Statistical estimation of the division rate of a size-structured population

M. Doumic, M. Hoffmann, P. Reynaud-Bouret, V. Rivoirard

INRIA Rocquencourt, ENSAE-CREST, Nice, Dauphine

10th ICOR, Habana, March 6-9, 2012

1 The problem

1. The problem
2. Goldenshluger and Lepski's method

The informal problem and the PDE translation for size-structured population

- A cell grows.
- Depending on its size x, the cell has a certain chance to divide itself in 2 offsprings, ie 2 cells of size $x/2$.
- We are interesting by the evolution of the whole population of cells, each of them having this behavior.

The informal problem and the PDE translation for size-structured population

- A cell grows.
- Depending on its size x, the cell has a certain chance to divide itself in 2 offsprings, ie 2 cells of size $x/2$.
- We are interesting by the evolution of the whole population of cells, each of them having this behavior.

Size-Structured Population Equation (finite time)

$$\begin{cases} \dfrac{\partial}{\partial t}(n(t,x)) + \kappa \dfrac{\partial}{\partial x}(g(x)n(t,x)) + B(x)n(t,x) = 4B(2x)n(t,2x), \\ n(t,x=0) = 0, \quad t > 0 \\ n(0,x) = n_0(x), \quad x \geq 0. \end{cases}$$

- $n(t,x)$ the "amount" of cells with size x ($\neq$ density),
- g the "qualitative" growth rate of one cell: linear is $g = 1$ ...
- B is the division rate, which depends on the size

Asymptotics of the PDE

It can be shown (Perthame Ryzhik 2005 for instance) that

- $n(t,.)$ grows exponentially fast ie $I_t = \int n(t,x)dx$ asymptotically proportional to $e^{\lambda t}$,
- the renormalized $n(t,x)/I_t$ tends to a density N, which satisfies

Size-Structured Population Equation (asymptotics)

$$\begin{cases} \kappa \dfrac{\partial}{\partial x}\big(g(x)N(x)\big) + \lambda N(x) = \mathcal{L}\big(BN\big)(x), \\ B(0)N(0) = 0, \qquad \int N(x)dx = 1, \end{cases}$$

where N step D step κ step L step H step B step

- for any real-valued function $x \rightsquigarrow \varphi(x)$, $\mathcal{L}\big(\varphi\big)(x) := 4\varphi(2x) - \varphi(x)$.
- $\kappa = \lambda \dfrac{\int_{\mathbb{R}_+} xN(x)dx}{\int_{\mathbb{R}_+} g(x)N(x)dx}$.

The inverse problem

Under the previous differential equation, we consider the inverse problem of finding B given a "noisy" version of N.

The inverse problem

Under the previous differential equation, we consider the inverse problem of finding B given a "noisy" version of N.

- Practical: biologists take a sample of, say, plankton in a lake, and they look at the respective size of the cells.

The inverse problem

Under the previous differential equation, we consider the inverse problem of finding B given a "noisy" version of N.

- Practical: biologists take a sample of, say, plankton in a lake, and they look at the respective size of the cells. Then they perform a preprocessing, by, say a kernel estimator. This is N_ϵ.

The inverse problem

Under the previous differential equation, we consider the inverse problem of finding B given a "noisy" version of N.

- Practical: biologists take a sample of, say, plankton in a lake, and they look at the respective size of the cells. Then they perform a preprocessing, by, say a kernel estimator. This is N_ϵ. (probably more approximation than that).

The inverse problem

Under the previous differential equation, we consider the inverse problem of finding B given a "noisy" version of N.

- Practical: biologists take a sample of, say, plankton in a lake, and they look at the respective size of the cells. Then they perform a preprocessing, by, say a kernel estimator. This is N_ϵ. (probably more approximation than that).
- Analytical point of view: N_ϵ is a noisy version of N, less regular than N (it is likely that no derivative exists) and $\|N - N_\epsilon\|_2 \leq \epsilon$. (see Perthame, Zubelli, etc)

The inverse problem

Under the previous differential equation, we consider the inverse problem of finding B given a "noisy" version of N.

- Practical: biologists take a sample of, say, plankton in a lake, and they look at the respective size of the cells. Then they perform a preprocessing, by, say a kernel estimator. This is N_ϵ. (probably more approximation than that).
- Analytical point of view: N_ϵ is a noisy version of N, less regular than N (it is likely that no derivative exists) and $\|N - N_\epsilon\|_2 \leq \epsilon$. (see Perthame, Zubelli, etc)
- Statistical point of view: we observe a n-sample $X_1, ..., X_n$ of iid variables with density N.

Pro and Con

Analytical point of view

Pro: taking into account maybe more approximations (but not all), results true for any N_ϵ.

Pro and Con

Analytical point of view

Pro: taking into account maybe more approximations (but not all), results true for any N_ϵ.
Con: N_ϵ is probably differentiable. If there are numerical methods which adapt to the regularity of N (discrepancy principle), they need to know ϵ.

Pro and Con

Analytical point of view

Pro: taking into account maybe more approximations (but not all), results true for any N_ϵ.
Con: N_ϵ is probably differentiable. If there are numerical methods which adapt to the regularity of N (discrepancy principle), they need to know ϵ.

Statistical point of view

Pro: Framework close to what biologists do, true inverse problem. We can adapt to the regularity, noise is given by the sample size.

Pro and Con

Analytical point of view

Pro: taking into account maybe more approximations (but not all), results true for any N_ϵ.
Con: N_ϵ is probably differentiable. If there are numerical methods which adapt to the regularity of N (discrepancy principle), they need to know ϵ.

Statistical point of view

Pro: Framework close to what biologists do, true inverse problem. We can adapt to the regularity, noise is given by the sample size.
Con: We only take one approximation into account and assume that we have access to the sample. Results true in expectation.

Pro and Con

Analytical point of view

Pro: taking into account maybe more approximations (but not all), results true for any N_ϵ.
Con: N_ϵ is probably differentiable. If there are numerical methods which adapt to the regularity of N (discrepancy principle), they need to know ϵ.

Statistical point of view

Pro: Framework close to what biologists do, true inverse problem. We can adapt to the regularity, noise is given by the sample size.
Con: We only take one approximation into account and assume that we have access to the sample. Results true in expectation.

Assumptions

1. For the considered nonnegative functions g and B and for $\kappa > 0$, there exists a unique solution (λ, N) of SSPS
2. This solution satisfies, for all $p \geq 0$, $\int x^p N(x)dx < \infty$ and $0 < \int g(x)N(x)dx < \infty$.
3. The functions N and gN belong to $\mathcal{W}^{s+1}$ with $s \geq 1$
   $\mathcal{W}^{s+1}$ denotes the Sobolev space of regularity $s+1$ measured in $\mathbb{L}^2$-norm.
4. We have $g \in \mathbb{L}^\infty(\mathbb{R}_+)$ with $\mathbb{R}_+ = [0, \infty)$.

Assumptions

1. For the considered nonnegative functions g and B and for $\kappa > 0$, there exists a unique solution (λ, N) of SSPS
2. This solution satisfies, for all $p \geq 0$, $\int x^p N(x)dx < \infty$ and $0 < \int g(x)N(x)dx < \infty$.
3. The functions N and gN belong to $\mathcal{W}^{s+1}$ with $s \geq 1$ $\mathcal{W}^{s+1}$ denotes the Sobolev space of regularity $s+1$ measured in $\mathbb{L}^2$-norm.
4. We have $g \in \mathbb{L}^\infty(\mathbb{R}_+)$ with $\mathbb{R}_+ = [0, \infty)$.

The statistical methodology is based on kernel rules. Classical assumptions on kernels are made (not specified in the sequel).

Estimation of N

Given K a kernel, we set $K_h(x) = \frac{1}{h}K\left(\frac{x}{h}\right)$ and

$$\hat{N}_h(x) := \frac{1}{n}\sum_{i=1}^{n} K_h(x - X_i)$$

Bias-Variance decomposition

$$\mathbb{E}\left[\left\|N - \hat{N}_h\right\|_2\right] \leq \|N - K_h \star N\|_2 + \frac{1}{\sqrt{nh}}\|K\|_2,$$

where $K_h \star N = \mathbb{E}(\hat{N}_h)$

For $\mathcal{H}$ a family of bandwidths, the "best choice" is the oracle:

$$\bar{h} := \mathrm{argmin}_{h\in\mathcal{H}}\left\{\|N - K_h \star N\|_2 + \frac{1}{\sqrt{nh}}\|K\|_2\right\}.$$

How to select the bandwidth h only based on data?

Estimation of N

Given K a kernel, we set $K_h(x) = \frac{1}{h}K\left(\frac{x}{h}\right)$ and

$$\hat{N}_h(x) := \frac{1}{n}\sum_{i=1}^{n} K_h(x - X_i)$$

Bias-Variance decomposition

$$\mathbb{E}\left[\left\|N - \hat{N}_h\right\|_2\right] \leq \|N - K_h \star N\|_2 + \frac{1}{\sqrt{nh}}\|K\|_2,$$

where $K_h \star N = \mathbb{E}(\hat{N}_h)$

For $\mathcal{H}$ a family of bandwidths, the "best choice" is the oracle:

$$\bar{h} := \mathrm{argmin}_{h\in\mathcal{H}}\left\{\|N - K_h \star N\|_2 + \frac{1}{\sqrt{nh}}\|K\|_2\right\}.$$

How to select the bandwidth h only based on data? Recent work of Goldenshluger and Lepski (2009, 2010)!!!

Estimation of N

Given K a kernel, we set $K_h(x) = \frac{1}{h}K\left(\frac{x}{h}\right)$ and

$$\hat{N}_h(x) := \frac{1}{n}\sum_{i=1}^{n} K_h(x - X_i)$$

Bias-Variance decomposition

$$\mathbb{E}\left[\left\|N - \hat{N}_h\right\|_2\right] \leq \|N - K_h \star N\|_2 + \frac{1}{\sqrt{nh}}\|K\|_2,$$

where $K_h \star N = \mathbb{E}(\hat{N}_h)$

For $\mathcal{H}$ a family of bandwidths, the "best choice" is the oracle:

$$\bar{h} := \mathrm{argmin}_{h\in\mathcal{H}}\left\{\|N - K_h \star N\|_2 + \frac{1}{\sqrt{nh}}\|K\|_2\right\}.$$

How to select the bandwidth h only based on data? Recent work of Goldenshluger and Lepski (2009, 2010)!!! Here just a "toy" version, but that's exactly what we needed.

Bandwidth selection by the GL method

SSPE Set for any x and any $h, h' > 0$,

$$\hat{N}_{h,h'}(x) := (K_h \star \hat{N}_{h'})(x) = \frac{1}{n}\sum_{i=1}^{n}(K_h \star K_{h'})(x - X_i),$$

Bandwidth selection by the GL method

SSPE Set for any x and any $h, h' > 0$,

$$\hat{N}_{h,h'}(x) := (K_h \star \hat{N}_{h'})(x) = \frac{1}{n}\sum_{i=1}^{n}(K_h \star K_{h'})(x - X_i),$$

"Estimator" of the bias term

$$A(h) := \sup_{h' \in \mathcal{H}} \left\{ \|\hat{N}_{h,h'} - \hat{N}_{h'}\|_2 - \frac{\chi}{\sqrt{nh'}}\|K\|_2 \right\}_+$$

where, given $\varepsilon > 0$, $\chi := (1+\varepsilon)(1+\|K\|_1)$.

Bandwidth selection by the GL method

SSPE Set for any x and any $h, h' > 0$,

$$\hat{N}_{h,h'}(x) := (K_h \star \hat{N}_{h'})(x) = \frac{1}{n}\sum_{i=1}^{n}(K_h \star K_{h'})(x - X_i),$$

"Estimator" of the bias term

$$A(h) := \sup_{h' \in \mathcal{H}} \left\{ \|\hat{N}_{h,h'} - \hat{N}_{h'}\|_2 - \frac{\chi}{\sqrt{nh'}}\|K\|_2 \right\}_+$$

where, given $\varepsilon > 0$, $\chi := (1+\varepsilon)(1+\|K\|_1)$.

$$\hat{h} := \arg\min_{h \in \mathcal{H}} \left\{ A(h) + \frac{\chi}{\sqrt{nh}}\|K\|_2 \right\} \quad \text{and} \quad \hat{N} := \hat{N}_{\hat{h}}.$$

Some explanations for this choice ...

$$
\begin{aligned}
A(h) &= \sup_{h' \in \mathcal{H}} \left\{ \|\hat{N}_{h,h'} - \hat{N}_{h'}\|_2 - \frac{\chi}{\sqrt{nh'}} \|K\|_2 \right\}_+ \\
&\leq \sup_{h' \in \mathcal{H}} \left\{ \|\mathbb{E}(\hat{N}_{h,h'}) - \mathbb{E}(\hat{N}_{h'})\|_2 \right\} + \zeta
\end{aligned}
$$

where

$$
\zeta = \sup_{h' \in \mathcal{H}} \left\{ \|\hat{N}_{h,h'} - \mathbb{E}(\hat{N}_{h,h'}) - \left(\hat{N}_{h'} - \mathbb{E}(\hat{N}_{h'})\right)\|_2 - \frac{\chi}{\sqrt{nh'}} \|K\|_2 \right\}_+ .
$$

Some explanations for this choice ...

$$\begin{aligned} A(h) &= \sup_{h' \in \mathcal{H}} \left\{ \|\hat{N}_{h,h'} - \hat{N}_{h'}\|_2 - \frac{\chi}{\sqrt{nh'}} \|K\|_2 \right\}_+ \\ &\leq \sup_{h' \in \mathcal{H}} \left\{ \|\mathbb{E}(\hat{N}_{h,h'}) - \mathbb{E}(\hat{N}_{h'})\|_2 \right\} + \zeta \end{aligned}$$

where

$$\zeta = \sup_{h' \in \mathcal{H}} \left\{ \|\hat{N}_{h,h'} - \mathbb{E}(\hat{N}_{h,h'}) - \left(\hat{N}_{h'} - \mathbb{E}(\hat{N}_{h'})\right)\|_2 - \frac{\chi}{\sqrt{nh'}} \|K\|_2 \right\}_+ .$$

$$\mathbb{E}\left(\hat{N}_{h,h'}(x)\right) - \mathbb{E}\left(\hat{N}_{h'}(x)\right) = \int (K_h \star K_{h'})(x-u) N(u) du - \int K_{h'}(x-v) N(v) dv$$

Some explanations for this choice ...

$$A(h) = \sup_{h' \in \mathcal{H}} \left\{ \|\hat{N}_{h,h'} - \hat{N}_{h'}\|_2 - \frac{\chi}{\sqrt{nh'}} \|K\|_2 \right\}_+ \leq \sup_{h' \in \mathcal{H}} \left\{ \|\mathbb{E}(\hat{N}_{h,h'}) - \mathbb{E}(\hat{N}_{h'})\|_2 \right\} + \zeta$$

where

$$\zeta = \sup_{h' \in \mathcal{H}} \left\{ \|\hat{N}_{h,h'} - \mathbb{E}(\hat{N}_{h,h'}) - (\hat{N}_{h'} - \mathbb{E}(\hat{N}_{h'}))\|_2 - \frac{\chi}{\sqrt{nh'}} \|K\|_2 \right\}_+ .$$

$$\mathbb{E}(\hat{N}_{h,h'}(x)) - \mathbb{E}(\hat{N}_{h'}(x)) = \int (K_h \star K_{h'})(x-u)N(u)du - \int K_{h'}(x-v)N(v)dv$$

$$= \int \int K_h(x-u-t)K_{h'}(t)N(u)dtdu - \int K_{h'}(x-v)N(v)dv$$

Some explanations for this choice ...

$$\begin{aligned} A(h) &= \sup_{h' \in \mathcal{H}} \left\{ \|\hat{N}_{h,h'} - \hat{N}_{h'}\|_2 - \frac{\chi}{\sqrt{nh'}} \|K\|_2 \right\}_+ \\ &\leq \sup_{h' \in \mathcal{H}} \left\{ \|\mathbb{E}(\hat{N}_{h,h'}) - \mathbb{E}(\hat{N}_{h'})\|_2 \right\} + \zeta \end{aligned}$$

where

$$\zeta = \sup_{h' \in \mathcal{H}} \left\{ \|\hat{N}_{h,h'} - \mathbb{E}(\hat{N}_{h,h'}) - \left(\hat{N}_{h'} - \mathbb{E}(\hat{N}_{h'})\right)\|_2 - \frac{\chi}{\sqrt{nh'}} \|K\|_2 \right\}_+ .$$

$$\begin{aligned} \mathbb{E}(\hat{N}_{h,h'}(x)) - \mathbb{E}(\hat{N}_{h'}(x)) &= \int (K_h \star K_{h'})(x-u) N(u) du - \int K_{h'}(x-v) N(v) dv \\ &= \int \int K_h(x-u-t) K_{h'}(t) N(u) dt du - \int K_{h'}(x-v) N(v) dv \\ &= \int \int K_h(v-u) K_{h'}(x-v) N(u) du dv - \int K_{h'}(x-v) N(v) dv \end{aligned}$$

Some explanations for this choice ...

$$
\begin{aligned}
A(h) &= \sup_{h' \in \mathcal{H}} \left\{ \|\hat{N}_{h,h'} - \hat{N}_{h'}\|_2 - \frac{\chi}{\sqrt{nh'}} \|K\|_2 \right\}_+ \\
&\leq \sup_{h' \in \mathcal{H}} \left\{ \|\mathbb{E}(\hat{N}_{h,h'}) - \mathbb{E}(\hat{N}_{h'})\|_2 \right\} + \zeta
\end{aligned}
$$

where

$$
\zeta = \sup_{h' \in \mathcal{H}} \left\{ \|\hat{N}_{h,h'} - \mathbb{E}(\hat{N}_{h,h'}) - \left(\hat{N}_{h'} - \mathbb{E}(\hat{N}_{h'})\right)\|_2 - \frac{\chi}{\sqrt{nh'}} \|K\|_2 \right\}_+ .
$$

$$
\mathbb{E}(\hat{N}_{h,h'}(x)) - \mathbb{E}(\hat{N}_{h'}(x)) = \int (K_h \star K_{h'})(x-u) N(u) du - \int K_{h'}(x-v) N(v) dv
$$

$$
= \int \int K_h(x-u-t) K_{h'}(t) N(u) dt du - \int K_{h'}(x-v) N(v) dv
$$

$$
= \int \int K_h(v-u) K_{h'}(x-v) N(u) du dv - \int K_{h'}(x-v) N(v) dv
$$

$$
= \int K_{h'}(x-v) \left(\int K_h(v-u) N(u) du - N(v) \right) dv
$$

Some explanations for this choice ...

$$
\begin{aligned}
A(h) &= \sup_{h' \in \mathcal{H}} \left\{ \|\hat{N}_{h,h'} - \hat{N}_{h'}\|_2 - \frac{\chi}{\sqrt{nh'}} \|K\|_2 \right\}_+ \\
&\leq \sup_{h' \in \mathcal{H}} \left\{ \|\mathbb{E}(\hat{N}_{h,h'}) - \mathbb{E}(\hat{N}_{h'})\|_2 \right\} + \zeta
\end{aligned}
$$

where

$$
\zeta = \sup_{h' \in \mathcal{H}} \left\{ \|\hat{N}_{h,h'} - \mathbb{E}(\hat{N}_{h,h'}) - (\hat{N}_{h'} - \mathbb{E}(\hat{N}_{h'}))\|_2 - \frac{\chi}{\sqrt{nh'}} \|K\|_2 \right\}_+ .
$$

- $\|\mathbb{E}(\hat{N}_{h,h'}) - \mathbb{E}(\hat{N}_{h'})\|_2 \leq \|K\|_1 \|K_h \star N - N\|_2$
- ζ is a residual controlled by "Uniform bounds". It is small (n^{-1}) if χ is large enough.

First result

Oracle inequality

If $\mathcal{H} = \{1/\ell \;/\; \ell = 1, ..., \ell_{max}\}$ and if $\ell_{max} = \delta n$, if moreover $\|N\|_\infty < \infty$,
then for any $q \geq 1$,

$$\mathbb{E}\left(\|\hat{N} - N\|_2^{2q}\right) \leq \square_q \chi^{2q} \inf_{h \in \mathcal{H}} \left\{ \|K_h \star N - N\|_2^{2q} + \frac{\|K\|_2^{2q}}{(hn)^q} \right\} + \square_{q,\varepsilon,\delta,\|K\|_2,\|K\|_1,\|N\|_\infty} \frac{1}{n^q}.$$

Estimation of $D = \frac{\partial}{\partial x}\big(g(x)N(x)\big)$

SSPE

Estimation of $D = \frac{\partial}{\partial x}\big(g(x)N(x)\big)$

SSPE If K is differentiable, $\int K = 1$ and $\int |K'|^2 < \infty$.

$$\hat{D}_h(x) := \frac{1}{n}\sum_{i=1}^{n} g(X_i)K_h'(x - X_i)$$

Estimation of $D = \frac{\partial}{\partial x}(g(x)N(x))$

SSPE If K is differentiable, $\int K = 1$ and $\int |K'|^2 < \infty$.

$$\hat{D}_h(x) := \frac{1}{n}\sum_{i=1}^{n} g(X_i)K'_h(x - X_i)$$

Bias-Variance decomposition:

$$\mathbb{E}(\left\| D - \hat{D}_h \right\|_2) \leq \|D - K_h \star D\|_2 + \frac{1}{\sqrt{nh^3}}\|g\|_\infty \|K'\|_2.$$

Estimation of $D = \frac{\partial}{\partial x}(g(x)N(x))$

SSPE If K is differentiable, $\int K = 1$ and $\int |K'|^2 < \infty$.

$$\hat{D}_h(x) := \frac{1}{n}\sum_{i=1}^{n} g(X_i)K'_h(x - X_i)$$

GL's trick

$\hat{D}_{h,h'}(x) := \frac{1}{n}\sum_{i=1}^{n} g(X_i)(K_h \star K_{h'})'(x - X_i)$,

$$\tilde{A}(h) := \sup_{h' \in \tilde{\mathcal{H}}} \left\{ \|\hat{D}_{h,h'} - \hat{D}_{h'}\|_2 - \frac{\tilde{\chi}}{\sqrt{nh'^3}}\|g\|_\infty \|K'\|_2 \right\}_+ ,$$

where, given $\tilde{\varepsilon} > 0$, $\tilde{\chi} := (1 + \tilde{\varepsilon})(1 + \|K\|_1)$.

Estimation of $D = \frac{\partial}{\partial x}(g(x)N(x))$

SSPE If K is differentiable, $\int K = 1$ and $\int |K'|^2 < \infty$.

$$\hat{D}_h(x) := \frac{1}{n}\sum_{i=1}^{n} g(X_i)K'_h(x - X_i)$$

GL's trick

$\hat{D}_{h,h'}(x) := \frac{1}{n}\sum_{i=1}^{n} g(X_i)(K_h \star K_{h'})'(x - X_i)$,

$$\tilde{A}(h) := \sup_{h' \in \tilde{\mathcal{H}}} \left\{ \|\hat{D}_{h,h'} - \hat{D}_{h'}\|_2 - \frac{\tilde{\chi}}{\sqrt{nh'^3}}\|g\|_\infty \|K'\|_2 \right\}_+ ,$$

where, given $\tilde{\varepsilon} > 0$, $\tilde{\chi} := (1 + \tilde{\varepsilon})(1 + \|K\|_1)$.

Finally, we estimate D by using $\hat{D} := \hat{D}_{\tilde{h}}$ with

$$\tilde{h} := \mathrm{argmin}_{h \in \tilde{\mathcal{H}}} \left\{ \tilde{A}(h) + \frac{\tilde{\chi}}{\sqrt{nh^3}}\|g\|_\infty \|K'\|_2 \right\}.$$

Result for the derivative D

Oracle inequality for D

If $\tilde{\mathcal{H}} = \{1/\ell \;/\; \ell = 1, \ldots, \ell_{max}\}$ and if $\ell_{max} = \sqrt{\delta' n}$, if moreover $\|N\|_\infty$ and $\|g\|_\infty < \infty$, then for any $q \geq 1$,

$$\mathbb{E}\left(\|\hat{D} - D\|_2^{2q}\right) \leq \square_q \tilde{\chi}^{2q} \inf_{h \in \tilde{\mathcal{H}}} \left\{ \|K_h \star D - D\|_2^{2q} + \left[\frac{\|g\|_\infty \|K'\|_2}{\sqrt{nh^3}}\right]^{2q} \right\}$$

$$+ \square_{q, \tilde{\varepsilon}, \delta', \|K'\|_2, \|K\|_1, \|K'\|_1, \|N\|_\infty, \|g\|_\infty} \frac{1}{n^q}.$$

Estimation of λ and κ

SSPE

Estimation of λ and κ

SSPE λ is estimated via another (or simultaneous experiment).

Estimation of λ and κ

SSPE λ is estimated via another (or simultaneous experiment).

Assumption on $\hat{\lambda}$

There exist some $q > 1$ such that

- $\varepsilon_\lambda = \mathbb{E}[|\sqrt{n}(\hat{\lambda} - \lambda)|^q] < \infty$,
- $R_\lambda = \mathbb{E}(\hat{\lambda}^{2q}) < \infty$.

Estimation of λ and κ

SSPE λ is estimated via another (or simultaneous experiment).

Assumption on $\hat{\lambda}$

There exist some $q > 1$ such that

- $\varepsilon_\lambda = \mathbb{E}[|\sqrt{n}(\hat{\lambda} - \lambda)|^q] < \infty$,
- $R_\lambda = \mathbb{E}(\hat{\lambda}^{2q}) < \infty$.

Let $c > 0$,

$$\hat{\kappa} = \hat{\lambda} \frac{\sum_{i=1}^n X_i}{\sum_{i=1}^n g(X_i) + c}.$$

The inversion of $\mathcal{L}$

In SSPE, it remains to (approximately) invert $\mathcal{L}$. (see Perthame, Zubelli, Doumic (2009))

The inversion of $\mathcal{L}$

In SSPE, it remains to (approximately) invert $\mathcal{L}$. (see Perthame, Zubelli, Doumic (2009))

Define $T > 0$, an integer $k \geq 1$ and the regular grid on $[0, T]$ with mesh $k^{-1}T$ defined by

$0 = x_{0,k} < x_{1,k} < \cdots < x_{i,k} := \frac{i}{k}T < \ldots < x_{k,k} = T.$

Set $\varphi_{i,k} =: \frac{k}{T}\int_{x_{i,k}}^{x_{i+1,k}} \varphi(x)dx$ for $i = 0, \ldots, k-1$, and define by induction the sequence

$$H_{i,k}(\varphi) := \frac{1}{4}(H_{i/2,k}(\varphi) + \varphi_{i/2,k}) \text{ with } \begin{cases} H_0(\varphi) := & \frac{1}{3}\varphi_{1,k}, \\ H_1(\varphi) := & \frac{4}{21}\varphi_{0,k} + \frac{1}{7}\varphi_{1,k} \end{cases}$$

The inversion of $\mathcal{L}$

In SSPE, it remains to (approximately) invert $\mathcal{L}$. (see Perthame, Zubelli, Doumic (2009))

Define $T > 0$, an integer $k \geq 1$ and the regular grid on $[0, T]$ with mesh $k^{-1}T$ defined by

$0 = x_{0,k} < x_{1,k} < \cdots < x_{i,k} := \frac{i}{k}T < \ldots < x_{k,k} = T.$

Set $\varphi_{i,k} =: \frac{k}{T}\int_{x_{i,k}}^{x_{i+1,k}} \varphi(x)dx$ for $i = 0, \ldots, k-1$, and define by induction the sequence

$$H_{i,k}(\varphi) := \frac{1}{4}(H_{i/2,k}(\varphi)+\varphi_{i/2,k}) \text{ with } \begin{cases} H_0(\varphi) := & \frac{1}{3}\varphi_{1,k}, \\ H_1(\varphi) := & \frac{4}{21}\varphi_{0,k} + \frac{1}{7}\varphi_{1,k} \end{cases}$$

for any sequence $u_i, i = 1, 2, \ldots,$

$$u_{i/2} := \begin{cases} u_{i/2} & \text{if } i \text{ is even} \\ \frac{1}{2}(u_{(i-1)/2} + u_{(i+1)/2}) & \text{otherwise.} \end{cases}$$

The inversion of $\mathcal{L}$

In SSPE, it remains to (approximately) invert $\mathcal{L}$. (see Perthame, Zubelli, Doumic (2009))

Define $T > 0$, an integer $k \geq 1$ and the regular grid on $[0, T]$ with mesh $k^{-1}T$ defined by

$0 = x_{0,k} < x_{1,k} < \cdots < x_{i,k} := \frac{i}{k}T < \ldots < x_{k,k} = T$.

Set $\varphi_{i,k} =: \frac{k}{T}\int_{x_{i,k}}^{x_{i+1,k}} \varphi(x)dx$ for $i = 0, \ldots, k-1$, and define by induction the sequence

$$H_{i,k}(\varphi) := \frac{1}{4}(H_{i/2,k}(\varphi) + \varphi_{i/2,k}) \text{ with } \begin{cases} H_0(\varphi) := & \frac{1}{3}\varphi_{1,k}, \\ H_1(\varphi) := & \frac{4}{21}\varphi_{0,k} + \frac{1}{7}\varphi_{1,k} \end{cases}$$

Finally, we define

$$\mathcal{L}_k^{-1}(\varphi)(x) := \sum_{i=0}^{k-1} H_{i,k}(\varphi) 1_{[x_{i,k}, x_{i+1,k})}(x).$$

The inversion of $\mathcal{L}$

In SSPE, it remains to (approximately) invert $\mathcal{L}$. (see Perthame, Zubelli, Doumic (2009))

Define $T > 0$, an integer $k \geq 1$ and the regular grid on $[0, T]$ with mesh $k^{-1}T$ defined by

$0 = x_{0,k} < x_{1,k} < \cdots < x_{i,k} := \frac{i}{k}T < \ldots < x_{k,k} = T$.

Set $\varphi_{i,k} =: \frac{k}{T}\int_{x_{i,k}}^{x_{i+1,k}} \varphi(x)dx$ for $i = 0, \ldots, k-1$, and define by induction the sequence

$$H_{i,k}(\varphi) := \frac{1}{4}(H_{i/2,k}(\varphi) + \varphi_{i/2,k}) \text{ with } \begin{cases} H_0(\varphi) := & \frac{1}{3}\varphi_{1,k}, \\ H_1(\varphi) := & \frac{4}{21}\varphi_{0,k} + \frac{1}{7}\varphi_{1,k} \end{cases}$$

Finally, we define

$$\mathcal{L}_k^{-1}(\varphi)(x) := \sum_{i=0}^{k-1} H_{i,k}(\varphi) 1_{[x_{i,k}, x_{i+1,k})}(x).$$

The (approximative) inversion of $\mathcal{L}$

Finally, we define

$$\mathcal{L}_k^{-1}(\varphi)(x) := \sum_{i=0}^{k-1} H_{i,k}(\varphi) 1_{[x_{i,k}, x_{i+1,k})}(x).$$

Proposition

- $\mathcal{L}_k^{-1} : \mathbb{L}^2[0,T] \mapsto \mathbb{L}^2[0,T]$ is continuous
- $\|\mathcal{L}_k^{-1}(\varphi) - \mathcal{L}^{-1}(\varphi)\|_{2,T} \leq C \frac{T}{\sqrt{k}} \|\varphi\|_{\mathcal{W}^1}$, with $C < \frac{1}{\sqrt{6}}$.

We estimate $H = BN$ by

$$\hat{H} = \mathcal{L}_k^{-1}(\hat{\kappa}\hat{D} + \hat{\lambda}\hat{N}).$$

Oracle inequality for the estimation of $H = BN$

We establish an oracle inequality for $H = BN$ which is true under all previous assumptions.

Theorem

$$\mathbb{E}\left[\left\|\hat{H} - H\right\|_{2,T}^{q}\right] \leq C\left\{E_D + E_N + E_\lambda + E_{\mathcal{L}} + n^{-\frac{q}{2}}\right\}$$

Oracle inequality for the estimation of $H = BN$

We establish an oracle inequality for $H = BN$ which is true under all previous assumptions.

Theorem

$$\mathbb{E}\left[\left\|\hat{H} - H\right\|_{2,T}^{q}\right] \leq C\left\{E_D + E_N + E_\lambda + E_{\mathcal{L}} + n^{-\frac{q}{2}}\right\}$$

with

- $E_D = \sqrt{R_\lambda}\,\inf_{h\in\tilde{\mathcal{H}}}\left\{\|K_h \star D - D\|_2^q + \left(\frac{\|g\|_\infty \|K'\|_2}{\sqrt{nh^3}}\right)^q\right\}$

Oracle inequality for the estimation of $H = BN$

We establish an oracle inequality for $H = BN$ which is true under all previous assumptions.

Theorem

$$\mathbb{E}\left[\left\|\hat{H} - H\right\|_{2,T}^{q}\right] \leq C\left\{E_D + E_N + E_\lambda + E_{\mathcal{L}} + n^{-\frac{q}{2}}\right\}$$

with

- $E_D = \sqrt{R_\lambda}\inf_{h\in\tilde{\mathcal{H}}}\left\{\|K_h \star D - D\|_2^q + \left(\frac{\|g\|_\infty\|K'\|_2}{\sqrt{nh^3}}\right)^q\right\}$
- $E_N = \inf_{h\in\mathcal{H}}\left\{\|K_h \star N - N\|_2^q + \left(\frac{\|K\|_2}{\sqrt{nh}}\right)^q\right\}$

Oracle inequality for the estimation of $H = BN$

We establish an oracle inequality for $H = BN$ which is true under all previous assumptions.

Theorem

$$\mathbb{E}\left[\left\|\hat{H} - H\right\|_{2,T}^{q}\right] \leq C\left\{E_D + E_N + E_\lambda + E_{\mathcal{L}} + n^{-\frac{q}{2}}\right\}$$

with

- $E_D = \sqrt{R_\lambda}\inf_{h\in\tilde{\mathcal{H}}}\left\{\|K_h \star D - D\|_2^q + \left(\frac{\|g\|_\infty\|K'\|_2}{\sqrt{nh^3}}\right)^q\right\}$
- $E_N = \inf_{h\in\mathcal{H}}\left\{\|K_h \star N - N\|_2^q + \left(\frac{\|K\|_2}{\sqrt{nh}}\right)^q\right\}$
- $E_\lambda = \varepsilon_\lambda n^{-\frac{q}{2}}$

Oracle inequality for the estimation of $H = BN$

We establish an oracle inequality for $H = BN$ which is true under all previous assumptions.

Theorem

$$\mathbb{E}\left[\left\|\hat{H} - H\right\|_{2,T}^{q}\right] \leq C\left\{E_D + E_N + E_\lambda + E_{\mathcal{L}} + n^{-\frac{q}{2}}\right\}$$

with

- $E_D = \sqrt{R_\lambda}\inf_{h\in\tilde{\mathcal{H}}}\left\{\|K_h \star D - D\|_2^q + \left(\frac{\|g\|_\infty \|K'\|_2}{\sqrt{nh^3}}\right)^q\right\}$
- $E_N = \inf_{h\in\mathcal{H}}\left\{\|K_h \star N - N\|_2^q + \left(\frac{\|K\|_2}{\sqrt{nh}}\right)^q\right\}$
- $E_\lambda = \varepsilon_\lambda n^{-\frac{q}{2}}$
- $E_{\mathcal{L}} = \left((\|N\|_{\mathcal{W}^1} + \|gN\|_{\mathcal{W}^2})\frac{T}{\sqrt{k}}\right)^q$

Rate of convergence for the estimation of B

here We finally set $\hat{B} = \hat{H}/\hat{N}$ and $\tilde{B} = \max(\min(\hat{B}, \sqrt{n}), -\sqrt{n})$.

Rate of convergence for the estimation of B

here We finally set $\hat{B} = \hat{H}/\hat{N}$ and $\tilde{B} = \max(\min(\hat{B}, \sqrt{n}), -\sqrt{n})$. If $B \in \mathcal{W}_s$ ($s > 1/2$) and $g \in \mathcal{W}_{s+1}$, then (under suitable assumptions and enough moments for the kernel) $N \in \mathcal{W}_{s+1}$.

Rate of convergence for the estimation of B

here We finally set $\hat{B} = \hat{H}/\hat{N}$ and $\tilde{B} = \max(\min(\hat{B}, \sqrt{n}), -\sqrt{n})$.
If $B \in \mathcal{W}_s$ ($s > 1/2$) and $g \in \mathcal{W}_{s+1}$, then (under suitable assumptions and enough moments for the kernel) $N \in \mathcal{W}_{s+1}$.

Theorem

one can choose a family of $\mathcal{H}$ and $\mathcal{H}'$ independent of s such that for any compact $[a, b]$ of $[0, T]$ (under technical assumptions),

$$\mathbb{E}\left[\left\|(\tilde{B} - B)1_{[a,b]}\right\|_2^q\right] = O\left(n^{-\frac{qs}{2s+3}}\right).$$

Why is it the good rate?(1)

In the deterministic set-up

- we observe $N_\epsilon = N + \epsilon\zeta$, with $\|\zeta\|_2 \leq 1$ and

$$BN = \mathcal{L}^{-1}\left(\kappa\partial_x\big(g(x)N(x)\big) + \lambda N(x)\right).$$

- Since $\mathcal{L}^{-1}$ is continuous and the recovery of $\partial_x N$ is a more difficult inverse problem than the recovery of N, hence the ill-posedness is only due to ∂N (degree of ill-posedness $=1$)
- Hence if $N \in \mathcal{W}^s$, error in $\epsilon^{\frac{s}{s+1}}$.

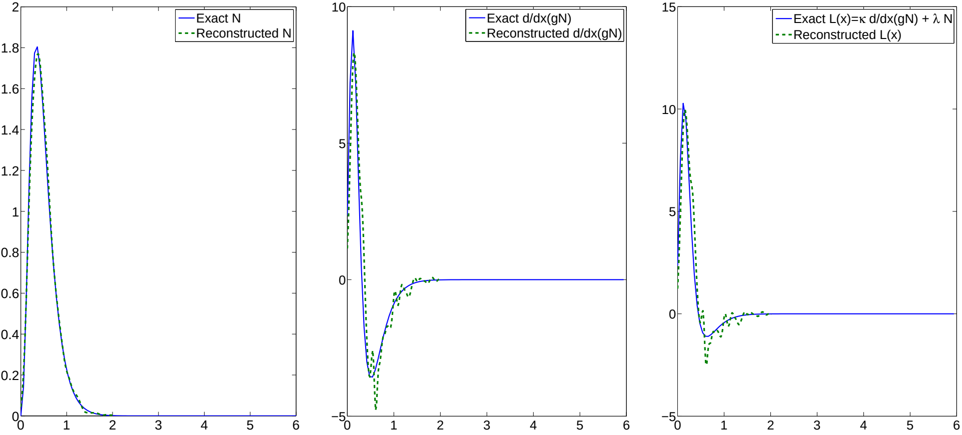
Why is it the good rate?(2)

In the n-sample set-up

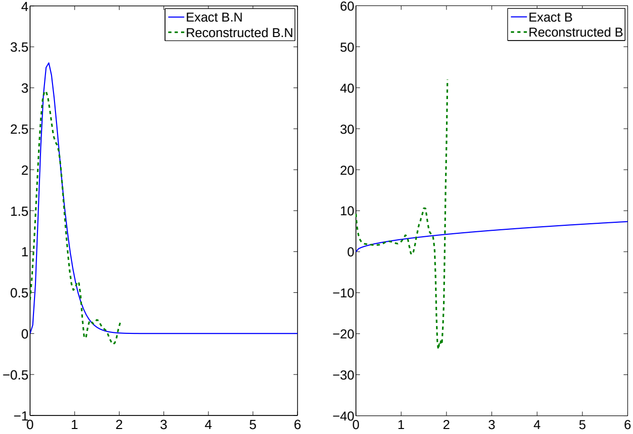
- problem well approximated by $N_\epsilon = N + \epsilon\mathbb{B}$ with $\mathbb{B}$ Gaussian white noise and $\epsilon = n^{-1/2}$.
- $\mathbb{B}$ is not in $\mathbb{L}_2$ but in $\mathcal{W}^{-1/2}$,
- Hence one needs to integrate ie $Z_\epsilon = \mathcal{I}^{1/2}N + \epsilon\mathcal{I}^{1/2}\mathbb{B}$ to have a noise in $\mathbb{L}_2$.
- Hence $Z_\epsilon = \mathcal{I}^{3/2}(\partial N) + \epsilon\mathcal{I}^{1/2}\mathbb{B}$ is of degree of ill-posedness 3/2.
- Hence if $N \in \mathcal{W}^s$, error in $\epsilon^{\frac{s}{s+3/2}} = n^{-\frac{s}{2s+3}}$.

Simulations

n=5000, Gaussian kernel, $B = 3\sqrt{x}$, $g = 1$.



Simulations



Perspectives

- Calibration and numerical optimization of the GL's method

Perspectives

- Calibration and numerical optimization of the GL's method
- To take into account noise in the measurements: Replace observations X_i with $X_i + Z_i$

Perspectives

- Calibration and numerical optimization of the GL's method
- To take into account noise in the measurements: Replace observations X_i with $X_i + Z_i$
- Extensions to fit with a more realistic biological model:

Perspectives

- Calibration and numerical optimization of the GL's method
- To take into account noise in the measurements: Replace observations X_i with $X_i + Z_i$
- Extensions to fit with a more realistic biological model:
  - The division law is given by a kernel $k(x, y)$:

$$... = 2 \int_x^\infty B(y)k(x,y)n(t,y)dy - B(x)n(t,x),$$

Division of the cell of size y into 2 cells of size x and $y - x$ with probability density$=k(x, y)$. Equal mitosis: $k(x,y) = \delta_{x=\frac{y}{2}}$, so $2\int_x^\infty B(y)k(x,y)n(t,y)dy = 4B(2x)n(t,2x)$

Perspectives

- Calibration and numerical optimization of the GL's method
- To take into account noise in the measurements: Replace observations X_i with $X_i + Z_i$
- Extensions to fit with a more realistic biological model:
  - The division law is given by a kernel $k(x, y)$:

$$... = 2 \int_x^\infty B(y)k(x,y)n(t,y)dy - B(x)n(t,x),$$

    Division of the cell of size y into 2 cells of size x and $y - x$ with probability density$=k(x,y)$. Equal mitosis: $k(x,y) = \delta_{x=\frac{y}{2}}$, so $2\int_x^\infty B(y)k(x,y)n(t,y)dy = 4B(2x)n(t,2x)$
  - Construct a microscopic stochastic system that matches with the PDE's approximation and that take advantage of richer observation schemes (Probabilistic works in progress studied by B. Cloez, V. Bansaye, M. Doumic, M. Hoffmann, N. Krell, T. Lepoutre, L. Robert,...)

References

Doumic, M. and Gabriel, P. (2010) *Eigenelements of a General Aggregation-Fragmentation Model*. Math. Models Methods Appl. Sci. 20(5), 757–783.

Doumic, M., Hoffmann, M., Reynaud-Bouret, P. and Rivoirard, V. (2011) Nonparametric estimation of the division rate of a size-structured population. To appear in SIAM J. Numer. Anal.

Doumic, M., Perthame, B. and Zubelli, J. (2009) *Numerical Solution of an Inverse Problem in Size-Structured Population Dynamics*. Inverse Problems, 25, 25pp.

Goldenshluger, A. and Lepski, O. (2009) *Uniform bounds for norms of sums of independent random functions* arXiv:0904.1950.

Goldenshluger, A. and Lepski, O. (2011) *Bandwidth selection in kernel density estimation: oracle inequalities and adaptive minimax optimality*. Ann. Statist. 39(3), 1608–1632.

Perthame, B. (2007) Transport equations in biology. In *Frontiers in Mathematics*, Frontiers in Mathematics. Birckhauser.

Perthame, B. and Ryzhik, L. (2005) *Exponential decay for the fragmentation or cell-division equation,* J. of Diff. Eqns, 210, 155–177 .

Perthame, B. and Zubelli, J. P. (2007) *On the inverse problem for a size-structured population model,* Inverse Problems, 23(3), 1037–1052.

Simulation study with Gaussian kernel

$n = 50\,000$, $\kappa = 1$, $g(x) = 1$ and 3 different functions B:

- $B_1(x) = 1$
- $B_2(x) = 1_{x \leq 1.5} + \text{affine part} + 5 \times 1_{x \geq 1.7}$ (B_2 continuous)
- $B_3(x) = \exp(-8(x-2)^2) + 1$

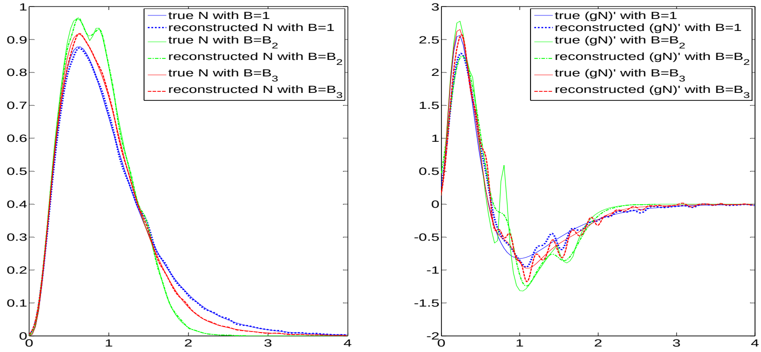


Figure: Reconstruction of N (left) and of D (right)

Simulation study with Gaussian kernel

$\kappa = 1, \; g(x) = x, \; B(x) = x^2$

SSPS

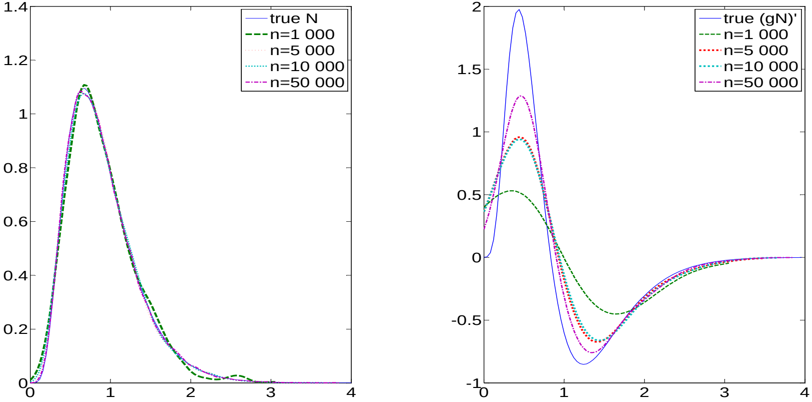


Figure: Reconstruction of N (left) and of D (right)

Simulation study with Gaussian kernel

$n = 50\,000$, $\kappa = 1$, $g(x) = 1$ and 3 different functions B:

- $B_1(x) = 1$
- $B_2(x) = 1_{x \leq 1.5} +$ affine part $+ 5 \times 1_{x \geq 1.7}$ (B_2 continuous)
- $B_3(x) = \exp(-8(x-2)^2) + 1$

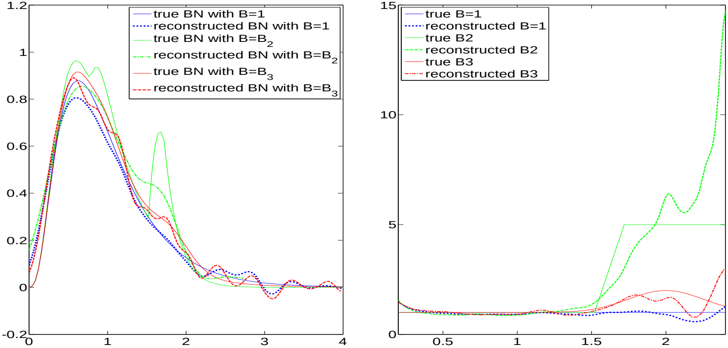


Figure: Reconstruction of BN (left) and of B (right)

Simulation study with Gaussian kernel

$\kappa = 1,\ g(x) = x,\ B(x) = x^2$

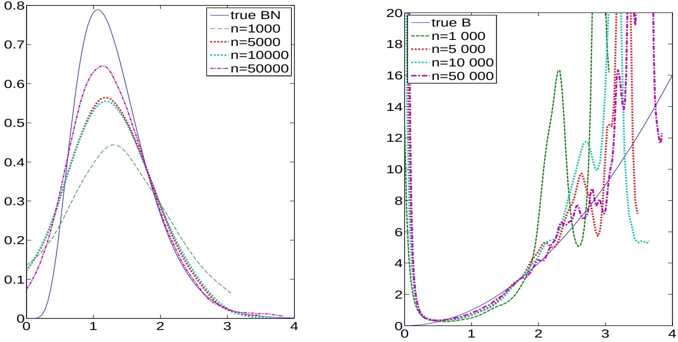


Figure: Reconstruction of BN (left) and of B (right)