Présentation orale - 11 Janvier 2024
Jérémy Bettinger

- Séminaire de recherche -

Modélisation de la croissance bactérienne

Sous la direction de Nathalie Krell
IRMAR, Université de Rennes.

Sommaire

Introduction

J'ai étudié un article coécrit par Nathalie Krell dont voici les références :

Scientific journal : Bernoulli, 21, 1760–1799.
Article : Statistical estimation of a growth-fragmentation model observed on a genealogical tree.
Author : Marie Doumic, Marc Hoffmann, Nathalie Krell and Lydia Robert.
First published : 2015.
Link : https://arxiv.org/pdf/1210.3240v2.pdf

Objectifs de l'article

Le but de ce séminaire est d'estimer statistiquement un modèle de croissance fragmentation d'une colonie de bactéries type Escherichia coli.

Objectifs de l'article

Le but de ce séminaire est d'estimer statistiquement un modèle de croissance fragmentation d'une colonie de bactéries type Escherichia coli.

Une équation classique en biologie est l'équation de croissance fragmentation :

Objectifs de l'article

Le but de ce séminaire est d'estimer statistiquement un modèle de croissance fragmentation d'une colonie de bactéries type Escherichia coli.

Une équation classique en biologie est l'équation de croissance fragmentation :

$$\begin{cases} \partial_t n(t,x) + \tau \partial_x \big(x\, n(t,x)\big) + B(x) n(t,x) = 4B(2x) n(t,2x) \\ \\ n(0,x) = n^{(0)}(x), x \geq 0, \end{cases}$$

où $n(t,x)$ désigne la proportion de cellules de taille x au temps t et B désigne le taux de division d'une cellule de taille x.

Objectifs de l'article

$$\begin{cases} \partial_t n(t,x) + \tau \partial_x \big(x\, n(t,x) \big) + B(x) n(t,x) = 4B(2x) n(t,2x) \\ \\ n(0,x) = n^{(0)}(x), x \geq 0, \end{cases}$$

Cette équation modélise une population dans laquelle les individus se divisent exactement en deux parties de tailles égales.

C'est le cas par exemple de la division cellulaire pour laquelle chaque cellule fille a la moitié de la taille de la cellule mère.

Objectifs de l'article

$$
\begin{cases}
\partial_t n(t,x) + \tau \partial_x \big(x\, n(t,x)\big) + B(x) n(t,x) = 4B(2x) n(t,2x) \\
\\
n(0,x) = n^{(0)}(x), x \geq 0,
\end{cases}
$$

Cette équation modélise une population dans laquelle les individus se divisent exactement en deux parties de tailles égales.

C'est le cas par exemple de la division cellulaire pour laquelle chaque cellule fille a la moitié de la taille de la cellule mère.

On suppose que la taille d'une cellule $x = x(t)$ au temps t évolue de manière exponentielle via l'équation différentielle :
$\mathrm{d}x(t) = \tau x(t)\,\mathrm{d}t$ où $\tau > 0$ est le taux de croissance d'une cellule.

- La constante τ est assimilée à une constante qui quantifie la capacité d'absorption de nutriments des cellules.

Ce que l'on va étudier

Hypothèses :

Ce que l'on va étudier

Hypothèses :

- Chaque cellule se divise en deux cellules filles selon un taux de division noté B, qui dépend de sa taille x.
  On note donc $B(x)$ le taux de division de la cellule mère de taille x.

Ce que l'on va étudier

Hypothèses :

- Chaque cellule se divise en deux cellules filles selon un taux de division noté B, qui dépend de sa taille x.
  On note donc $B(x)$ le taux de division de la cellule mère de taille x.
- Chaque cellule de taille x qui se divise se divisera en deux cellules de même taille $x/2$.

Ce que l'on va étudier

Hypothèses :

- Chaque cellule se divise en deux cellules filles selon un taux de division noté B, qui dépend de sa taille x.
  On note donc $B(x)$ le taux de division de la cellule mère de taille x.
- Chaque cellule de taille x qui se divise se divisera en deux cellules de même taille $x/2$.
- La taille de chaque cellule augmente de façon exponentielle avec le temps, mais à un rythme qui varie selon chaque individu. On prend τ **aléatoire** (et non pas déterministe).

Ce que l'on va étudier

Hypothèses :

- Chaque cellule se divise en deux cellules filles selon un taux de division noté B, qui dépend de sa taille x.
  On note donc $B(x)$ le taux de division de la cellule mère de taille x.
- Chaque cellule de taille x qui se divise se divisera en deux cellules de même taille $x/2$.
- La taille de chaque cellule augmente de façon exponentielle avec le temps, mais à un rythme qui varie selon chaque individu. On prend τ **aléatoire** (et non pas déterministe).

But : Estimer statistiquement la fonction B

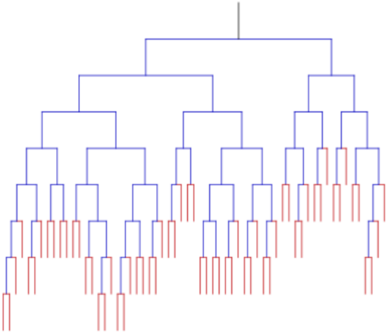
Cadre statistique

Soit

$$\mathcal{U} = \bigcup_{k=0}^{\infty} \{0,1\}^k$$

un arbre généalogique binaire (avec la convention $\{0,1\}^0 := \{\varnothing\}$).

On identifie chaque noeud $u \in \mathcal{U}$ avec une cellule qui a une taille à la naissance ξ_u et une durée de vie ζ_u.



Cadre statistique

- Notre but est d'estimer $\big(B(x), x \in [0,\infty)\big)$ sur les ensembles compacts de $(0,\infty)$.
- On notera $\mathcal{E}$ un tel ensemble compact et $\mathcal{S} = [0,+\infty) \times \mathcal{E}$.

Cadre statistique

- Notre but est d'estimer $\big(B(x), x \in [0,\infty)\big)$ sur les ensembles compacts de $(0,\infty)$.
- On notera $\mathcal{E}$ un tel ensemble compact et $\mathcal{S} = [0,+\infty) \times \mathcal{E}$.
- Notre travail statistique est basé sur les observations $\big((\xi_u, \zeta_u), u \in \mathcal{U}_n\big)$ où $\mathcal{U}_n \subset \mathcal{U}$ désigne un sous-ensemble connexe de taille n contenant la racine $u = \emptyset$.
  On fera une estimation en faisant $n \to +\infty$.

Cadre statistique

- Notre but est d'estimer $\big(B(x), x \in [0,\infty)\big)$ sur les ensembles compacts de $(0,\infty)$.
- On notera $\mathcal{E}$ un tel ensemble compact et $\mathcal{S} = [0,+\infty) \times \mathcal{E}$.
- Notre travail statistique est basé sur les observations $\big((\xi_u, \zeta_u), u \in \mathcal{U}_n\big)$ où $\mathcal{U}_n \subset \mathcal{U}$ désigne un sous-ensemble connexe de taille n contenant la racine $u = \emptyset$.
  On fera une estimation en faisant $n \to +\infty$.
- On suppose que taux de croissance τ peut varier avec chaque cellule $u \in \mathcal{U}$. Nous supposons qu'une cellule donnée u a un taux de croissance aléatoire $\tau_u = v \in \mathcal{E}$. Cette valeur v est héritée du taux de croissance v' de son parent selon une distribution $\rho(v', \mathrm{d}v)$.
- On note u^- le parent de u.

Hypothèses et théorème d'invariance

On suppose dans la suite les hypothèses :

1. Le taux de division B est continu.
2. On a $B(0) = 0$ et $\int_0^{+\infty} x^{-1} B(x)\, \mathrm{d}x = +\infty$.
3. Le noyau markovien $\rho(v, dv')$ est défini sur un ensemble compact $\mathcal{E}$ de $]0, +\infty[$

Hypothèses et théorème d'invariance

On suppose dans la suite les hypothèses :

1. Le taux de division B est continu.
2. On a $B(0) = 0$ et $\int_0^{+\infty} x^{-1} B(x)\, \mathrm{d}x = +\infty$.
3. Le noyau markovien $\rho(v, dv')$ est défini sur un ensemble compact $\mathcal{E}$ de $]0, +\infty[$

Théorème :

Sous les hypothèses précédentes, la probabilité de transition $\mathcal{P}$ du couple (ξ_u, τ_u) admet une probabilité invariante ν_B et vérifie :

$$\nu_B(y) = \frac{B(2y)}{y} \mathbb{E}_{\nu_B}\left[\frac{1}{\tau_{u^-}} \mathbf{1}_{\{\xi_{u^-} \leq 2y,\, \xi_u \geq y\}}\right]$$

où $\mathbb{E}_{\nu_B}[\cdot]$ est l'espérance avec la condition initiale $(\xi_\emptyset, \tau_\emptyset)$ de loi ν_B.

Utilisation du théorème à visée statistique

Ce qu'on peut réécire par :

$$B(y) = \frac{y}{2} \frac{\nu_B(y/2)}{\mathbb{E}_{\nu_B}\left[\frac{1}{\tau_{u^-}} \mathbf{1}_{\{\xi_{u^-} \leq y,\, \xi_u \geq y/2\}}\right]}.$$

Utilisation du théorème à visée statistique

Ce qu'on peut réécire par :

$$B(y) = \frac{y}{2} \frac{\nu_B(y/2)}{\mathbb{E}_{\nu_B}\left[\frac{1}{\tau_{u^-}} \mathbf{1}_{\{\xi_{u^-} \leq y,\, \xi_u \geq y/2\}}\right]}.$$

- On estime chaque terme du quotient.

Utilisation du théorème à visée statistique

Ce qu'on peut réécire par :

$$B(y) = \frac{y}{2} \frac{\nu_B(y/2)}{\mathbb{E}_{\nu_B}\left[\frac{1}{\tau_{u^-}} \mathbf{1}_{\{\xi_{u^-} \leq y,\, \xi_u \geq y/2\}}\right]}.$$

- On estime chaque terme du quotient.
- On ne sait pas si notre dénominateur est petit ou non...

Utilisation du théorème à visée statistique

Ce qu'on peut réécire par :

$$B(y) = \frac{y}{2} \frac{\nu_B(y/2)}{\mathbb{E}_{\nu_B}\left[\frac{1}{\tau_{u^-}} \mathbf{1}_{\{\xi_{u^-} \leq y,\, \xi_u \geq y/2\}}\right]}.$$

- On estime chaque terme du quotient.
- On ne sait pas si notre dénominateur est petit ou non...
- On va prendre le max avec une quantité $\varpi > 0$ petite arbitraire.

Estimation de chaque membre

$$B(y) = \frac{y}{2} \frac{\nu_B(y/2)}{\mathbb{E}_{\nu_B}\left[\frac{1}{\tau_{u^-}} \mathbf{1}_{\{\xi_{u^-} \leq y,\, \xi_u \geq y/2\}}\right]}.$$

- Le numérateur est estimé par estimation non paramétrique à l'aide d'un noyau

$$K : [0, \infty) \to \mathbb{R}, \quad \int_{[0,\infty)} K(y)dy = 1,$$

renormalisé $K_h(y) = h^{-1} K\left(h^{-1} y\right)$ pour $y \in [0, \infty)$ et $h > 0$.

Estimation de chaque membre

$$B(y) = \frac{y}{2} \frac{\nu_B(y/2)}{\mathbb{E}_{\nu_B}\left[\frac{1}{\tau_{u^-}} \mathbf{1}_{\{\xi_{u^-} \leq y,\, \xi_u \geq y/2\}}\right]}.$$

- Le numérateur est estimé par estimation non paramétrique à l'aide d'un noyau

$$K : [0, \infty) \to \mathbb{R}, \quad \int_{[0,\infty)} K(y) dy = 1,$$

renormalisé $K_h(y) = h^{-1} K\left(h^{-1} y\right)$ pour $y \in [0, \infty)$ et $h > 0$.

- Le dénominateur est quant à lui estimé par la loi des grands nombres.

Estimateur sélectionné

$$B(y) = \frac{y}{2} \frac{\nu_B(y/2)}{\mathbb{E}_{\nu_B}\left[\frac{1}{\tau_{u^-}} \mathbf{1}_{\{\xi_{u^-} \leq y,\, \xi_u \geq y/2\}}\right]}.$$

On obtient l'estimateur :

$$\widehat{B}_n(y) = \frac{y}{2} \frac{n^{-1} \sum_{u \in \mathcal{U}_n} K_h(\xi_u - y/2)}{n^{-1} \sum_{u \in \mathcal{U}_n} \frac{1}{\tau_{u^-}} \mathbf{1}_{\{\xi_{u^-} \leq y, \xi_u \geq y/2\}} \bigvee \varpi}.$$

Estimateur sélectionné

$$B(y) = \frac{y}{2} \frac{\nu_B(y/2)}{\mathbb{E}_{\nu_B}\left[\frac{1}{\tau_{u^-}} \mathbf{1}_{\{\xi_{u^-} \leq y,\, \xi_u \geq y/2\}}\right]}.$$

On obtient l'estimateur :

$$\widehat{B}_n(y) = \frac{y}{2} \frac{n^{-1} \sum_{u \in \mathcal{U}_n} K_h(\xi_u - y/2)}{n^{-1} \sum_{u \in \mathcal{U}_n} \frac{1}{\tau_{u^-}} \mathbf{1}_{\{\xi_{u^-} \leq y, \xi_u \geq y/2\}} \bigvee \varpi}.$$

- **Notre but est d'obtenir une vitesse satisfaisante pour $\widehat{B}_n$.**

Estimateur sélectionné

$$B(y) = \frac{y}{2} \frac{\nu_B(y/2)}{\mathbb{E}_{\nu_B}\left[\frac{1}{\tau_{u^-}} \mathbf{1}_{\{\xi_{u^-} \leq y,\, \xi_u \geq y/2\}}\right]}.$$

On obtient l'estimateur :

$$\widehat{B}_n(y) = \frac{y}{2} \frac{n^{-1} \sum_{u \in \mathcal{U}_n} K_h(\xi_u - y/2)}{n^{-1} \sum_{u \in \mathcal{U}_n} \frac{1}{\tau_{u^-}} \mathbf{1}_{\{\xi_{u^-} \leq y, \xi_u \geq y/2\}} \bigvee \varpi}.$$

- **Notre but est d'obtenir une vitesse satisfaisante pour $\widehat{B}_n$.**

Hypothèses supplémentaires :

1. K est à support compact
2. Pour $n_0 \geq 1$, on a pour $0 \leq k \leq n_0$:

$$\int_{[0,\infty)} x^k K(x)\, \mathrm{d}x = \mathbf{1}_{\{k=0\}}.$$

Définitions

Définition

Soit $\lambda > 0$ *et un vecteur de constantes stictement positives* $\mathfrak{c} = (r, m, \ell, L)$. *On introduit la classe de fonctions continues* $\mathcal{F}^\lambda(\mathfrak{c})$ $B : [0, \infty) \to [0, \infty)$ *qui vérifient :*

$$\int_0^{r/2} x^{-1} B(2x) dx \leq L, \quad \int_{r/2}^{r} x^{-1} B(2x) dx \geq \ell, \tag{1}$$

et

$$B(x) \geq m\, x^\lambda \quad \textit{pour} \quad x \geq r. \tag{2}$$

Définitions

Définition

Soit $\lambda > 0$ et un vecteur de constantes stictement positives $\mathfrak{c} = (r, m, \ell, L)$. On introduit la classe de fonctions continues $\mathcal{F}^{\lambda}(\mathfrak{c})$ $B : [0, \infty) \to [0, \infty)$ qui vérifient :

$$\int_0^{r/2} x^{-1} B(2x) dx \leq L, \quad \int_{r/2}^{r} x^{-1} B(2x) dx \geq \ell, \tag{1}$$

et

$$B(x) \geq m\, x^{\lambda} \;\; \textit{pour} \;\; x \geq r. \tag{2}$$

Définition

Soit $\rho_{\min}, \rho_{\max}$ deux probabilités sur $\mathcal{E}$. On définit $\mathcal{M}(\rho_{\min}, \rho_{\max})$ comme la classe des transitions de Markov $\rho(v, dv')$ sur $\mathcal{E}$ telle que

$$\rho_{\min}(A) \leq \rho(v, A) \leq \rho_{\max}(A), \quad A \subset \mathcal{E}, v \in \mathcal{E}. \tag{3}$$

Définitions

Définition

Donnons nous dans ce qui suit un vecteur (à coordonnées positives) $\mathfrak{c} = (r, m, \ell, L)$ *et* $0 < e_{\min} \leq e_{\max}$ *tel que* $\mathcal{E} \subset [e_{\min}, e_{\max}]$*. On introduit la fonction de Lyapunov :*

$$\mathbb{V}(x, v) = \mathbb{V}(x) = \exp\left(\frac{m}{e_{\min}\lambda} x^{\lambda}\right) \quad \textit{pour} \quad (x, v) \in \mathcal{S} \tag{4}$$

et le taux

$$\delta = \delta(\mathfrak{c}) := \frac{1}{1 - 2^{-\lambda}} \exp\left(- \left(1 - 2^{-\lambda}\right) \frac{m}{e_{\max}\lambda} r^{\lambda}\right).$$

Définitions

Définition

On note $\gamma_{B,\mathbb{V}}$ *le rayon spectral de l'opérateur* $\mathcal{P}_B - 1 \otimes \nu_B$ *qui agit sur l'espace de Banach des fonctions* $g : \mathcal{S} \to \mathbb{R}$ *telles que*

$$\sup\{|g(\boldsymbol{x})|/\mathbb{V}(\boldsymbol{x}), \boldsymbol{x} \in \mathcal{S}\} < \infty.$$

Définitions

Définition

On note $\gamma_{B,\mathbb{V}}$ le rayon spectral de l'opérateur $\mathcal{P}_B - 1 \otimes \nu_B$ qui agit sur l'espace de Banach des fonctions $g : \mathcal{S} \to \mathbb{R}$ telles que

$$\sup\{|g(\boldsymbol{x})|/\mathbb{V}(\boldsymbol{x}), \boldsymbol{x} \in \mathcal{S}\} < \infty.$$

On fait une hypothèse supplémentaire dans la suite :

On suppose dans la suite $\delta(\mathfrak{c}) < \frac{1}{2}$ et

$$\sup_{B \in \mathcal{F}^{\lambda}(\mathfrak{c})} \gamma_{B,\mathbb{V}} < \frac{1}{2}. \tag{5}$$

Définition

Pour $s > 0$, avec $s = \lfloor s \rfloor + \{s\}$, $0 < \{s\} \leq 1$ et $\lfloor s \rfloor$ un entier , on introduit l'espace de Holder $\mathcal{H}^s(\mathcal{D})$ de fonctions $f : \mathcal{D} \to \mathbb{R}$ qui possèdent une dérivée d'ordre $\lfloor s \rfloor$ satisfaisant :

$$|f^{\lfloor s \rfloor}(y) - f^{\lfloor s \rfloor}(x)| \leq c(f)|x - y|^{\{s\}}. \tag{6}$$

La constante minimale $c(f)$ vérifiant (6) définie une semie norme $|f|_{\mathcal{H}^s(\mathcal{D})}$. On muni l'espace $\mathcal{H}^s(\mathcal{D})$ de la norme

$$\|f\|_{\mathcal{H}^s(\mathcal{D})} = \|f\|_{L^\infty(\mathcal{D})} + |f|_{\mathcal{H}^s(\mathcal{D})}$$

et on pose la boule de rayon $M > 0$ dans l'espace de Holder :

$$\mathcal{H}^s(\mathcal{D}, M) = \{B,\ \|B\|_{\mathcal{H}^s(\mathcal{D})} \leq M\}.$$

Idées et notations

- On notera $a_n \lesssim b_n$ s'il existe $C > 0$ tel que $a_n \leq C\, b_n$.

Idées et notations

- On notera $a_n \lesssim b_n$ s'il existe $C > 0$ tel que $a_n \leq C\, b_n$.
- On va prendre $\varpi = \varpi_n$ qui tend vers 0 afin d'obtenir une estimation satisfaisante de B.

On a les expressions :

$$\widehat{B}_n(2y) = y \frac{n^{-1} \sum_{u \in \mathcal{U}_n} K_{h_n}(\xi_u - y)}{n^{-1} \sum_{u \in \mathcal{U}_n} \frac{1}{\tau_{u^-}} \mathbf{1}_{\{\xi_{u^-} \leq 2y,\, \xi_u \geq y\}} \bigvee \varpi_n}$$

et

$$B(2y) = y \frac{\nu_B(y)}{\mathbb{E}_{\nu_B}\left[\frac{1}{\tau_{u^-}} \mathbf{1}_{\{\xi_{u^-} \leq 2y,\, \xi_u \geq y\}}\right]}.$$

Idées et notations

- On notera $a_n \lesssim b_n$ s'il existe $C > 0$ tel que $a_n \leq C\, b_n$.
- On va prendre $\varpi = \varpi_n$ qui tend vers 0 afin d'obtenir une estimation satisfaisante de B.

On a les expressions :

$$\widehat{B}_n(2y) = y \frac{n^{-1} \sum_{u \in \mathcal{U}_n} K_{h_n}(\xi_u - y)}{n^{-1} \sum_{u \in \mathcal{U}_n} \frac{1}{\tau_{u^-}} \mathbf{1}_{\{\xi_{u^-} \leq 2y,\, \xi_u \geq y\}} \bigvee \varpi_n}$$

et

$$B(2y) = y \frac{\nu_B(y)}{\mathbb{E}_{\nu_B}\left[\frac{1}{\tau_{u^-}} \mathbf{1}_{\{\xi_{u^-} \leq 2y,\, \xi_u \geq y\}}\right]}.$$

Cherchons une vitesse de convergence de $\widehat{B}_n$ vers B.

Découpage

On a l'égalité $\widehat{B}_n(2y) - B(2y) = y(I + II + III)$ avec :

Découpage

On a l'égalité $\widehat{B}_n(2y) - B(2y) = y(I + II + III)$ avec :

$$I = \frac{K_{h_n} \star \nu_B(y) - \nu_B(y)}{D(y)}\,; \quad II = \frac{K_{h_n} \star \widehat{\nu}_n(y) - K_{h_n} \star \nu_B(y)}{D_n(y)_{\varpi_n}}\,;$$

$$III = \frac{K_{h_n} \star \nu_B(y)}{D_n(y)_{\varpi_n} D(y)} \left(D(y) - D_n(y)_{\varpi_n}\right)$$

Découpage

On a l'égalité $\widehat{B}_n(2y) - B(2y) = y(I + II + III)$ avec :

$$I = \frac{K_{h_n} \star \nu_B(y) - \nu_B(y)}{D(y)}\ ; \quad II = \frac{K_{h_n} \star \widehat{\nu}_n(y) - K_{h_n} \star \nu_B(y)}{D_n(y)_{\varpi_n}}\ ;$$

$$III = \frac{K_{h_n} \star \nu_B(y)}{D_n(y)_{\varpi_n} D(y)} \big(D(y) - D_n(y)_{\varpi_n}\big)$$

où on a posé pour tout $y \in (0, \infty)$ et $u \in \mathcal{U}$ avec $|u| \geq 1$:

$$D(y) = \mathbb{E}_{\nu_B}\Big[\tfrac{1}{\tau_{u^-}} \mathbf{1}_{\{\xi_{u^-} \leq 2y,\, \xi_u \geq y\}}\Big], \tag{7}$$

Découpage

On a l'égalité $\widehat{B}_n(2y) - B(2y) = y(I + II + III)$ avec :

$$I = \frac{K_{h_n} \star \nu_B(y) - \nu_B(y)}{D(y)}\ ; \quad II = \frac{K_{h_n} \star \widehat{\nu}_n(y) - K_{h_n} \star \nu_B(y)}{D_n(y)_{\varpi_n}}\ ;$$

$$III = \frac{K_{h_n} \star \nu_B(y)}{D_n(y)_{\varpi_n} D(y)} \big(D(y) - D_n(y)_{\varpi_n}\big)$$

où on a posé pour tout $y \in (0, \infty)$ et $u \in \mathcal{U}$ avec $|u| \geq 1$:

$$D(y) = \mathbb{E}_{\nu_B}\Big[\tfrac{1}{\tau_{u^-}} \mathbf{1}_{\{\xi_{u^-} \leq 2y,\, \xi_u \geq y\}}\Big], \tag{7}$$

son estimée

$$D_n(y)_{\varpi_n} = n^{-1} \sum_{u \in \mathcal{U}_n} \tfrac{1}{\tau_{u^-}} \mathbf{1}\{\xi_{u^-} \leq 2y, \xi_u \geq y\} \bigvee \varpi_n, \tag{8}$$

Découpage

On a l'égalité $\widehat{B}_n(2y) - B(2y) = y(I + II + III)$ avec :

$$I = \frac{K_{h_n} \star \nu_B(y) - \nu_B(y)}{D(y)}\,; \quad II = \frac{K_{h_n} \star \widehat{\nu}_n(y) - K_{h_n} \star \nu_B(y)}{D_n(y)_{\varpi_n}}\,;$$

$$III = \frac{K_{h_n} \star \nu_B(y)}{D_n(y)_{\varpi_n} D(y)} \big(D(y) - D_n(y)_{\varpi_n}\big)$$

où on a posé pour tout $y \in (0, \infty)$ et $u \in \mathcal{U}$ avec $|u| \geq 1$:

$$D(y) = \mathbb{E}_{\nu_B}\Big[\tfrac{1}{\tau_{u^-}} \mathbf{1}_{\{\xi_{u^-} \leq 2y,\, \xi_u \geq y\}}\Big], \tag{7}$$

son estimée

$$D_n(y)_{\varpi_n} = n^{-1} \sum_{u \in \mathcal{U}_n} \tfrac{1}{\tau_{u^-}} \mathbf{1}_{\{\xi_{u^-} \leq 2y,\, \xi_u \geq y\}} \bigvee \varpi_n, \tag{8}$$

et la mesure empirique des observations $(\xi_u, u \in \mathcal{U}_n)$:

$$\widehat{\nu}_n(\mathrm{d}y) = \frac{1}{n} \sum_{u \in \mathcal{U}_n} \delta_{\xi_u}(\mathrm{d}y). \tag{9}$$

Découpage

Par changement de variable on a :

$$\|\widehat{B}_n - B\|^2_{L^2(\mathcal{D})} = 2\int_{\frac{1}{2}\mathcal{D}} \left(\widehat{B}_n(2y) - B(2y)\right)^2 dy \lesssim IV + V + VI,$$

avec

$$IV = \int_{\frac{1}{2}\mathcal{D}} \left(K_{h_n} \star \nu_B(y) - \nu_B(y)\right)^2 \tfrac{y^2}{D(y)^2} dy$$

$$V = \int_{\frac{1}{2}\mathcal{D}} \left(K_{h_n} \star \widehat{\nu}_n(y) - K_{h_n} \star \nu_B(y)\right)^2 D_n(y)_{\varpi}^{-2} y^2 dy$$

$$VI = \int_{\frac{1}{2}\mathcal{D}} \left(D_n(y)_{\varpi} - D(y)\right)^2 \left(K_{h_n} \star \nu_B(y)\right)^2 \left(D_n(y)_{\varpi} D(y)\right)^{-2} y^2 dy.$$

Démonstration d'une majoration

Proposition

On a l'estimation

$$IV = \int_{\frac{1}{2}\mathcal{D}} \left(K_{h_n} \star \nu_B(y) - \nu_B(y)\right)^2 \frac{y^2}{D(y)^2} dy \lesssim h_n^{2s} \tag{10}$$

uniformément en $B \in \mathcal{H}^s(\mathcal{D}, M)$.

Démonstration d'une majoration

Proposition

On a l'estimation

$$IV = \int_{\frac{1}{2}\mathcal{D}} \left(K_{h_n} \star \nu_B(y) - \nu_B(y)\right)^2 \frac{y^2}{D(y)^2} dy \lesssim h_n^{2s} \qquad (10)$$

uniformément en $B \in \mathcal{H}^s(\mathcal{D}, M)$.

Démonstration (idée).

Démonstration d'une majoration

Proposition

On a l'estimation

$$IV = \int_{\frac{1}{2}\mathcal{D}} \big(K_{h_n} \star \nu_B(y) - \nu_B(y)\big)^2 \frac{y^2}{D(y)^2}\, dy \lesssim h_n^{2s} \qquad (10)$$

uniformément en $B \in \mathcal{H}^s(\mathcal{D}, M)$.

Démonstration (idée).

- On borne uniformément en B le terme $1/D(y)$.
  Pour cela on montre que $\inf\{D(y)\,;\, y \in \mathcal{D}\}$ est strictement positif pour toute les classes $B \in \mathcal{H}^s(\mathcal{D}/2, M) \cap \mathcal{F}^\lambda(\mathfrak{c})$.

Démonstration d'une majoration

Proposition

On a l'estimation

$$IV = \int_{\frac{1}{2}\mathcal{D}} \left(K_{h_n} \star \nu_B(y) - \nu_B(y)\right)^2 \frac{y^2}{D(y)^2} dy \lesssim h_n^{2s} \qquad (10)$$

uniformément en $B \in \mathcal{H}^s(\mathcal{D}, M)$.

Démonstration (idée).

- On borne uniformément en B le terme $1/D(y)$.
  Pour cela on montre que $\inf\{D(y)\,;\, y \in \mathcal{D}\}$ est strictement positif pour toute les classes $B \in \mathcal{H}^s(\mathcal{D}/2, M) \cap \mathcal{F}^\lambda(\mathfrak{c})$.
- On borne y en utilisant le fait que $\mathcal{D}$ soit borné.

Démonstration d'une majoration

Proposition

On a l'estimation

$$IV = \int_{\frac{1}{2}\mathcal{D}} \big(K_{h_n} \star \nu_B(y) - \nu_B(y)\big)^2 \frac{y^2}{D(y)^2}\, dy \lesssim h_n^{2s} \qquad (10)$$

uniformément en $B \in \mathcal{H}^s(\mathcal{D}, M)$.

Démonstration (idée).

- On borne uniformément en B le terme $1/D(y)$.
  Pour cela on montre que $\inf\{D(y)\,;\, y \in \mathcal{D}\}$ est strictement positif pour toute les classes $B \in \mathcal{H}^s(\mathcal{D}/2, M) \cap \mathcal{F}^\lambda(\mathfrak{c})$.
- On borne y en utilisant le fait que $\mathcal{D}$ soit borné.
- Avec l'inégalité d'estimation par noyau, on obtient :

$$IV \lesssim \|K_{h_n} \star \nu_B - \nu_B\|^2_{L^2(\mathcal{D}/2)} \lesssim |\nu_B|^2_{\mathcal{H}^s(\mathcal{D}/2)}\ h_n^{2s}.$$

Démonstration d'une majoration

Proposition

On a l'estimation

$$IV = \int_{\frac{1}{2}\mathcal{D}} \left(K_{h_n} \star \nu_B(y) - \nu_B(y)\right)^2 \frac{y^2}{D(y)^2} dy \lesssim h_n^{2s} \qquad (11)$$

uniformément en $B \in \mathcal{H}^s(\mathcal{D}, M)$.

Démonstration (idée).

- Avec l'inégalité d'estimation par noyau, on obtient :

$$IV \lesssim \|K_{h_n} \star \nu_B - \nu_B\|^2_{L^2(\mathcal{D}/2)} \lesssim |\nu_B|^2_{\mathcal{H}^s(\mathcal{D}/2)}\ h_n^{2s}.$$

- On montre que $|\nu_B|^2_{\mathcal{H}^s(\mathcal{D}/2)}$ est majorée par une constante qui ne dépend que de e_{min}, e_{max}, $\mathcal{D}$ et de $\|B\|_{\mathcal{H}^s(\mathcal{D})}$.

$\square$

Propositions admises

Proposition

On a l'estimation

$$\mathbb{E}_\mu[V] \lesssim \varpi_n^{-2} |\log h_n| (nh_n)^{-1} \tag{12}$$

uniformément en $B \in \mathcal{F}^\lambda(\mathfrak{c})$.

Propositions admises

Proposition

On a l'estimation

$$\mathbb{E}_\mu[V] \lesssim \varpi_n^{-2} |\log h_n| (nh_n)^{-1} \tag{12}$$

uniformément en $B \in \mathcal{F}^\lambda(\mathfrak{c})$.

Proposition

On a l'estimation

$$\mathbb{E}_\mu\left[VI\right] \lesssim \varpi_n^{-2} n^{-1} \tag{13}$$

uniformément en $B \in \mathcal{F}^\lambda(\mathfrak{c})$.

Vitesse de convergence

Mis bout à bout (11), (12) et (13). On obtient :

$$\mathbb{E}_\mu\left[\|\widehat{B}_n - B\|^2_{L^2(\mathcal{D})}\right] \lesssim h_n^{2s} + \varpi_n^{-2}|\log h_n|(nh_n)^{-1} + \varpi_n^{-2}n^{-1}$$

uniformément en $B \in \mathcal{F}^\lambda(\mathfrak{c}) \cap \mathcal{H}^s(\mathcal{D}, M)$.

Vitesse de convergence

Mis bout à bout (11), (12) et (13). On obtient :

$$\mathbb{E}_\mu\left[\|\widehat{B}_n - B\|^2_{L^2(\mathcal{D})}\right] \lesssim h_n^{2s} + \varpi_n^{-2}|\log h_n|(nh_n)^{-1} + \varpi_n^{-2}n^{-1}$$

uniformément en $B \in \mathcal{F}^\lambda(\mathfrak{c}) \cap \mathcal{H}^s(\mathcal{D}, M)$.

Rester à 'égaliser' les vitesses.
Pour cela on prend le pas $h_n = n^{-1/(2s+1)}$ et $\varpi_n = (\log n)^{-1}$:

Vitesse de convergence

Mis bout à bout (11), (12) et (13). On obtient :

$$\mathbb{E}_\mu\left[\|\widehat{B}_n - B\|^2_{L^2(\mathcal{D})}\right] \lesssim h_n^{2s} + \varpi_n^{-2}|\log h_n|(nh_n)^{-1} + \varpi_n^{-2}n^{-1}$$

uniformément en $B \in \mathcal{F}^\lambda(\mathfrak{c}) \cap \mathcal{H}^s(\mathcal{D}, M)$.

Rester à 'égaliser' les vitesses.
Pour cela on prend le pas $h_n = n^{-1/(2s+1)}$ et $\varpi_n = (\log n)^{-1}$:

$$\begin{aligned}\mathbb{E}_\mu\left[\|\widehat{B}_n - B\|^2_{L^2(\mathcal{D})}\right] &\lesssim \left(\frac{1}{n}\right)^{\frac{2s}{(2s+1)}} + \frac{(\log n)^3}{2s+1}\left(\frac{1}{n}\right)^{\frac{2s}{(2s+1)}} + \frac{(\log n)^2}{n} \\ &\lesssim \left(\frac{1}{n}\right)^{\frac{2s}{(2s+1)}}(\log n)^3\end{aligned}$$

Vitesse de convergence

Mis bout à bout (11), (12) et (13). On obtient :

$$\mathbb{E}_\mu\left[\|\widehat{B}_n - B\|^2_{L^2(\mathcal{D})}\right] \lesssim h_n^{2s} + \varpi_n^{-2}|\log h_n|(nh_n)^{-1} + \varpi_n^{-2}n^{-1}$$

uniformément en $B \in \mathcal{F}^\lambda(\mathfrak{c}) \cap \mathcal{H}^s(\mathcal{D}, M)$.

Rester à 'égaliser' les vitesses.
Pour cela on prend le pas $h_n = n^{-1/(2s+1)}$ et $\varpi_n = (\log n)^{-1}$:

$$\begin{aligned}\mathbb{E}_\mu\left[\|\widehat{B}_n - B\|^2_{L^2(\mathcal{D})}\right] &\lesssim \left(\frac{1}{n}\right)^{\frac{2s}{(2s+1)}} + \frac{(\log n)^3}{2s+1}\left(\frac{1}{n}\right)^{\frac{2s}{(2s+1)}} + \frac{(\log n)^2}{n} \\ &\lesssim \left(\frac{1}{n}\right)^{\frac{2s}{(2s+1)}} (\log n)^3\end{aligned}$$

soit

$$\sup_{\rho, B} \mathbb{E}_\mu\left[\|\widehat{B}_n - B\|^2_{L^2(\mathcal{D})}\right]^{1/2} \lesssim (\log n)^{3/2}\ \ n^{-s/(2s+1)}. \quad \square$$

Théorème

Théorème :
Plaçons nous sous nos hypothèses, le choix d'un noyau K, et d'un pas h_n vérifiant pour $n_0 > 0$:

$$h_n = c_0 n^{-1/(2s+1)}, \quad \varpi_n = (\log n)^{-1}.$$

Alors pour tout $M > 0$ il existe $c_0 = c_0(\mathfrak{c}, M)$ et $d(\mathfrak{c}) \geq 0$ tel que pour $0 < s < n_0$ et tout compact $\mathcal{D} \subset (d(\mathfrak{c}), \infty)$ tel que $\inf \mathcal{D} \geq r/2$, *on ait :*

$$\boxed{\sup_{\rho, B} \mathbb{E}_\mu \left[\| \widehat{B}_n - B \|^2_{L^2(\mathcal{D})} \right]^{1/2} \lesssim (\log n)^{3/2} \, n^{-s/(2s+1)}}$$

où le sup est pris sur $\rho \in \mathcal{M}(\rho_{\min}, \rho_{\max})$ et $B \in \mathcal{F}^\lambda(\mathfrak{c}) \cap \mathcal{H}^s(\mathcal{D}, M)$.

De plus, $\mathbb{E}_\mu[\cdot]$ est l'espérance par rapport à la distribution initiale $\mu(d\boldsymbol{x})$ pour $(\xi_\emptyset, \tau_\emptyset)$ sur $\mathcal{S}$ avec la condition $\int_{\mathcal{S}} \mathbb{V}(\boldsymbol{x})^2 \mu(d\boldsymbol{x}) < \infty$.

Remarques

$$\sup_{\rho,B} \mathbb{E}_\mu \left[\| \widehat{B}_n - B \|^2_{L^2(\mathcal{D})} \right]^{1/2} \lesssim (\log n)^{3/2}\, n^{-s/(2s+1)}$$

Remarque 1 :

Ce résultat nous fournit une vitesse, au sens du risque quadratique, en $n^{-s/(2s+1)}$ avec des termes en $\log n$.

Remarques

$$\sup_{\rho,B} \mathbb{E}_\mu \left[\| \widehat{B}_n - B \|^2_{L^2(\mathcal{D})} \right]^{1/2} \lesssim (\log n)^{3/2} \, n^{-s/(2s+1)}$$

Remarque 1 :

Ce résultat nous fournit une vitesse, au sens du risque quadratique, en $n^{-s/(2s+1)}$ avec des termes en $\log n$.

Remarque 2 :

On a le même énoncé en gardant $\varpi_n \xrightarrow[n\to+\infty]{} 0$. On a l'estimation :

$$\sup_{\rho,B} \mathbb{E}_\mu \left[\| \widehat{B}_n - B \|^2_{L^2(\mathcal{D})} \right]^{1/2} \lesssim \varpi_n^{-1} \, (\log n)^{1/2} \, n^{-s/(2s+1)}.$$

Remarques

$$\sup_{\rho,B} \mathbb{E}_\mu \left[\| \widehat{B}_n - B \|^2_{L^2(\mathcal{D})} \right]^{1/2} \lesssim (\log n)^{3/2} \, n^{-s/(2s+1)}$$

Remarque 1 :

Ce résultat nous fournit une vitesse, au sens du risque quadratique, en $n^{-s/(2s+1)}$ avec des termes en $\log n$.

Remarque 2 :

On a le même énoncé en gardant $\varpi_n \xrightarrow[n \to +\infty]{} 0$. On a l'estimation :

$$\sup_{\rho,B} \mathbb{E}_\mu \left[\| \widehat{B}_n - B \|^2_{L^2(\mathcal{D})} \right]^{1/2} \lesssim \varpi_n^{-1} \, (\log n)^{1/2} \, n^{-s/(2s+1)}.$$

Remarque 3 :

Le choix de ϖ_n en $(\log n)^{-1}$ est satisfaisant car ϖ_n tend vers 0 lentement, et d'un point de vue statistique, bien que $\log n$ tend vers l'infini, ce terme s'apparente presque à une constante au vu du nombre limité de données.

Merci de votre attention !