0} \Lambda^\ell \sum_{j=1}^N A_N^\ell(i,j) \approx \left(1 + \frac{\Lambda}{1-\Lambda p} \frac{1}{N} \sum_{j=1}^N \theta_{ij}\right)$$

since for $\ell \geq 1$

$$\sum_{j=1}^N A_N^\ell(i,j) = \frac{1}{N^\ell} \sum_j \sum_{i_1,\dots,i_{\ell-1}} \theta_{i,i_1}\theta_{i_1,i_2}\dots\theta_{i_{\ell-1},j}$$

$$= \frac{1}{N^\ell} \sum_{i_1} \theta_{i,i_1} \underbrace{\sum_{i_2,\dots,i_{\ell-1},j} \theta_{i_1,i_2}\dots\theta_{i_{\ell-1},j}}_{\approx N^{\ell-1}p^{\ell-1} \text{ if } \ell \ll N} \approx p^{\ell-1} \frac{1}{N} \sum_j \theta_{ij}$$

Not related to known eigenvalues problems, no way to use moments, because all expectations are infinite, because A_N does not exist on a small event.

Subcritical case 4

Thus $Z_t^N(i) \simeq \mu(1 + \frac{\Lambda}{1-\Lambda p} L_N(i))t$ with $L_N(i) = \frac{1}{N}\sum_{j=1}^N \theta_{ij}$.

First estimator:

$$\boxed{\mathcal{E}_t^N = \frac{\bar{Z}_t^N}{t}} \simeq \mu(1 + \frac{\Lambda}{1-\Lambda p} p) = \frac{\mu}{1-\Lambda p}.$$

Subcritical case 5

Thus $Z_t^N(i) \simeq \mu(1 + \frac{\Lambda}{1-\Lambda p} L_N(i))t$ with $L_N(i) = \frac{1}{N}\sum_{j=1}^N \theta_{ij}$.

Second estimator:

$$\boxed{\mathcal{V}_t^N := \sum_{i=1}^N \Big(\frac{1}{t} Z_t^{i,N} - \frac{1}{t}\bar{Z}_t^N\Big)^2 - \frac{N}{t}\bar{Z}_t^N}$$

$$\frac{1}{N}\sum_{i=1}^N \Big(\frac{1}{t} Z_t^{i,N} - \frac{1}{t}\bar{Z}_t^N\Big)^2 \approx \mathbb{V}\mathrm{ar}\Big(\frac{1}{t} Z_t^{1,N}\Big)$$

$$= \mathbb{V}\mathrm{ar}\big(\mathsf{E}_\theta[\frac{1}{t} Z_t^{1,N}]\big) + \frac{1}{t^2}\mathsf{E}\big[\mathbb{V}\mathrm{ar}_\theta(Z_t^{1,N})\big]$$

and $\mathbb{V}\mathrm{ar}_\theta(Z_t^{1,N}) \approx \mathsf{E}_\theta[Z_t^{1,N}]$ since $Z_t^{1,N} \approx$ Poisson. Thus

$$\mathcal{V}_t^N \approx \frac{N}{t^2}\mathbb{V}\mathrm{ar}(\mathsf{E}_\theta[Z_t^{1,N}]) \approx \frac{\mu^2\Lambda^2}{(1-\Lambda p)^2} p(1-p)$$

Subcritical case 6

Third estimator: temporal empirical variance

$$\mathcal{W}^N_{\Delta,t} = \frac{N}{t}\sum_{k=1}^{t/\Delta}\Big(\bar{Z}^N_{k\Delta} - \bar{Z}^N_{(k-1)\Delta} - \frac{\Delta}{t}\bar{Z}^N_t\Big)^2$$

where $1 \ll \Delta \ll t$ (theory: $\Delta = t^{0.0001}$)

$$\mathcal{W}^N_{\Delta,t} \approx N\frac{1}{\Delta}\mathbb{V}\mathrm{ar}_\theta(\bar{Z}^N_\Delta) \approx \frac{\mu}{(1-\Lambda p)^3}$$

More complicated. Really uses that $\bar{Z}^N$ is not Poisson. Actually, $N\bar{Z}^N$ resembles, very roughly, an autonomous (1D) Hawkes process with parameters $N\mu$ and $p\varphi$.

Subcritical case 7

Remark: everything starts from $\Gamma(t) := \int_0^t s\varphi(t-s)ds \simeq \Lambda t$.

With e.g. $\varphi = e^{-t}$, we have $\int_0^t s\varphi(t-s)ds = \Lambda t - 1 + e^{-t}$. So $\Gamma(2t) - \Gamma(t)$ resembles Λt considerably much more precisely than $\Gamma(t)$ (always true when φ has a fast decay).

We thus modify the 3 estimators. For example, we use

$$\mathcal{E}_t^N = \frac{\bar{Z}_{2t}^N - \bar{Z}_t^N}{t} \quad \text{instead of} \quad \mathcal{E}_t^N = \frac{\bar{Z}_t^N}{t}$$

This is crucial to get the *nearly optimal (??)* precision.

Supercritical case 1

We expect that $Z_t^{i,N} \simeq H_N \mathsf{E}_\theta[Z_t^{i,N}]$ for some r.v. $H_N > 0$;

and that $\mathsf{E}_\theta[Z_t^{i,N}] \simeq \gamma_N(i) e^{\alpha_N t}$. But

$$\mathsf{E}_\theta[Z_t^{i,N}] = \mu t + \frac{1}{N} \sum_{j=1}^{N} \theta_{ij} \int_0^t \varphi(t-s) \mathsf{E}_\theta[Z_s^{j,N}] ds$$

So, with $A_N(i,j) = \frac{1}{N}\theta_{ij}$,

$$\gamma_N = A_N \gamma_N \int_0^\infty e^{-\alpha_N s} \varphi(s) ds.$$

The vector γ_N being positive, it is a Perron-Frobenius eigenvector of A_N, so that $\rho_N = (\int_0^\infty e^{-\alpha_N s}\varphi(s)ds)^{-1}$ is its Perron-Frobenius eigenvalue. Since $A_N(i,j) \simeq p$, we conclude that $\rho_N \simeq p$ and thus $\alpha_N \simeq \alpha_0$.

We consider the Perron-Frobenius eigenvector V_N such that $\sum_{i=1}^N (V_N(i))^2 = N$ and conclude that (with another r.v. $K_N > 0$)

$$Z_t^{i,N} \simeq K_N V_N(i) e^{\alpha_0 t}$$

Supercritical case 2

Thus $Z_t^{i,N} \simeq K_N V_N(i) e^{\alpha_0 t}$. Single estimator:

$$\boxed{\mathcal{U}_t^N = \frac{1}{(\bar{Z}_t^N)^2}\Big[\sum_1^N (Z_t^{i,N} - \bar{Z}_t^N)^2 - N\bar{Z}_t^N\Big]} \simeq \frac{1}{(\bar{V}_N)^2}\sum_1^N (V_N(i) - \bar{V}_N)^2$$

Indeed,

$$\begin{aligned}\frac{1}{N}\sum_{i=1}^N \Big(Z_t^{i,N} - \bar{Z}_t^N\Big)^2 &\approx \mathbb{V}\mathrm{ar}\Big(Z_t^{1,N}\Big) \\ &= \mathbb{V}\mathrm{ar}\big(\mathsf{E}_\theta[Z_t^{1,N}]\big) + \mathsf{E}\big[\mathbb{V}\mathrm{ar}_\theta(Z_t^{1,N})\big]\end{aligned}$$

and $\mathbb{V}\mathrm{ar}_\theta(Z_t^{1,N}) \approx \mathsf{E}_\theta[Z_t^{1,N}]$ since $Z_t^{1,N} \approx$ Poisson. Thus

$$\mathcal{U}_t^N \approx \frac{N}{(\bar{Z}_t^N)^2}\mathbb{V}\mathrm{ar}(\mathsf{E}_\theta[Z_t^{1,N}]) \approx \frac{1}{(\bar{V}_N)^2}\sum_1^N (V_N(i) - \bar{V}_N)^2$$

Supercritical case 3

$\mathcal{U}_t^N \approx \frac{1}{(\bar{V}_N)^2} \sum_1^N (V_N(i) - \bar{V}_N)^2.$

But V_N is almost colinear to L_N (with $L_N(i) = \sum_1^N A_N(i,j)$):

very roughly, the matrix A_N^2 is almost constant ($A_N^2(i,j) \simeq p$), so that its Perron Frobenius eigenvector (V_N) is almost constant ($V_N(i) \simeq 1$), so that $A_N V_N = \rho_N V_N$ gives $V_N(i) \simeq \rho_N^{-1} \sum_j A_N(i,j) = \rho_N^{-1} L_N(i)$ (not very convincing).

Since L_N is a vector of i.i.d. Binomial(N, p), we conclude that

$$\mathcal{U}_t^N \approx \frac{1}{(\bar{L}_N)^2} \sum_1^N (L_N(i) - \bar{L}_N)^2 \simeq \frac{p(1-p)}{p^2} = \frac{1}{p} - 1.$$

Choice between sub and super

We thus have two different estimators $p_1(N,t) \simeq p$ (if $\Lambda p < 1$) and $p_2(N,t) \simeq p$ (if $\Lambda p > 1$). If we do not know, we set

$$p(N,t) = p_1(N,t)\mathbf{1}_{\{\bar{Z}_t^N \leq \exp((\log t)^2)\}} + p_2(N,t)\mathbf{1}_{\{\bar{Z}_t^N > \exp((\log t)^2)\}}$$

(does not affect the precision).

Optimality ? A Gaussian toy model

$\Gamma > 0$ and $p \in (0, 1]$ unknown, $(\theta_{ij})_{i,j=1,\dots,N}$ i.i.d. Ber(p) and the observations are

$$Z_t^{i,N} \sim \text{Poisson}\Big(\Gamma t \sum_{j=1}^{N} \theta_{ij}\Big), \quad i = 1, \dots, N$$

Then roughly,

$$X_t^{i,N} = \frac{1}{t} Z_t^{i,N} \sim \text{Normal}\Big(\Gamma p, \frac{\Gamma^2 p(1-p)}{N} + \frac{1}{t}\Gamma p\Big).$$

Assume that Γp is known (this can only increase the precision).

$S_t^N = N^{-1} \sum_1^N (X_t^{i,N} - \Gamma p)^2$ is the *B.E.* of $N^{-1}\Gamma^2 p(1-p) + t^{-1}\Gamma p$.

Thus $T_t^N = N(\Gamma p)^{-2}(S_t^N - t^{-1}\Gamma p)$ is the *B.E.* of $(1/p - 1)$.

And $\mathbb{V}\mathrm{ar}\, T_t^N \simeq (N^{-1/2} + t^{-1}N^{1/2})^2$.

Thus optimal precision in $N^{-1/2} + t^{-1}N^{1/2}$.

Mersi pour votre attansion.