ESTIMATION NON-PARAMÉTRIQUE DE LA FONCTION DE HASARD AVEC VARIABLE EXPLICATIVE FONCTIONNELLE: CAS DES DONNÉES SPATIALES

ALI LAKSACI et BOUBAKEUR MECHAB

In this paper, we study a kernel estimator of the conditional hazard function when the covariates take values in some abstract function space. The almost completely convergence (with rate) of this estimate is obtained when the sample considered is collected in spatial order with mixing structure. These results are extensions to spatial dependent data of the results given by Ferraty *et al.* (Rev. Roumaine Math. Pures Appl. **53** (2008), 1–18).

AMS 2010 Subject Classification: 62G05, 62G20, 62N02.

Key words: données spatiales, champ aléatoire fonctionnel, estimation non-paramétrique, fonction de hasard conditionnelle, probabilités de petites boules.

1. INTRODUCTION

Les problèmes statistiques liés à la modélisation et à l'étude de données spatiales connaissent ces derniers temps un grand intérêt en statistique. L'importance de ce thème de recherche est motivée par la croissance du nombre de problèmes concrets pour lesquels les données sont récoltées dans un ordre spatial. De tels problèmes se rencontrent dans de nombreux domaines tels l'épidémiologie, l'économétrie, les sciences de l'environnement et de la terre, l'agronomie, l'imagerie, ... En statistique non paramétrique la modélisation des données spatiales est relativement récente par rapport au cas paramétrique. En effet, les premiers résultats ont été obtenus par Tran [21], tandis que les références les plus récentes incluent celles de Dabo-Niang et Yao [6], Carbon *et al.* [4] et Li et Tran [17]. L'objectif de ce travail est d'étudier l'estimation non paramétrique de la fonction de hasard conditionnelle lorsque les observations sont à la fois spatialement dépendantes et la coraviable est dans un espace semi-métrique de dimension éventuellement infinie.

L'estimation de la fonction de hasard joue un rôle très important en statistique. En effet, elle est utilisée dans l'analyse de risque ou pour l'étude des phénomènes de survie dans de nombreux domaines tels (médecine, géophysique, fiabilité, ...). La littérature sur l'estimation de la fonction de hasard est très abondante, lorsque les observations sont vectorielles. Citons, par exemple, Watson et Leadbetter [22], Roussas [20], Lecoutre et Ould-Said [15], Estèvez *et al.* [7] et Quintela-del-Rio [13] pour des références récentes. Dans tous ces travaux les auteurs considèrent des observations indépendantes ou des données dépendantes issues de séries temporelles. Dans le cas spatial, la littérature est très restreinte, il n'y a qu'un seul travail de Li et Tran [16]. Ces derniers ont obtenu, dans un contexte spatial, la normalité asymptotique d'un estimateur à noyau de la fonction de hasard. Les premiers résultats sur l'estimation non paramétrique de ce modèle, en statistique fonctionnelle, ont été obtenus par Ferraty *et al.* [9]. Ils ont étudié la convergence presque complète d'un estimateur à noyau pour la fonction de hasard d'une variable aléatoire réelle conditionnée par une variable explicative fonctionnelle. La normalité asymptotique de ce dernier estimateur a été obtenue, dans le cas α-mélangeant, par Quintela-del-Rio [14]. Nous renvoyons à Ferraty *et al.* [10] pour la convergence presque complète uniforme sur la composante fonctionnelle de ce modèle non paramétrique.

L'apport principal de ce travail est la généralisation des résultats de Ferraty *et al.* [9] au cas des observations spatialement dépendantes. Sous des conditions de mélange assez générales, on établit la convergence presque complète (avec vitesse) d'un estimateur à noyau pour la fonction de hasard d'une variable aléatoire réelle conditionnée par une variable explicative fonctionnelle. Notons que, comme toutes les propriétés asymptotiques en statistique non paramétrique fonctionnelle, notre résultat est lié au phénomène de concentration de la mesure de probabilité de la variable explicative et à la régularité de l'espace fonctionnel du modèle. Mentionnons également qu'en statistique spatiale, on distingue deux types d'asymptotiques (voir Cressie [5]). L'asymptotique extensive et l'asymptotique intensive. La première traite le cas où la taille des observations croît avec celle du domaine d'observation. En pratique, cette asymptotique est utilisée lorsque les observations sont collectées par des stations de mesure séparées. On trouvera des exemples de données fonctionnelles, adaptées à cet asymptotique, en économie (les courbes de consommation d'un produit quelconque dans différentes villes), environnement (les courbes de concentration d'un gaz polluant dans différentes régions) ou en agronomie (les courbes de pluviométrie dans différentes localités). L'asymptotique intensive examine la situation d'observations qui se densifient dans un domaine borné fixé. Cela est le cas, par exemple, en prospection minière ou en analyse radiographique.

Nous présentons l'estimateur de notre modèle spatial dans le Paragraphe 2. Dans le Paragraphe 3 nous donnons les hypothèses, puis nous étudions la convergence presque complète de cet estimateur. Comme application, nous traitons, dans le Paragraphe 4, l'estimation de risque maximum conditionnellement à une variable explicative fonctionnelle.

2. LE MODÈLE ET SON ESTIMATEUR

Soit N un entier naturel dans $\mathbb{N}^*$. On considère le champ aléatoire $Z_{\mathbf{i}} = (X_{\mathbf{i}}, Y_{\mathbf{i}})$, $\mathbf{i} \in \mathbb{N}^N$ à valeurs dans $\mathcal{F} \times \mathbb{R}$, où $(\mathcal{F}, d)$ est un espace semi-métrique de dimension éventuellement infinie. Dans ce contexte, $(X_{\mathbf{i}})_{\mathbf{i}\in\mathbb{N}^N}$ peut être une variable aléatoire fonctionnelle. Notons que, depuis une bonne dizaine d'années, la communauté statistique s'est préoccupée du développement de modèles et méthodes adaptés à ce contexte de données fonctionnelles. Alors que les premières études dans cette direction se sont essentiellement concentrées sur des modèles linéaires (voir Bosq [2], Ramsay et Silverman [19]), les développements récents (voir Ferraty et Vieu [11]) font état de modèles non-paramétriques adaptés à ce type de données.

Par la suite, on fixe un point x dans $\mathcal{F}$ (respectivement, un compact $\mathcal{S} \in \mathbb{R}$), on suppose que les observations spatiales $(X_{\mathbf{i}}, Y_{\mathbf{i}})_{\mathbf{i}\in\mathbb{N}^N}$ ont la même distribution que $Z := (X, Y)$ et que la version régulière de la probabilité conditionnelle de Y sachant $X = x$ existe et admet une densité bornée par rapport à la mesure de Lebesgue sur $\mathbb{R}$, notée f^x. Le paramètre fonctionnel étudié dans cet article, noté h^x, est défini, pour tout $y \in \mathbb{R}$ tel que $F^x(y) < 1$, par

$$h^x(y) = \frac{f^x(y)}{1 - F^x(y)},$$

où F^x est la fonction de répartition conditionnelle de Y sachant $X = x$.

Par ailleurs, nous supposons que le champ aléatoire fonctionnel est observé sur l'ensemble $I_{\mathbf{n}} = \{\mathbf{i} = (i_1, \ldots, i_N) \in \mathbb{N}^N,\ 1 \le i_k \le n_k,\ k = 1, \ldots, N\}$, $\mathbf{n} = (n_1, \ldots, n_N) \in \mathbb{N}^N$ et on estime la fonction de répartition conditionnelle par

$$\widehat{F}^x(y) = \frac{\sum_{\mathbf{i}\in I_{\mathbf{n}}} K(h_K^{-1} d(x, X_{\mathbf{i}})) H(h_H^{-1}(y - Y_{\mathbf{i}}))}{\sum_{\mathbf{i}\in I_{\mathbf{n}}} K(h_K^{-1} d(x, X_{\mathbf{i}}))} \quad \forall y \in \mathbb{R},$$

où K est un noyau, H une fonction de répartition et $h_K = h_{K,\mathbf{n}}$ (respectivement, $h_H = h_{H,\mathbf{n}}$) une suite de nombres réels positifs.

De $\widehat{F}^x$, on déduit un estimateur de la densité conditionnelle, noté $\widehat{f}^x$, défini par

$$\widehat{f}^x(y) = \frac{h_H^{-1} \sum_{\mathbf{i}\in I_{\mathbf{n}}} K(h_K^{-1} d(x, X_{\mathbf{i}})) H'(h_H^{-1}(y - Y_{\mathbf{i}}))}{\sum_{\mathbf{i}\in I_{\mathbf{n}}} K(h_K^{-1} d(x, X_{\mathbf{i}}))} \quad \forall y \in \mathbb{R},$$

où H' est la dérivée de H. Notons que, si $N = 1$, on obtient les mêmes estimateurs de Ferraty *et al.* [8]. L'estimateur naturel de la fonction de hasard conditionnelle, noté $\widehat{h}^x$, est

$$\widehat{h}^x(y) = \frac{\widehat{f}^x(y)}{1 - \widehat{F}^x(y)} \quad \forall y \in \mathbb{R}.$$

Le but de ce travail est d'étudier la convergence presque complète de l'estimateur $\widehat{h}^x$ vers h^x, lorsque le champ aléatoire fonctionnel $(Z_{\mathbf{i}})_{\mathbf{i}\in\mathbb{N}^N}$ vérifie la condition de mélange suivante.

$$\begin{cases} \text{Il existe une fonction } \varphi(t) \downarrow 0 \text{ quand } t \to \infty, \text{ telle que} \\ \forall\, E,\ E' \text{ sous ensemble de } \mathbb{N}^N \text{ de cardinal fini} \\ \alpha\left(\mathcal{B}(E),\, \mathcal{B}(E')\right) = \sup\limits_{B\in\mathcal{B}(E),\, C\in\mathcal{B}(E')} \left|\mathbb{P}\left(B\cap C\right) - \mathbb{P}\left(B\right)\mathbb{P}\left(C\right)\right| \leq \\ \leq \psi\left(\mathrm{Card}\left(E\right), \mathrm{Card}(E')\right)\varphi\left(\mathrm{dist}\left(E, E'\right)\right), \end{cases}$$

où $\mathcal{B}(E)$ (resp. $\mathcal{B}(E')$) la tribu borélienne engendrée par $(Z_{\mathbf{i}},\ \mathbf{i}\in E)$ $\big($resp. $(Z_{\mathbf{i}},\ \mathbf{i}\in E')\big)$, $\mathrm{Card}(E)$ $\big($resp. $\mathrm{Card}(E')\big)$ est le cardinal de E $\big($resp. $E'\big)$, $\mathrm{dist}(E, E')$ désigne la distance euclidien entre E et E' et ψ une fonction symétrique : $\mathbb{N}^2 \to \mathbb{R}^+$, décroissante par rapport aux deux variables séparément et satisfait l'une des conditions suivantes

$$\psi\left(n, m\right) \leq C \min\left(n, m\right), \quad n, m \in \mathbb{N} \tag{1}$$

ou

$$\psi\left(n, m\right) \leq C\left(n + m + 1\right)^{\widetilde{\beta}}, \quad n, m \in \mathbb{N} \tag{2}$$

pour certains $\widetilde{\beta} \geq 1$ et $C > 0$.

Notons que ces conditions ont été utilisées par Tran [21] et elle sont vérifiées par de nombreux modèles spatiaux (voir Guyon [12]).

3. PROPRIÉTÉS ASYMPTOTIQUES

Dans la suite, nous noterons par C et/ou C' des constantes strictement positives quelconques et nous supposons que: $\forall y \in \mathcal{S},\ F^x(y) < 1$. Rappelons que dans ce contexte spatial, $\mathbf{n} \to \infty$ signifie que $\min\{n_k\} \to \infty$ et que pour chaque $1 \leq j, k \leq N$ on a $|\frac{n_j}{n_k}| < C < \infty$. Introduisons les hypothèses suivantes

(H1) $\mathbb{P}(X \in B(x, r)) = \phi_x(r) > 0$ où $B(x, r)$ est la boule fermée, centrée en x et de rayon r.

(H2) La fonction φ vérifie $\sum\limits_{i=1}^{\infty} i^\delta \varphi(i) < \infty$, $\delta > 5N$.

(H3) $0 < \sup\limits_{\mathbf{i}\neq\mathbf{j}} \mathbb{P}\left[(X_{\mathbf{i}}, X_{\mathbf{j}}) \in B(x, h) \times B(x, h)\right] \leq C(\phi_x(h))^{(a+1)/a}$, avec $1 < a < \delta N^{-1}$.

(H4) Il existe un voisinage de x, noté N_x, tel que: $\forall(x_1,x_2) \in N_x^2$, $\forall(y_1,y_2) \in \mathcal{S}^2$

$$|F^{x_1}(y_1) - F^{x_2}(y_2)| \leq C\big(d^{b_1}(x_1,x_2) + |y_1 - y_2|^{b_2}\big), \quad b_1 > 0,\, b_2 > 0$$
$$|f^{x_1}(y_1) - f^{x_2}(y_2)| \leq C\big(d^{b_3}(x_1,x_2) + |y_1 - y_2|^{b_4}\big), \quad b_3 > 0,\, b_4 > 0.$$

(H5) K est de support $[0,1]$, vérifie $C\mathbb{1}_{[0,1]}(\cdot) \leq K(\cdot) \leq C'\mathbb{1}_{[0,1]}(\cdot)$, où $\mathbb{1}_{[0,1]}$ est la fonction indicatrice de $[0,1]$.

(H6) H est une fonction de classe C^2 et à support compact.

(H7) Il existe $0 < \alpha < (\delta - 5N)/3N$ et $\eta_0 > 0$, tels que:

$$\lim_{\mathbf{n}\to\infty} \widehat{\mathbf{n}}^{\alpha}\, h_H = \infty \quad \text{et} \quad C\,\widehat{\mathbf{n}}^{\frac{(5+3\alpha)N-\delta}{\delta}+\eta_0} \leq h_H \phi_x(h),$$

où $\widehat{\mathbf{n}} = n_1 \dots n_N$.

Nos hypothèses sont relativement standard, puisque les conditions (H1), (H3)–(H7) sont très similaires à celles utilisées par Ferraty *et al.* [9], tandis que l'hypothèse (H2) est la même que dans Tran [21].

THÉORÈME 1. *Sous les hypothèses* (1), (2) *et* (H1)–(H7), *on a,*

$$\sup_{y\in\mathcal{S}} |\widehat{h}^x(y) - h^x(y)| = O\left(h_K^{b_1} + h_H^{b_2}\right) + O\left(\left(\frac{\log\widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\, h_H \phi_x(h_K)}\right)^{\frac{1}{2}}\right), \quad p.co.$$

Il est intéressant de noter que si $N = 1$ on obtient la même vitesse de convergence de Ferraty *et al.* [9].

Démonstration. On considère la décomposition

$$\widehat{h}^x(y) - h^x(y) = \frac{1}{1-\widehat{F}^x(y)}\left[\widehat{f}^x(y) - f^x(y)\right] + \frac{h^x(y)}{1-\widehat{F}^x(y)}\left[\widehat{F}^x(y) - F^x(y)\right].$$

La démonstration du Théorème 1 repose sur les résultats suivants

$$\sup_{y\in\mathcal{S}} |\widehat{F}^x(y) - F^x(y)| = O\left(h_K^{b_1} + h_H^{b_2}\right) + O\left(\left(\frac{\log\widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\, \phi_x(h_K)}\right)^{\frac{1}{2}}\right) \quad \text{p.co.}$$

$$\sup_{y\in\mathcal{S}} |\widehat{f}^x(y) - f^x(y)| = O\left(h_K^{b_1} + h_H^{b_2}\right) + O\left(\left(\frac{\log\widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\, h_H \phi_x(h_K)}\right)^{\frac{1}{2}}\right) \quad \text{p.co.}$$

$$\exists \eta > 0 \text{ tel que } \sum_{\mathbf{n}\in\mathbb{N}^N} \mathbb{P}\left\{\inf_{y\in\mathcal{S}} |1 - \widehat{F}^x(y)| < \eta\right\} < \infty,$$

dont les démonstrations sont basées, respectivement, sur les décompositions

$$\widehat{F}^x(y) - F^x(y) = \frac{1}{\widehat{F}_D^x}\left\{\left(\widehat{F}_N^x(y) - \mathbb{E}\widehat{F}_N^x(y)\right) - \left(F^x(y) - \mathbb{E}\widehat{F}_N^x(y)\right)\right\} + \\ + \frac{F^x(y)}{\widehat{F}_D^x}\left(\mathbb{E}\widehat{F}_D^x - \widehat{F}_D^x\right)$$

et

$$\widehat{f}^x(y) - f^x(y) = \frac{1}{\widehat{F}_D^x}\left\{\left(\widehat{f}_N^x(y) - \mathbb{E}\widehat{f}_N^x(y)\right) - \left(f^x(y) - \mathbb{E}\widehat{f}_N^x(y)\right)\right\} + \\ + \frac{f^x(y)}{\widehat{F}_D^x}\left(\mathbb{E}\widehat{F}_D^x - \widehat{F}_D^x\right)$$

où

$$\widehat{F}_D^x = \frac{1}{\widehat{\mathbf{n}}\mathbb{E}\left[K(h_K^{-1}d(x, X_{\mathbf{1}}))\right]}\sum_{\mathbf{i}\in I_{\mathbf{n}}} K(h_K^{-1}d(x, X_{\mathbf{i}})),$$

$$\widehat{F}_N^x(y) = \frac{1}{\widehat{\mathbf{n}}\mathbb{E}\left[K(h_K^{-1}d(x, X_{\mathbf{1}}))\right]}\sum_{\mathbf{i}\in I_{\mathbf{n}}} K(h_K^{-1}d(x, X_{\mathbf{i}}))H(h_H^{-1}(y - Y_{\mathbf{i}}),$$

et

$$\widehat{f}_N^x(y) = \frac{1}{\widehat{\mathbf{n}}h_H\mathbb{E}\left[K(h_K^{-1}d(x, X_{\mathbf{1}}))\right]}\sum_{\mathbf{i}\in I_{\mathbf{n}}} K(h_K^{-1}d(x, X_{\mathbf{i}}))H'(h_H^{-1}(y - Y_{\mathbf{i}}),$$

où $\mathbf{1}$ est l'indice spatial des composantes fixées à 1.

Finalement, le Théorème 1 est une conséquence des résultats suivants.

Lemme 2. *Sous les hypothèses* (1), (2), (H1)–(H3), (H5) *et* (H7)*, on a,*

$$\widehat{F}_D^x - \mathbb{E}\widehat{F}_D^x = O\left(\left(\frac{\log\widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\,\phi(h_K)}\right)^{\frac{1}{2}}\right), \quad p.co.$$

Corollaire 3. *Sous les hypothèses du Lemme* 2*, on a,*

$$\sum_{\mathbf{n}\in\mathbb{N}^N}\mathbb{P}\left(\widehat{F}_D^x < 1/2\right) < \infty.$$

Lemme 4. *Sous les hypothèses* (H1) *et* (H4)–(H7)*, on a,*

$$\sup_{y\in\mathcal{S}}\left|F^x(y) - \mathbb{E}\widehat{F}_N^x(y)\right| = O\left(h_K^{b_1} + h_H^{b_2}\right).$$

Lemme 5. *Sous les conditions du Théorème* 1*, on a,*

$$\sup_{y\in\mathcal{S}}\left|\widehat{F}_N^x(y) - \mathbb{E}\widehat{F}_N^x(y)\right| = O\left(\left(\frac{\log\widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\,\phi(h_K)}\right)^{\frac{1}{2}}\right), \quad p.co.$$

Lemme 6. *Sous les hypothèses* (H1) *et* (H4)–(H7)*, on a,*

$$\sup_{y\in\mathcal{S}}\left|f^x(y)-\mathbb{E}\widehat{f}_N^x(y)\right| = O\left(h_K^{b_3}+h_H^{b_4}\right).$$

Lemme 7. *Sous les conditions du Théorème* 1*, on a,*

$$\sup_{y\in\mathcal{S}}\left|\widehat{f}_N^x(y)-\mathbb{E}\widehat{f}_N^x(y)\right| = O\left(\left(\frac{\log\widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\,h_H\phi_x(h_K)}\right)^{\frac{1}{2}}\right), \quad p.co.$$

Lemme 8. *Sous les conditions du Théorème* 1*, on a,*

$$\exists\eta>0 \text{ tel que } \sum_{\mathbf{n}\in\mathbb{N}^N}\mathbb{P}\left\{\inf_{y\in\mathcal{S}}|1-\widehat{F}^x(y)|<\eta\right\}<\infty. \quad \square$$

4. APPLICATION: ESTIMATION DU POINT À HAUT RISQUE

Dans cette section, on se propose d'estimer le point à haut risque dans $\mathcal{S}$, noté $\theta(x)$, défini par

$$h^x(\theta(x)) = \max_{y\in\mathcal{S}} h^x(y). \tag{3}$$

Ce modèle a un grand intérêt en statistique, notamment dans l'analyse de risque sismique (voir Quintela-del-Rio [13]). Dans notre contexte fonctionnel, on suppose qu'il existe un point unique $\theta(x)$ dans $\mathcal{S}$ vérifiant (3). L'estimateur naturel de $\theta(x)$, noté $\widehat{\theta}(x)$, est tel que

$$\widehat{h}^x(\widehat{\theta}(x)) = \max_{y\in\mathcal{S}} \widehat{h}^x(y). \tag{4}$$

En général, cet estimateur n'est pas unique. Ainsi, dans toute la suite de cet article $\widehat{\theta}(x)$ désignera toute variable aléatoire vérifiant (4).

Afin d'étudier la convergence presque complète de l'estimateur $\widehat{\theta}(x)$, on garde les hypothèses de la section précédente et on suppose que la fonction h^x est de classe C^2 par rapport à y et telle que

$$h^{x'}(\theta(x)) = 0 \quad \text{et} \quad h^{x''}(\theta(x)) > 0. \tag{5}$$

Finalement, le théorème précédent nous permet de conclure le corollaire suivant:

Corollaire 9. *Sous les conditions* (1), (2), (H1)–(H7) *et* (5)*, on a,*

$$\widehat{\theta}(x)-\theta(x) = O\left(h_K^{b_1/2}+h_H^{b_2/2}\right)+O\left(\left(\frac{\log\widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\,h_H\phi_x(h_K)}\right)^{\frac{1}{4}}\right), \quad p.co.$$

5. APPENDICE

Dans les preuves suivantes nous noterons pour tout $\mathbf{i} \in I_{\mathbf{n}}$

$$K_{\mathbf{i}}(x) = K(h_K^{-1} d(x, X_{\mathbf{i}})), \quad H_{\mathbf{i}}(y) = H(h_H^{-1}(y - Y_{\mathbf{i}}))$$

et

$$H'_{\mathbf{i}}(y) = H'(h_H^{-1}(y - Y_{\mathbf{i}})).$$

Démonstration du Lemme 2. La démonstration est basée sur des idées similaires à celles utilisées par Carbon et al. [3]. En effet, on a

$$\widehat{F}_D^x(x) - \mathbb{E}[\widehat{F}_D^x(x)] = \frac{1}{\widehat{\mathbf{n}}\mathbb{E}\left[K_{\mathbf{1}}(x)\right]} \sum_{\mathbf{i} \in I_{\mathbf{n}}} \Delta_{\mathbf{i}}(x)$$

où $\Delta_{\mathbf{i}}(x) K_{\mathbf{i}}(x) - \mathbb{E}\left[K_{\mathbf{i}}(x)\right]$. On considère la décomposition spatiale de Tran [21] sur les variables $\Delta_{\mathbf{i}}(x)$, définie, pour un entier $p_{\mathbf{n}}$ fixé, de la manière suivante

$$U(1, \mathbf{n}, x, \mathbf{j}) = \sum_{\substack{i_k = 2j_k p_{\mathbf{n}} + 1 \\ k = 1, \ldots, N}}^{2j_k p_{\mathbf{n}} + p_{\mathbf{n}}} \Delta_{\mathbf{i}}(x),$$

$$U(2, \mathbf{n}, x, \mathbf{j}) = \sum_{\substack{i_k = 2j_k p_{\mathbf{n}} + 1 \\ k = 1, \ldots, N-1}}^{2j_k p_{\mathbf{n}} + p_{\mathbf{n}}} \sum_{i_N = 2j_N p_{\mathbf{n}} + p_{\mathbf{n}} + 1}^{2(j_N + 1) p_{\mathbf{n}}} \Delta_{\mathbf{i}}(x),$$

$$U(3, \mathbf{n}, x, \mathbf{j}) = \sum_{\substack{i_k = 2j_k p_{\mathbf{n}} + 1 \\ k = 1, \ldots, N-2}}^{2j_k p_{\mathbf{n}} + p_{\mathbf{n}}} \sum_{i_{N-1} = 2j_{N-1} p_{\mathbf{n}} + p_{\mathbf{n}} + 1}^{2(j_{N-1} + 1) p_{\mathbf{n}}} \sum_{i_N = 2j_N p_{\mathbf{n}} + 1}^{2j_N p_{\mathbf{n}} + p_{\mathbf{n}}} \Delta_{\mathbf{i}}(x),$$

$$U(4, \mathbf{n}, x, \mathbf{j}) = \sum_{\substack{i_k = 2j_k p_{\mathbf{n}} + 1 \\ k = 1, \ldots, N-2}}^{2j_k p_{\mathbf{n}}} \sum_{i_{N-1} = 2j_{N-1} p_{\mathbf{n}} + p_{\mathbf{n}} + 1}^{2(j_{N-1} + 1) p_{\mathbf{n}}} \sum_{i_N = 2j_N p_{\mathbf{n}} + p_{\mathbf{n}} + 1}^{2(j_N + 1) p_{\mathbf{n}}} \Delta_{\mathbf{i}}(x),$$

et ainsi de suite. Les deux derniers termes sont

$$U(2^{N-1}, \mathbf{n}, x, \mathbf{j}) = \sum_{\substack{i_k = 2j_k p_{\mathbf{n}} + p_{\mathbf{n}} + 1 \\ k = 1, \ldots, N-1}}^{2(j_k + 1) p_{\mathbf{n}}} \sum_{i_N = 2j_N p_{\mathbf{n}} + 1}^{2j_N p_{\mathbf{n}} + p_{\mathbf{n}}} \Delta_{\mathbf{i}}(x),$$

$$U(2^N, \mathbf{n}, x, \mathbf{j}) = \sum_{\substack{i_k = 2j_k p_{\mathbf{n}} + p_{\mathbf{n}} + 1 \\ k = 1, \ldots, N}}^{2(j_k + 1) p_{\mathbf{n}}} \Delta_{\mathbf{i}}(x).$$

Pour $r_i = 2^{-1}n_i p_{\mathbf{n}}^{-1}$, $i = 1,\ldots,N$ et $\mathcal{J} = \{0,\ldots,r_1-1\}\times\ldots\times\{0,\ldots,r_N-1\}$, on pose

$$T(\mathbf{n},x,i) = \sum_{\mathbf{j}\in\mathcal{J}} U(i,\mathbf{n},x,\mathbf{j}).$$

Sans perte de généralité, on peut écrire

$$|\widehat{F}_D^x(x) - \mathbb{E}[\widehat{F}_D^x(x)]| = \frac{1}{\widehat{\mathbf{n}}\mathbb{E}\left[K_{\mathbf{1}}(x)\right]}\sum_{i=1}^{2^N} T(\mathbf{n},x,i), \tag{6}$$

parce que même si $n_{\mathbf{i}}$ n'est pas exactement égal à $2r_{\mathbf{i}}p_{\mathbf{n}}$, on regroupe les variables restantes dans un bloc $T(\mathbf{n},x,2^N+1)$ (ceci ne changera pas la démonstration voir Biau et Cadre [1]).

Maintenant, en vertu de la dernière équation (6), pour chaque $\eta > 0$ on a

$$\mathbb{P}\left(|\widehat{F}_D^x(x) - \mathbb{E}[\widehat{F}_D^x(x)]| \geq \eta\right) \leq 2^N \max_{i=1,\ldots} \mathbb{P}\left(T(\mathbf{n},x,i) \geq \eta\widehat{\mathbf{n}}\mathbb{E}\left[K_{\mathbf{1}}(x)\right]\right).$$

Ainsi, il suffit de calculer

$$\mathbb{P}\left(T(\mathbf{n},x,i) \geq \eta\widehat{\mathbf{n}}\mathbb{E}\left[K_{\mathbf{1}}(x)\right]\right) \quad \text{pour chaque } i = 1,\ldots,2^N.$$

On traite seulement le cas $i = 1$. Pour cela, on re-numérote les variables $(U(1,\mathbf{n},x,\mathbf{j});\, \mathbf{j}\in\mathcal{J})$ et on applique le Lemme 4.5 de Carbon et al. [3] sur les variables re-numérotées. Les variables avec la nouvelle numérotation seront notées $Z_1,\ldots,Z_M$ où $M = \prod_{k=1}^N r_k = 2^{-N}\widehat{\mathbf{n}}p_{\mathbf{n}}^{-N} \leq \widehat{\mathbf{n}}p_{\mathbf{n}}^{-N}$. Remarquons que, pour chaque Z_j il existe un certain $\mathbf{j}$ dans $\mathcal{J}$ tel que

$$Z_j = \sum_{\mathbf{i}\in I(1,\mathbf{n},x,\mathbf{j})} \Delta_{\mathbf{i}}(x)$$

où $I(1,\mathbf{n},x,\mathbf{j}) = \{\mathbf{i} : 2j_k p_{\mathbf{n}} + 1 \leq i_k \leq 2j_k p_{\mathbf{n}} + p_{\mathbf{n}};\ k = 1,\ldots,N\}$. La distance séparant ces ensembles est supérieure à $p_{\mathbf{n}}^N$ et chaque ensemble contient $p_{\mathbf{n}}^N$ éléments.

Le Lemme 4.5 de Carbon et al. [3] nous permet d'approximer Z_1,Z_2, …, Z_M par des variables aléatoires indépendantes $Z_1^*,\ldots,Z_M^*$ de même loi que $Z_{j=1,\ldots,M}$ et telle que

$$\sum_{j=1}^M \mathbb{E}|Z_j - Z_j^*| \leq 2CM\big(p_{\mathbf{n}}^N\psi(M-1)p_{\mathbf{n}}^N, p_{\mathbf{n}}^N\big)\varphi(p_{\mathbf{n}}).$$

Alors, d'après les inégalités de Bernstein et de Markov, on a

$$\mathbb{P}\left(T(\mathbf{n},x,i) \geq \eta\widehat{\mathbf{n}}\mathbb{E}\left[K_{\mathbf{1}}(x)\right]\right) \leq B_1 + B_2,$$

où

$$B_1 = \mathbb{P}\left(\left|\sum_{j=1}^{M} Z_j^*\right| \geq \frac{M\eta\widehat{\mathbf{n}}\mathbb{E}\left[K_{\mathbf{1}}(x)\right]}{2M}\right) \leq$$
$$\leq 2\exp\left(-\frac{(\eta\widehat{\mathbf{n}}\mathbb{E}\left[K_{\mathbf{1}}(x)\right])^2}{M\mathrm{Var}\left[Z_1^*\right] + Cp_{\mathbf{n}}^N\eta\widehat{\mathbf{n}}\mathbb{E}\left[K_{\mathbf{1}}(x)\right]}\right),$$

$$B_2 = \mathbb{P}\left(\sum_{j=1}^{M}|Z_j - Z_j^*| \geq \frac{\eta\widehat{\mathbf{n}}\mathbb{E}\left[K_{\mathbf{1}}(x)\right]}{2}\right) \leq \frac{C}{\eta\widehat{\mathbf{n}}\mathbb{E}\left[K_{\mathbf{1}}(x)\right]}\sum_{j=1}^{M}\mathbb{E}|Z_j - Z_j^*| \leq$$
$$\leq CMp_{\mathbf{n}}^N(\eta\widehat{\mathbf{n}}\mathbb{E}\left[K_{\mathbf{1}}(x)\right])^{-1}\psi((M-1)p_{\mathbf{n}}^N, p_{\mathbf{n}}^N)\varphi(p_{\mathbf{n}}).$$

Comme $\widehat{\mathbf{n}} = 2^N Mp_{\mathbf{n}}^N$ et $\psi((M-1)p_{\mathbf{n}}^N, p_{\mathbf{n}}^N) \leq p_{\mathbf{n}}^N$ alors pour $\eta = \eta_0\sqrt{\frac{\log\widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\,\phi_x(h_K)}}$

$$B_2 \leq \widehat{\mathbf{n}}p_{\mathbf{n}}^N\left(\log\widehat{\mathbf{n}}\right)^{-1/2}\left(\widehat{\mathbf{n}}\phi_x(h_K)\right)^{-1/2}\varphi(p_{\mathbf{n}}).$$

On prend $p_{\mathbf{n}}C\left(\frac{\widehat{\mathbf{n}}\phi_x(h_K)}{\log\widehat{\mathbf{n}}}\right)^{1/2N}$, alors

$$B_2 \leq \left(\log\widehat{\mathbf{n}}\right)^{-1}\widehat{\mathbf{n}}\,\varphi(p_{\mathbf{n}}). \tag{7}$$

D'après (H7), on montre que $\sum_{\mathbf{n}\in I_{\mathbf{n}}}\widehat{\mathbf{n}}\,\varphi(p_{\mathbf{n}}) < \infty$.

On va traiter, maintenant, B_1. Pour cela, il suffit d'évaluer $\mathrm{Var}\left[Z_1^*\right]$. En effet,

$$\mathrm{Var}\left[Z_1^*\right] = \mathrm{Var}\left[\sum_{\mathbf{i}\in I(1,\mathbf{n},x,\mathbf{1})}\Delta_{\mathbf{i}}(x)\right] = \sum_{\mathbf{i},\mathbf{j}\in I(1,\mathbf{n},x,\mathbf{1})}\left|\mathrm{Cov}(\Delta_{\mathbf{i}}(x), \Delta_{\mathbf{j}}(x))\right|.$$

Posons $Q_{\mathbf{n}} = \sum_{\mathbf{i}\in I(1,\mathbf{n},x,\mathbf{1})}\mathrm{Var}\left[\Delta_{\mathbf{i}}(x)\right]$ et $R_{\mathbf{n}} = \sum_{\mathbf{i}\neq\mathbf{j}\in I(1,\mathbf{n},x,\mathbf{1})}|\mathrm{Cov}(\Delta_{\mathbf{i}}(x), \Delta_{\mathbf{j}}(x))|$. En vertu de (H1), on a $\mathrm{Var}[\Delta_{\mathbf{i}}(x)] \leq C(\phi_x(h_K) + (\phi_x(h_K))^2)$, donc $Q_{\mathbf{n}} = O\big(p_{\mathbf{n}}^N\phi_x(h_K)\big)$. En ce qui concerne $R_{\mathbf{n}}$, on utilise les techniques de Masry [18] et on considère les ensembles

$$S_1 = \{\mathbf{i}, \mathbf{j} \in I(1, \mathbf{n}, x, \mathbf{1}) : 0 < \|\mathbf{i} - \mathbf{j}\| \leq c_{\mathbf{n}}\},$$
$$S_2 = \{\mathbf{i}, \mathbf{j} \in I(1, \mathbf{n}, x, \mathbf{1}) : \|\mathbf{i} - \mathbf{j}\| > c_{\mathbf{n}}\},$$

où $c_{\mathbf{n}}$ est une suite réelle tendant vers $+\infty$. Ainsi,

$$R_{\mathbf{n}} = \sum_{(\mathbf{i},\mathbf{j})\in S_1}\left|\mathrm{Cov}\left(\Delta_{\mathbf{i}}(x), \Delta_{\mathbf{j}}(x)\right)\right| + \sum_{(\mathbf{i},\mathbf{j})\in S_2}\left|\mathrm{Cov}\left(\Delta_{\mathbf{i}}(x), \Delta_{\mathbf{j}}(x)\right)\right| = R_{\mathbf{n}}^1 + R_{\mathbf{n}}^2.$$

D' une part, on a

$$R_{\mathbf{n}}^1 = \sum_{(\mathbf{i},\mathbf{j})\in S_1} \left|\mathbb{E}\left[K_{\mathbf{i}}(x)K_{\mathbf{j}}(x)\right] - \mathbb{E}\left[K_{\mathbf{i}}(x)\right]\mathbb{E}\left[K_{\mathbf{j}}(x)\right]\right| \leq$$

$$\leq Cp_{\mathbf{n}}^N c_{\mathbf{n}}^N \phi_x(h_K)\left((\phi_x(h_K))^{1/a} + \phi_x(h_K)\right) \leq Cp_{\mathbf{n}}^N c_{\mathbf{n}}^N \phi_x(h_K)^{(a+1)/a}.$$

D'autre part, on a $R_{\mathbf{n}}^2 = \sum_{(\mathbf{i},\mathbf{j})\in S_2} \left|\mathrm{Cov}\left(\Delta_{\mathbf{i}}(x), \Delta_{\mathbf{j}}(x)\right)\right|$. Comme le noyau K est borné, du Lemme 2.1(ii) de Tran [21] on obtient

$$\left|\mathrm{Cov}\left(\Delta_{\mathbf{i}}(x), \Delta_{\mathbf{j}}(x)\right)\right| \leq C\varphi\left(\|\mathbf{i}-\mathbf{j}\|\right).$$

Ainsi

$$R_{\mathbf{n}}^2 \leq C \sum_{(\mathbf{i},\mathbf{j})\in S_2} \varphi\left(\|\mathbf{i}-\mathbf{j}\|\right) \leq Cp_{\mathbf{n}}^N \sum_{\mathbf{i}:\|\mathbf{i}\|\geq c_{\mathbf{n}}} \varphi\left(\|\mathbf{i}\|\right) \leq Cp_{\mathbf{n}}^N c_{\mathbf{n}}^{-Na} \sum_{\mathbf{i}:\|\mathbf{i}\|\geq c_{\mathbf{n}}} \|\mathbf{i}\|^{Na}\varphi\left(\|\mathbf{i}\|\right).$$

On prend $c_{\mathbf{n}} = (\phi_x(h_K))^{-1/Na}$, alors

$$R_{\mathbf{n}}^2 \leq Cp_{\mathbf{n}}^N c_{\mathbf{n}}^{-Na} \sum_{\mathbf{i}:\|\mathbf{i}\|\geq c_{\mathbf{n}}} \|\mathbf{i}\|^{Na}\varphi\left(\|\mathbf{i}\|\right) \leq Cp_{\mathbf{n}}^N \phi_x(h_K) \sum_{\mathbf{i}:\|\mathbf{i}\|\geq c_{\mathbf{n}}} \|\mathbf{i}\|^{Na}\varphi\left(\|\mathbf{i}\|\right).$$

D'après (H2) on peut écrire $R_{\mathbf{n}}^2 \leq Cp_{\mathbf{n}}^N \phi_x(h_K)$. De plus, avec ce choix de $c_{\mathbf{n}}$ on a

$$R_{\mathbf{n}}^1 \leq Cp_{\mathbf{n}}^N \phi_x(h_K).$$

D'où

$$\mathrm{Var}\left[Z_1^*\right] = O\left(p_{\mathbf{n}}^N \phi_x(h_K)\right).$$

En utilisant ce dernier résultat, avec les définitions de $p_{\mathbf{n}}$, M et η, on montre

$$B_1 \leq \exp\left(-C\eta_0 \log\widehat{\mathbf{n}}\right).$$

Par conséquent, un choix approprié de η_0 permet de déduire le résultat du Lemme. □

Démonstration du Corollaire 3. On a, $\mathbb{E}\left[\widehat{F}_D^x(x)\right] = 1$, donc

$$\mathbb{P}\left(\left|\widehat{F}_D^x(x)\right| \leq 1/2\right) \leq \mathbb{P}\left(\left|\widehat{F}_D^x(x) - \mathbb{E}\left[\widehat{F}_D^x(x)\right]\right| > 1/2\right)$$

et Corollaire 3 devient est une conséquence du Lemme 2. □

Démonstration du Lemme 4. Du fait que les variables aléatoires sont identiquement distribuées, on a

$$\forall y \in \mathcal{S}, \quad F^x(y) - \mathbb{E}\widehat{F}_N^x(y) = \mathbb{E}\left[K_1(x)\mathbb{1}_{B(x,h_K)}(X)\left[\mathbb{E}\left[H_1(y)/X\right] - F^x(y)\right]\right].$$

En intégrant par parties, on montre que

$$\mathbb{E}\left(H_1(y)/X\right) = \int_{\mathbb{R}} H(h_H^{-1}(y-z))f^X(z)\mathrm{d}z = h_H^{-1}\int_{\mathbb{R}} H'(h_H^{-1}(y-z))F^X(z)\mathrm{d}z.$$

En considérant le changement de variable usuel $t = \frac{y-z}{h_H}$ on obtient

$$\mathbb{E}\left(H_1(y)/X\right) = \int_{\mathbb{R}} H'(y) F^X(y - h_H t) \mathrm{d}t,$$

donc

$$|\mathbb{E}(H_1(y)/X) - F^x(y)| \leq \int_{\mathbb{R}} H'(y) |F^X(y - h_H t) - F^x(y)| \mathrm{d}t.$$

Sous (H4), on a

$$\forall y \in \mathcal{S}, \quad \mathbb{1}_{B(x,h_K)}(X) |\mathbb{E}(H_1(y)/X) - F^x(y)| \leq \int_{\mathbb{R}} H'(y) \big(h_K^{b_1} + |t|^{b_2} h_H^{b_2}\big) \mathrm{d}t.$$

Comme H' est une densité de probabilité, alors, l'hypothèse (H6) achève la démonstration du Lemme 4. $\square$

Démonstration du Lemme 5. On considère le recouvrement

$$\mathcal{S} \subset \bigcup_{k=1}^{z_{\mathbf{n}}} (t_j - l_{\mathbf{n}}, t_j + l_{\mathbf{n}})$$

avec $l_{\mathbf{n}} = \widehat{\mathbf{n}}^{-\alpha-\frac{1}{2}}$ et $z_{\mathbf{n}} \leq \widehat{\mathbf{n}}^{\alpha+\frac{1}{2}}$. En posant, $j(y) = \arg\min_{j\in\{1,2,\ldots,z_{\mathbf{n}}\}} |y - t_j|$, on obtient la décomposition

$$\left|\widehat{F}_N^x(y) - \mathbb{E}\widehat{F}_N^x(y)\right| \leq$$

$$\leq \underbrace{\left|\widehat{F}_N^x(y) - \widehat{F}_N^x(t_{j(y)})\right|}_{T_1} + \underbrace{\left|\widehat{F}_N^x(t_{j(y)}) - \mathbb{E}\widehat{F}_N^x(t_{j(y)})\right|}_{T_2} + \underbrace{\left|\mathbb{E}\widehat{F}_N^x(t_{j(y)}) - \mathbb{E}\widehat{F}_N^x(y)\right|}_{T_3}.$$

• En ce qui concerne (T_1) et (T_3) on utilise le fait que H' est bornée pour démonter

$$\left|\widehat{F}_N^x(y) - \widehat{F}_N^x(t_{j(y)})\right| \leq \frac{1}{\widehat{\mathbf{n}}\mathbb{E}\left[K_{\mathbf{1}}(x)\right]} \sum_{\mathbf{i}\in I_{\mathbf{n}}} K_{\mathbf{i}}(x) \left|H_{\mathbf{i}}(y) - H_{\mathbf{i}}(t_{j(y)})\right| \leq$$

$$\leq C \frac{1}{\widehat{\mathbf{n}}\mathbb{E}\left[K_{\mathbf{1}}(x)\right]} \sum_{\mathbf{i}\in I_{\mathbf{n}}} K_{\mathbf{i}}(x) \left|H_{\mathbf{i}}(y) - H_{\mathbf{i}}(t_{j(y)})\right| \leq C \frac{l_{\mathbf{n}}}{h_H} \widehat{F}_D^x.$$

De la définition de $l_{\mathbf{n}}$ et sous l'hypothèse (H7), on a $\frac{l_{\mathbf{n}}}{h_H} = o\left(\sqrt{\frac{\log \widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\,\phi(h_K)}}\right)$. Ainsi, la convergence presque complète de $\widehat{F}_D^x$, nous permet d'écrire

$$\left|\widehat{F}_N^x(y) - \widehat{F}_N^x(t_{j(y)})\right| = o\left(\sqrt{\frac{\log \widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\,\phi(h_K)}}\right), \quad \text{p.co.} \tag{8}$$

et

$$\left|\mathbb{E}\widehat{F}_N^x(t_{j(y)}) - \mathbb{E}\widehat{F}_N^x(y)\right| = o\left(\sqrt{\frac{\log \widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\,\phi(h_K)}}\right). \tag{9}$$

- Pour (T_3), $\forall\,\eta > 0$, on a

$$\mathbb{P}\left(\sup_{y\in\mathcal{S}}\left|\widehat{F}_N^x(t_{j(y)}) - \mathbb{E}\widehat{F}_N^x(t_{j(y)})\right| > \eta\sqrt{\frac{\log \widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\,\phi(h_K)}}\right) =$$

$$= \mathbb{P}\left(\max_{j\in\{1,2,\ldots,z_{\mathbf{n}}\}}\left|\widehat{F}_N^{x_k}(t_j) - \mathbb{E}\widehat{F}_N^{x_k}(t_j)\right| > \eta\sqrt{\frac{\log \widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\,\phi(h_K)}}\right) \leq$$

$$\leq z_{\mathbf{n}} \max_{j\in\{1,2,\ldots,z_{\mathbf{n}}\}} \mathbb{P}\left(\left|\widehat{F}_N^{x_k}(t_j) - \mathbb{E}\widehat{F}_N^{x_k}(t_j)\right| > \eta\sqrt{\frac{\log \widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\,\phi(h_K)}}\right).$$

Posons

$$\Delta'_{\mathbf{i}} = K_{\mathbf{i}}(x)H_{\mathbf{i}}(t_j) - \mathbb{E}\left[K_{\mathbf{i}}(x)H_{\mathbf{i}}(t_j)\right].$$

Il est clair que

$$\begin{cases} |\Delta'_{\mathbf{i}}| \leq C\,|\Delta_{\mathbf{i}}(x)|\,, \\ \mathrm{Var}\left[\Delta'_{\mathbf{i}}\right] \leq C\mathrm{Var}\left[\Delta_{\mathbf{i}}(x)\right] \text{ et} \\ \mathrm{Cov}(\Delta'_{\mathbf{i}}, \Delta'_{\mathbf{j}}) \leq C\mathrm{Cov}(\Delta_{\mathbf{i}}(x), \Delta_{\mathbf{j}}(x)) \quad \forall\,\mathbf{i} \neq \mathbf{j}. \end{cases}$$

En utilisant les mêmes arguments que dans le Lemme 2, on montre

$$\mathbb{P}\left(\max_{z\in\mathcal{G}_n}\left|\widehat{F}_N^x(z) - \mathbb{E}\big[\widehat{F}_N^x(z)\big]\right| > \eta_0\sqrt{\frac{\log \widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\,\phi_x(h_K)}}\right) \leq z_{\mathbf{n}}\,(B_1 + B_2)\,.$$

La définition de $z_{\mathbf{n}}$, l'équation (7) et l'hypothèse (H7) nous permettent d'écrire

$$\sum_{\mathbf{n}\in I_{\mathbf{n}}} \widehat{\mathbf{n}}^{\alpha+\frac{1}{2}} B_2 < \infty.$$

Tandis qu'un choix convenable de η_0 donne

$$\sum_{\mathbf{n}\in I_{\mathbf{n}}} \widehat{\mathbf{n}}^{\alpha+\frac{1}{2}} B_1 < \infty.$$

Finalement, le Lemme 5 est une conséquence de ces deux derniers résultats combinés avec les équations (8) et (9). $\square$

Démonstration du Lemme 6. En utilisant la stationnarité des observations, le conditionnement par la variable explicative et le changement de variable usuel $t = \frac{y-z}{h_H}$, on obtient

$$h_H^{-1}\mathbb{E}\left(H'_1(y)/X\right) = h_H^{-1}\int_{\mathbb{R}} H'(y) f^X(y - h_H t)\mathrm{d}t,$$

et on en déduit

$$|h_H^{-1}\mathbb{E}(H_1'(y)/X) - f^x(y)| \leq \int_{\mathbb{R}} H'(y)|f^X(y - h_H t) - f^x(y)|\mathrm{d}t.$$

Sous la deuxième partie de (H4)

$$\forall y \in \mathcal{S}, \quad \mathbb{I}_{B(x,h_K)}(X)|h_H^{-1}\mathbb{E}(H_1'(y)/X) - f^x(y)| \leq \int_{\mathbb{R}} H'(y)(h_K^{b_3} + |t|^{b_4} h_H^{b_2})\mathrm{d}t.$$

L'hypothèse (H6) achève la démonstration du Lemme 6. $\square$

Démonstration du Lemme 7. La démonstration est très similaire à celle du Lemme 5. En effet, on considère le recouvrement $\mathcal{S} \subset \bigcup_{k=1}^{z_{\mathbf{n}}}(t_j - l_{\mathbf{n}}, t_j + l_{\mathbf{n}})$ avec $l_{\mathbf{n}} = \widehat{\mathbf{n}}^{-\frac{3}{2}\alpha - \frac{1}{2}}$ et $z_{\mathbf{n}} \leq \widehat{\mathbf{n}}^{\frac{3}{2}\alpha + \frac{1}{2}}$. En posant $j(y) = \arg\min_{j\in\{1,2,\ldots,z_{\mathbf{n}}\}} |y - t_j|$, on obtient la décomposition

$$\begin{gathered}\left|\widehat{f}_N^x(y) - \mathbb{E}\widehat{f}_N^x(y)\right| \leq \\ \leq \underbrace{\left|\widehat{f}_N^x(y) - \widehat{f}_N^x(t_{j(y)})\right|}_{S_1} + \underbrace{\left|\widehat{f}_N^x(t_{j(y)}) - \mathbb{E}\widehat{f}_N^x(t_{j(y)})\right|}_{S_2} + \underbrace{\left|\mathbb{E}\widehat{f}_N^x(t_{j(y)}) - \mathbb{E}\widehat{f}_N^x(y)\right|}_{S_3}.\end{gathered}$$

• Étude de S_1 et S_3. De même que pour T_1 et T_3 dans le Lemme 5, on utilise le fait que H est une fonction de classe C^2 et à support compact. On a

$$\begin{aligned}\left|\widehat{f}_N^x(y) - \widehat{f}_N^x(t_{j(y)})\right| &\leq \frac{h_H^{-1}}{\widehat{\mathbf{n}}\mathbb{E}\left[K_{\mathbf{1}}(x)\right]} \sum_{\mathbf{i}\in I_{\mathbf{n}}} K_{\mathbf{i}}(x)\left|H_{\mathbf{i}}'(y) - H_{\mathbf{i}}'(t_{j(y)})\right| \leq \\ &\leq C\frac{h_H^{-1}}{\widehat{\mathbf{n}}\mathbb{E}\left[K_{\mathbf{1}}(x)\right]} \sum_{\mathbf{i}\in I_{\mathbf{n}}} K_{\mathbf{i}}(x)\left|H_{\mathbf{i}}'(y) - H_{\mathbf{i}}'(t_{j(y)})\right| \leq C\frac{l_{\mathbf{n}}}{h_H^2}\widehat{F}_D^x.\end{aligned}$$

La définition de $l_{\mathbf{n}}$ et l'hypothèse (H7) nous permet d'écrire $\frac{l_{\mathbf{n}}}{h_H^2} o\left(\sqrt{\frac{\log\widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\, h_H\phi(h_K)}}\right)$. La convergence presque complète de $\widehat{F}_D^x$ implique

$$\left|\widehat{f}_N^x(y) - \widehat{f}_N^x(t_{j(y)})\right| = o\left(\sqrt{\frac{\log\widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\, h_H\phi(h_K)}}\right), \quad \text{p.co.} \tag{10}$$

$$\left|\mathbb{E}\widehat{f}_N^x(t_{j(y)}) - \mathbb{E}\widehat{f}_N^x(y)\right| = o\left(\sqrt{\frac{\log\widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\, h_H\phi(h_K)}}\right). \tag{11}$$

• En ce qui concerne (S_2), on suit les mêmes étapes de T_3; $\forall\,\eta>0$, on a

$$\mathbb{P}\left(\sup_{y\in\mathcal{S}}\left|\widehat{f}_N^x(t_{j(y)})-\mathbb{E}\widehat{f}_N^x(t_{j(y)})\right|>\eta\sqrt{\frac{\log\widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\,h_H\phi(h_K)}}\right)=$$

$$=\mathbb{P}\left(\max_{j\in\{1,2,\ldots,z_{\mathbf{n}}\}}\left|\widehat{f}_N^{x_k}(t_j)-\mathbb{E}\widehat{f}_N^{x_k}(t_j)\right|>\eta\sqrt{\frac{\log\widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\,h_H\phi(h_K)}}\right)\leq$$

$$\leq z_{\mathbf{n}}\max_{j\in\{1,2,\ldots,z_{\mathbf{n}}\}}\mathbb{P}\left(\left|\widehat{f}_N^{x_k}(t_j)-\mathbb{E}\widehat{f}_N^{x_k}(t_j)\right|>\eta\sqrt{\frac{\log\widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\,h_H\phi(h_K)}}\right).$$

Posons

$$\Delta_{\mathbf{i}}''=K_{\mathbf{i}}(x)H_{\mathbf{i}}'(t_j)-\mathbb{E}\left[K_{\mathbf{i}}(x)H_{\mathbf{i}}'(t_j)\right].$$

Alors

$$\begin{cases}|\Delta_{\mathbf{i}}''|\leq C\,|\Delta_{\mathbf{i}}(x)|\,,\\ \operatorname{Var}[\Delta_{\mathbf{i}}'']=O(h_H\phi(h_K))\text{ et}\\ \mathbb{E}\big[H_{\mathbf{i}}'H_{\mathbf{j}}'|(X_{\mathbf{i}},X_{\mathbf{j}})\big]=O(h_H^2)\quad\forall\,\mathbf{i}\neq\mathbf{j}.\end{cases}$$

En utilisant les mêmes étapes du Lemme 2, avec une légère modification dans le choix de $c_{\mathbf{n}}$, on montre que

$$\operatorname{Cov}(\Delta_{\mathbf{i}}',\Delta_{\mathbf{j}}')=O(h_H\phi(h_K))\quad\forall\,\mathbf{i}\neq\mathbf{j}.$$

Ainsi,

$$\mathbb{P}\left(\max_{z\in\mathcal{G}_n}\left|\widehat{f}_N^x(z)-\mathbb{E}\big[\widehat{f}_N^x(z)\big]\right|>\eta_0\sqrt{\frac{\log\widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\,h_H\phi_x(h_K)}}\right)\leq z_{\mathbf{n}}\left(B_1'+B_2'\right)$$

où $B_1'\leq\exp\left(-C(\eta_0)\log\widehat{\mathbf{n}}\right)$ et

$$B_2'\leq\widehat{\mathbf{n}}p_{\mathbf{n}}^N\left(\log\widehat{\mathbf{n}}\right)^{-1/2}\left(\widehat{\mathbf{n}}h_H\phi_x(h_K)\right)^{-1/2}\varphi(p_{\mathbf{n}}).$$

Prenons $p_{\mathbf{n}}C\left(\frac{\widehat{\mathbf{n}}h_H\phi_x(h_K)}{\log\widehat{\mathbf{n}}}\right)^{1/2N}$, la définition de $z_{\mathbf{n}}$ sous l'hypothèse (H7), un choix convenable de η_0 nous permettent d'écrire

$$\sum_{\mathbf{n}\in I_{\mathbf{n}}}\widehat{\mathbf{n}}^{\frac{3}{2}\alpha+\frac{1}{2}}(B_2'+B_1')<\infty.$$

Par conséquent, le Lemme 7 est une conséquence de ce dernier résultat combiné avec les équations (10) et (11). □

Preuve du Lemme 8. Il suffit de remarquer que

$$\inf_{y\in\mathcal{S}}|1-\widehat{F}^x(y)|\leq\Big(1-\sup_{y\in\mathcal{S}}F^x(y)\Big)/2\Rightarrow\sup_{y\in\mathcal{S}}|\widehat{F}^x(y)-F^x(y)|\geq\Big(1-\sup_{y\in\mathcal{S}}F^x(y)\Big)/2.$$

En termes de probabilité, on obtient

$$\mathbb{P}\left\{\inf_{y\in\mathcal{S}}|1-\widehat{F}^x(y)|<\Big(1-\sup_{y\in\mathcal{S}}F^x(y)\Big)/2\right\}\leq$$
$$\leq\mathbb{P}\left\{\sup_{y\in\mathcal{S}}|\widehat{F}^x(y)-F^x(y)|\geq\Big(1-\sup_{y\in\mathcal{S}}F^x(y)\Big)/2\right\}<\infty.$$

Il suffit, donc, de prendre $\eta=\Big(1-\sup_{y\in\mathcal{S}}F^x(y)\Big)/2$ et appliquer les résultats des Lemmes 2–5 pour achever la démonstration du Lemme 8. $\square$

Démonstration du Corollaire 9. Sous l'hypothèse (5), le développement de Taylor de la fonction h^x au voisinage de $\theta(x)$, en particulier pour le point $\widehat{\theta}(x)$, est

$$h^x(\widehat{\theta}(x))=h^x(\theta(x))+(\widehat{\theta}(x)-\theta(x))^2\frac{h^{x''}(\theta^*)}{j!},$$

où θ^* est entre $\widehat{\theta}(x)$ et $\theta(x)$, il est donc nécessairement dans le compact $\mathcal{S}$. Ainsi

$$|\widehat{\theta}(x)-\theta(x)|^2\leq\frac{2!}{\min_{y\in\mathcal{S}}h^{x''}(y)}|h^x(\widehat{\theta}(x))-h^x(\theta(x))|$$

et à l'aide d'arguments analytiques on montre que

$$|h^x(\widehat{\theta}(x))-h^x(\theta(x))|\leq 2\sup_{y\in\mathcal{S}}|\widehat{h}^x(y)-h^x(y)|.$$

D'où

$$|\widehat{\theta}(x)-\theta(x)|^2\leq\frac{C}{\min_{y\in\mathcal{S}}h^{x''}(y)}\ \sup_{y\in\mathcal{S}}|\widehat{h}^x(y)-h^x(y)|.$$

En utilisant le résultat du Théorème 1, on montre que

$$|\widehat{\theta}(x)-\theta(x)|^2=O\left(h_K^{b_1}+h_H^{b_2}\right)+O\left(\left(\frac{\log\widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\,h_H\phi_x(h_K)}\right)^{\frac{1}{2}}\right).\quad\square$$

Remerciements. Le rapporteur anonyme est vivement remercié pour ses commentaires pertinents.

REFERENCES

[1] G. Biau and B. Cadre, *Nonparametric spatial prediction.* Stat. Inference Stoch. Process **7** (2004), 327–349.

[2] D. Bosq, *Linear Processes in Function Spaces: Theory and Applications.* Lecture Notes in Statistics **149**, Springer, 2000.

[3] M. Carbon, L.T. Tran and B. Wu, *Kernel density estimation for random fields.* Statist. Probab. Lett. **36** (1997), 115–125.
[4] M. Carbon, C. Francq and L.T. Tran, *Kernel regression estimation for random fields.* J. Statist. Plann. Inference **137** (2007), 778–798.
[5] N.A. Cressie, *Statistics for Spatial Data.* Wiley, New York, 1991.
[6] S. Dabo-Niang and A.F. Yao, *Kernel regression estimation for continuous spatial processes.* Math. Methods Statist **16** (2007), 298–317.
[7] G. Estévez-Pérez, A. Quintela-del-Ro and Ph. Vieu, *Convergence rate for cross-validatory bandwidth in kernel hazard estimation from dependent samples.* J. Statist. Plann. Inference **104** (2002), 1–30.
[8] F. Ferraty, A. Laksaci and Ph. Vieu, *Estimating some characteristics of the conditional distribution in nonparametric functional models.* Stat. Inference Stoch. Process **9** (2006), 47–76.
[9] F. Ferraty, A. Rabhi and Ph. Vieu, *Estimation non-paramétrique de la fonction de hasard avec variable explicative fonctionnelle.* Rev. Roumaine Math. Pures Appl. **53** (2008), 1–18.
[10] F. Ferraty, A. Tadj, A. Laksaci and Ph. Vieu, *Rate of uniform consistency for nonparametric estimates with functional variables.* J. Statist. Plann. Inference. (accepté)
[11] F. Ferraty and Ph. Vieu, *Nonparametric Functional Data Analysis.* Springer Series in Statistics. Springer, New York, 2006.
[12] X, Guyon, *Estimation d'un champ par pseudo-vraisemblance conditionnelle: Etude asymptotique et application au cas Markovien.* Proceedings of the Sixth Franco-Belgian Meeting of Statisticians, 1987.
[13] A. Quintela-del-Ro, *Nonparametric estimation of the maximum hazard under dependence conditions.* Statist. Probab. Lett. **76** (2006), 1117–1124.
[14] A. Quintela-del-Ro, *Hazard function given a functional variable: Non-parametric estimation under strong mixing conditions.* J. Nonparametr. Stat. **20** (2008), 413–430.
[15] J.P. Lecoutre and E. Ould-Said, *Estimation de la densité et de la fonction de hasard conditionnelle pour un processus fortement mélangeant avec censure.* C.R. Math. Acad. Sci. Paris **314** (1992), 295–300.
[16] J. Li and L.T. Tran, *Hazard rate estimation on random fields.* J. Multivariate Anal. **98** (2007), 1337–1355.
[17] J.Li and L.T. Tran, *Nonparametric estimation of conditional expectation.* J. Statist. Plann. Inference **139** (2009), 164–175.
[18] E. Masry, *Recursive probability density estimation for weakly dependent stationary processes.* IEEE Trans. Inform. Theory **32** (1986), 254–267.
[19] J.O. Ramsay and B.W. Silverman, *Applied Functional Data Analysis; Methods and Case Studies.* Springer-Verlag, New York, 2002.
[20] G. Roussas, *Hazard rate estimation under dependence conditions.* J. Statist. Plann. Inference **22** (1989), 81–93.
[21] L.T. Tran, *Kernel density estimation on random fields.* J. Multivariate Anal. **34** (1990), 37–53.
[22] G.S. Watson and M.R. Leadbetter, *Hazard analysis.* Sankhyia **26** (1964), 101–116.

Received 29 March 2009

Université Djilali Liabes
Laboratoire de Mathématiques
BP 89, Sidi Bel Abbes, 22000, Algérie
alilak@yahoo.fr