République Algérienne Démocratique et Populaire

Ministère de l'enseignement supérieure et de la recherche scientifique

N° Attribué par la bibliothèque

Année univ.: 2016/2017

Estimation non paramétrique de la fonction de Hasard avec variable explicative fonctionnelle

Master Académique

Université Dr Tahar Moulay - Saïda

Discipline : MATHEMATIQUES

Spécialité : analyse stochastique, statistique des processus et applications

par

Sehibi **F**atima Zohra [1]

Sous la direction de

Dr N. Hchemi

Soutenue le 21 Mai 2017 devant le jury composé de

M. Mmmmmmm	Université Dr Tahar Moulay - Saïda	Président
M. Mmmmmmm	Université Dr Tahar Moulay - Saïda	Rapporteur
M. Mmmmmmm	Université Dr Tahar Moulay - Saïda	Examinateur
M. Mmmmmmm	Université Dr Tahar Moulay - Saïda	Examinateur

1. e-mail : azerty@gmail.com

Dedicace

Je dédis ce mémoire :

– $\mathcal{A}$ la mémoir de mon père.

– $\mathcal{A}$ ma trés chère mère, qui m'a encouragé à aller de l'avant et qui m'a donné tous son amour pour reprendre mes étude.

– $\mathcal{A}$ ma trés chère sœur Leila, son mari et leurs enfants(Hadile, Houcine et Ibtihal).

– $\mathcal{A}$ mon chèr frère que dieu t'assistes.

– $\mathcal{A}$ ma petite sœur chèrie Sara, je te souhaite une bonne continuat.

– $\mathcal{A}$ tout mes oncles et tantes qui m'ont toujours soutenue et époulée dans la vie.

– $\mathcal{A}$ tout mes cousins.

– $\mathcal{A}$ mes chères amies Naima, Hadjer, Nassima, Noura, Imen, Khadidja, Bouchera.

– $\mathcal{A}$ ma chère collègues Kaouther.

– $\mathcal{A}$ tout qui prend une place de mon cœur.

Remerciments

– Nous remercions Dieu, le tout puissant, le miséricordieux, qui nous a donné l'opportunité de mener à bien ce travail.

– Nous voudrons tout d'abord adresser toute nos gratitude à notre encadreur : N.Hachemi pour sa patience, sa disponibilité et surtout ses judicieux conseils, qui ont contribué à alimenter notre réflexion.

– Nos vifs remerciements vont également aux membres du jury pour l'intérêt qu'ils ont porté à notre recherche en acceptant d'examiner notre travail et de l'enrichir par leurs propositions.

– Nous désirons aussi remercier nos professeurs qui nous ont enseigné durant nos études.

– Nous remercions nos amis et camarades de promotion pour ces cinq années passées ensemble, dans les meilleurs moments comme dans les pires.

– A la fin il nous est agréable d'adresser nos vifs remerciments : à tous ceux qui nous ont aidé de prés ou loin à elaborer cet ouvrage.

Table des matières

Introduction générale

La modèlisation statistique de données fonctionnelles est actuellement au coeur d'une dynamique qui touche de nombreux statisiens, qu'ils soient théorique ou appliquées. L'engouement de cette thématique est justifié par la potentialité en termes d'applications qu'elle offre. En effet, dans de nombreux domaines les observations sont représentées par des variables aléatoires fonctionnelles. A ce sujet, l'estimation de la fonction de Hasard joue un role cruciale, de part la variété de ses possibilités d'applications, est une question importante en statistique. Ce sujet peut être abordé sous plusieurs angles selon la complexité du problème posé : présence éventuelle de censure dans l'èchantillon observé (phénomène conrant dans les applications médicales par exemple), présence éventuelle de dépendance entre les variables observées ou bien présence de variables explicatives. De nombreuses techniques pour traiter de cas différentes situation mais toutes ne traitent que de variables alèatoires explicatives réelles ou multidimensinnelles. Ainsi, l'estimation d'un taux de hasard en pésence de variable explicative fonctionnelle est une question d'actualité à laquelle ce travail propose d'apporter un élément de réponse. Comme toute étude asymptotique en statistique non paramétrique fonctionnelle la difficulté technique se situe dans le fait de la non-existence de la mesure de Lebesgue dans les espaces de dimension infinie. Cette étude met en évidence la propriété de concentration de la mesure de probabilité de la variable fonctionnelle considérée sur les petites boules. Hystoriquement, l'estimation de la fonction de hasard est très abondante, lorsque les observations sont vectorielles. Citons, par exemple, Watson et Leadbetter [7], Roussas [8], Lecoutre et Ould-Said [11], Estèvez et al[9] et Quintela-del-Rio [14] pour des références récentes. Dans tous ces travaux les auteurs considérent des observations indépendantes ou des données dépendantes issues de séries temporelles. Dans le cas spatial, la littérature est trés restreinte, il n'y a qu'un seul travail de Li et Tran [12]. 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 [5]. 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 été obtenue, dans le cas α-mélangeant, par Quintela-del-Rio [14]. Nous renvoyons à Ferraty et al [3]. pour la convergence presque complète uniforme sur la composante fonctionnelle de ce modèle non paramétrique. Ferraty et al [6] ont étudiée le cas des observations spatialement dépendantes La littérature sur cette estimation est relativement restreinte en statistique fonctionnelle. Ezzahrioui [2] (2007) a étudié la normalité asymptotique. Le cas α-mélangeant a été traité par Quintela-Del-Rio [14] (2010). Ce dernier a établi la convergence presque complète et la normalité asymptotique de l'estimateur proposé par Ferraty et al. (2008). Ferraty et al [5] (2008) ont établi la convergence presque complète d'un estimateur à noyau de la fonction de hasard conditionnelle, lorsque les observations sont indépendantes et identiquement distribuées. Laksaci et Mechab [1] (2010) sur l'estimation de la fonction de hasard conditionnelle pour des données fonctionnelles spatialement dépendantes.

Le présent document est divisé en trois chapitres et il considère les deux cas des observations sont indépendantes et identiquement distribuées et le mélange fort. Le premier chapitre est consacré à l'introduction des définitions et outils techniques utilisées pour l'élaboration de nos résultats. En particulier, nous rappelons la définition de processus de mélange, les définitions de différents modes de convergence. On trouvera, aussi, dans ce chapitre un nombre trés important d'outils nécessaires pour l'élaboration des nos résultats. Dans le deuxième chapitre, on présente le problème de prévision d'une manière générale et on montre que l'approche non paramétrique devient naturelle, notamment, lorsqu'on ne dispose pas a priori d'aucune information sur la loi des observations. Dans ce Chapitre, on a estimé la fonction de hasard conditionnelle, par la méthode du noyau d'une réponse réelle conditionnellement à une variable fonctionnelle lorsque les données sont complètes et nous avons démontré convergence presque complète uniforme sur un compact fixe lorsque les observations sont fonctionnelles et indépendant. On trouve aussi la convergence uniforme lorsque les données sont imcomplètes dans le chapitre trois.

Chapitre 1

Notations et définitions

Soient $(\Omega, \mathcal{A}, P)$ un espace de probabilité et $\{\Delta_i\}_{i\in\mathcal{Z}}$ une famille des variables aléatoires définie sur $(\Omega, \mathcal{A}, P)$ à valeurs dans un espace probabilisable $(\mathbb{E}, \xi)$. On note $(\sigma_i^j)_{i\neq j}$ dans $\mathbb{Z}\cup \{-\infty, +\infty\}$, la tribu engendrée par $\{\Delta_k, i < k < j\}$ et par $L_2(\sigma_i^j)$ l'espace des variables aléatoires σ_i^j- mesurable et de carrée sommable.

Définition 1.0.1 *Soit* $\{\Delta_i\}_{i\in\mathcal{Z}}$ *une famille des variables aléatoires définie sur* $(\Omega, \mathcal{A}, P)$ *à valeurs dans un espace probabilisable* $(\mathbb{E}, \xi)$. *On dit que la famille* $\{\Delta_i, i \in \mathbb{Z}\ \}$ *est* α*- mélangent si la suite*

$$\alpha_n = \sup_{\left\{k\in\mathcal{Z}, A\in\sigma_{-\infty}^k, B\in\sigma_{n+k}^{+\infty}\right\}} |P(A\cap B) - P(A)P(B)|$$

tend vers 0 *quand* n *tend vers l'infinie. La suite* α_n *est appelée coefficient de mélange forte.*

Définition 1.0.2 *On dit que qu'une famille* $\{\Delta_i, i \in Z\}$ *de variables aléatoire á valeurs dans un même espace probabilisable* $(\mathbb{E},\varepsilon)$ *est algébriquement* α*-mélangeante, s'il existe deux constantes* $c\in\mathbb{R}^{*+}$ *et* $a\in\mathbb{R}^{*+}$ *telles que les coefficients de melange verifient*

$$\alpha(n) \leq cn^{-a}.$$

Définition 1.0.3 *On dit que qu'une famille* $\{\Delta_i, i \in Z\}$ *de variables aléatoire á valeurs dans un même espace probabilisable* $(\mathbb{E},\varepsilon)$ *est géometriquement* α*-mélangeante, s'il existe deux constantes* $s\in\mathbb{R}^{*+}$ *et* $t\in]0,1[$ *telles que les coefficients de melange verifient*

$$\alpha(n) \leq st^n.$$

Définition 1.0.4 *Une fonction K de $\mathbb{R}^p$ dans $\mathbb{R}$ est dite noyau d'ordre k, $k \in \mathbb{N}^*$, si :*

$$T_{i_1,\dots,i_p}(K) = 0,\ \forall (i_1,\dots,i_p) \in \mathbb{R}^p * \quad \text{vérifiant} \ \ i_j < k,\ 1 \le j \le p.$$

et

$$T_j(K) \ne 0,\ \forall j \le k.$$

où

$$T_{i_1,\dots,i_p}(K) = \int_{\mathbb{R}} u_1^{i_1},\dots,u_p^{i_p} K(u_1 \dots u_p) du_1,\dots,du_j.$$

et

$$T_j(K) = \int_{\mathbb{R}} u_j^k K(u_1 \dots u_p) du_1,\dots,du_j.$$

1.1 Outils

Lemme 1.1.1 *([4], 2004) Soit $\Delta_1,\dots,\Delta_n$ des variables aléatoires réelles centrées, indépendantes et identiquement distribuées, telles qu'il existe deux réels positifs d et δ vérifiant :*

$$|\Delta_1| \le d \ \ et \quad E\Delta_1^2 \le \delta^2.$$

Alors, pour tout $\varepsilon \in]0, \frac{\delta^2}{d}[$ on a

$$p\left[n^{-1}\left|\sum_{i=1}^n \Delta_i\right| > \varepsilon\right] \le 2e^{-\frac{n\varepsilon^2}{4\delta^2}}.$$

cette inégalité a été donnée par W.Hoeffding en ((1.1.1), [4] ,1963). Les lemmes suivants donnent les deux version de l'inégalité de Fuk Nagaeve

Lemme 1.1.2 *([4], 2004) Soit $\{\Delta_i, i\epsilon\mathbb{N}\}$ une suite des variables aléatoires réelles α-mélangeante, de coefficient de mélange α_n vérifiant :*

$$\exists c \in \mathbb{R}^{*+}, \quad a \in \mathbb{R}^{*+} \qquad \alpha(n) \le cn^{-a}$$

et si $\forall i \|\Delta_i\|_\infty < \infty$, alors, pour tout $\varepsilon > 0$ et $r > 0$, on a

$$P\left[|\sum_{k=1}^n \Delta_k| > 4\varepsilon\right] \le \left(1 + \frac{\varepsilon^2}{rS_n^2}\right)^{\frac{-r}{2}} + 2ncr^{-1}\left(\frac{2r}{\varepsilon}\right)^{a+1}. \tag{1.1}$$

Lemme 1.1.3 *([4], 2004) Soit* $\{\Delta_i, i\epsilon\mathbb{N}\}$ *une suite des variables aléatoires réelles* α*-mélangeante, de coefficient de mélange* α_n *vérifiant :*

$$\exists c \in \mathbb{R}^{*+}, \quad a \in \mathbb{R}^{*+} \qquad \alpha(n) \leq cn^{-a}$$

et si $\|\Delta_i\|_\infty < \infty, \forall i$*, alors, pour tout* $\lambda > 0$ *et* $r > 1$*, on a*

$$\begin{aligned} P\left[|\sum_{i=1}^n \Delta_i| > 4\lambda n\mathbb{E}K_1(x)\right] &= P\left[|s_n| > 4\lambda n\mathbb{E}K_1(x)\right] \\ &\leq C\left(1 + \frac{\lambda^2 n^2(\mathbb{E}K_1(x))^2}{rs^{n,l}}\right)^{\frac{-r}{2}} \\ &\quad + C\frac{n}{r}\left(\frac{r}{\lambda n\mathbb{E}K_1(x)}\right)^{\frac{p(a+1)}{a+p}}. \end{aligned}$$

où

$$s_n = \sum_{i=1}^n \Delta_i.$$

Pour calculer l'expression de S_n^2, définie dans le lemme précédent, on utilise le lemme suivant :

Lemme 1.1.4 *([4], 2004) Soit* $\{\Delta_i, i\varepsilon\mathbb{N}\}$ *une suite des variables aléatoires réelles* α*-mélangeante, de coefficient de mélange* α_n *vérifiant :*

$$\exists c \in \mathbb{R}^{*+}, \quad a \in \mathbb{R}^{*+} \qquad \alpha(n) \leq cn^{-a}$$

et si $\|\Delta_i\|_\infty < \infty, \forall i$*, alors, pour tout* $\varepsilon > 0$ *et* $r > 0$

$$S_n^2 = \sum_{i=1}^n \sum_{j=1}^n |cov(\Delta_i, \Delta_j)|.$$

pour calculer l'expression de S_n^2, définie dans le lemme précédent, on utilise le lemme suivant :

Lemme 1.1.5 *([4], 2004) Soit* $\{\Delta_i, i\varepsilon\mathbb{N}\}$ *une suite des variables aléatoires réelle* α*-mélangeante, de coefficient des mélange* α_n*, telle que* $\|\Delta_i\|_\infty < \infty, \forall i$*. On a pour tout* $i \neq j$ *:*

$$|cov(\Delta_i, \Delta_j)| \leq 4\|\Delta_i\|_\infty\|\Delta_j\|_\infty\alpha_{|i-j|}.$$

et dans le cas fonctionnelle on introduit

$$|cov(\Delta_i, \Delta_j)| \leq C(\alpha_{|i-j|})^{\frac{(p-2)}{p}}.$$

Lemme 1.1.6 *([4], 2004) Soit* $\{\Delta_i, i\epsilon\mathbb{N}\}$ *une suite des variables aléatoires réelle α-mélangeante, de coefficient des mélange* α_n*, telle que* $\|\Delta_i\|_\infty < \infty, \forall i$*. On a pour tout* $i \neq j$ *:*

$$|cov(\Delta_i, \Delta_j)| \leq C(\alpha_{|i-j|})^{\frac{(p-2)}{p}}.$$

Chapitre 2

Données complètes

L'objective de ce chapitre est montrer la convergence presque complète de l'estimateur à noyau de la fonction de hasard conditionnelle lorsque les observations sont fonctionnelles et complètes. Comme toute étude asymptotique non paramétrique fonctionnelle Cette étude met en évidence la propriété de concentration de la mesure de probabilité de la variable fonctionnelle considérée sur les petites boules. Ce chapitre est dévisé en trois Sections. La première Section est consacrée à la présentation du modèle et à la construction de l'estimation de la fonction de hasard conditionnelle, dans la deuxième Section, on s'intèresse à la convergence presque complète unforme sur un compact fixe S de l'estimateur construit dans le cas où les observations sont indépendantes identiquement distribuées (i.i.d). Tandits que, dans la dernière Section on traitera le cas où les observations sont α-mélangeantes.

2.1 Modèle

Soit (Z, X) un couple de variable aléatoire à valeur dans $\mathcal{F} \times \mathbb{R}$ où $\mathcal{F}$ est un espace semi-métrique. Ce chapitre est consacré au problème général de l'estimation d'une fonction de hasard conditionnelle d'une variable aléatoire réelle X sachant une variable aléatoire Z à valeur dans un même espace probabilisé $(\Omega, A, \mathbb{P})$. On désigne par F^z la fonction de répartition conditionnelle de X sachant $Z = z$ par

$$F^z(x) = \mathbb{P}(X \leq x)|Z = z)$$

on suppose que cette distribution est absolument continue par rapport à la mesure de Lebesgue dont, on désigne par f^x la densité conditionnelle. Par ailleurs, on estime la fonction de répartition conditionnelle par

$$\widehat{F}^z(x) = \frac{\sum_{i=1}^{n} K(h_K^{-1}d(z,Z_i))H(h_H^{-1}(x-X_i))}{\sum_{i=1}^{n} K(h_K^{-1}d(z,Z_i))}$$

où K est un noyau, H est une fonction de répartition et $h_K = h_{K,n}$(resp.$h_H = h_{H,n}$) est une suite de réels positifs. De $\hat{F}^z(x)$, on déduit un estimateur de la densité conditionnelle, notée $\hat{f}^z(x)$, défini par

$$\widehat{f}^z(x) = \frac{\sum_{i=1}^{n} K(h_K^{-1}d(z,Z_i))H^{'}(h_H^{-1}(x-X_i))}{h_H \sum_{i=1}^{n} K(h_K^{-1}d(z,Z_i))} \quad \forall x \in \mathbb{R}$$

où H' est la dérivée de H. En ce qui concerne l'estimateur de la fonction de hasard conditionnelle, notée $\hat{h}^z$, (parfois appelé aussi taux de survie ou fonction de risque) s'interpréte comme la probabilité instantanée de sortir de l'état que l'on observe à la date x, sachant que le sujet est encore dans cet états en x,

$$h^Z(x) = \lim_{\Delta x \to 0} \frac{\mathbb{P}(X \leq x + \Delta x | X \geq x, Z)}{\Delta}, \qquad x > 0$$

Son estimateur peut se construire de la manière suivantes :

$$\hat{h}^z(x) = \frac{\hat{f}^z(x)}{1 - \hat{F}^z(x)}$$

2.2 Cas i.i.d.

Le but de cette section est d'étudié le modèle de la fonction de hasard conditionnelle dans lequel les observations sont indépendantes et identiquement distribuées (i.i.d). Ainsi, lorsque la dimension de l'espace $\mathcal{F}$ n'est pas nécessairement finie.

Hypothèses

On fixe un point x de l'espace fonctionnel $\mathcal{F}$ et on considère les hypothèses suivantes :

(H1) Sur la variables explicative

$$\forall h > 0, \phi_z(h_K) = \mathbb{P}(Z \in \mathrm{B}(z, h_K)) = \mathbb{P}(Z \in \{z' \in \mathcal{F}, d(z, z') < h\}) > 0$$

(H2) Sur le modèle non paramétrique

$$\exists A_z < \infty, \exists b_1, b_2 > 0, \forall (x_1, x_2) \in \mathcal{S}^2, \forall (Z_1, Z_2) \in \mathrm{N}_\mathrm{z}^2$$

$$|F^{z_1}(x_1) - F^{z_2}(x_2)| \leq A_z(d(z_1, z_2)^{b_1} + |x_1 - x_2)^{b_2}|$$

$$|f^{z_1}(x_1) - f^{z_2}(x_2)| \leq A_z(d(z_1, z_2)^{b_1} + |x_1 - x_2|^{b_2})$$

(H3) $\exists \nu < \infty, \forall (x, z') \in \mathcal{S} \times \mathrm{N}_z, f^{z'}(x) \leq \nu$

(H4) $\exists \beta < \infty, \forall (x, z') \in \mathcal{S} \times \mathrm{N}_z, F^{z'}(x) \leq 1 - \beta$

(H5) le noyau H est dérivable tel que

- i)$\exists A < \infty, \forall (x_1, x_2) \in \mathbb{R}^2, \quad |H'(x_1) - H'(x_2)| \leq A|x_1 - x_2|$
- ii)H' est de support compact $[-1, 1]$ et $H'(t) > 0, \forall t \in [0, 1]$
- iii) $\forall (t_1, t_2) \in \mathbb{R}^2, \quad |H(t_1) - H(t_2)| \leq C|t_1 - t_2|$
- v)$\exists \nu > 0, \lim_{t \in \infty} |y|^{1+\nu}|H(t)| = 0$
- vi) $\int |t|^{b_2} H^{(1)}(t)dt < \infty$

(H6) Le noyau fonctionnel K vérifié les conditions suivantes

- i) K est à support compact $[0, 1]$,
- ii)$\exists A_1, A_2, \forall t \in (0, 1), 0 < A_1 < K(t) < A_2 < \infty$

(H7) La longeur de la fenêtre h_K vérifie les conditions :

$$\lim_{n \to 0} h_K = 0, \quad \text{et} \quad \lim_{n \to 0} \frac{\log n}{n h_H \phi_n h_K} = 0$$

(H8) La longeur de la fenêtre h_H vérifie les conditions suivantes :

$$\lim_{n \to \infty} h_H = 0, \quad \text{et} \quad \exists a > 0, \quad \lim_{n \to \infty} n^\alpha h_H = \infty$$

Remarque 2.2.1 *Les hypothèses $(H1)$ sur la fonction de concentration et $(H2) - (H4)$ sont des conditions de régularité portant sur la loi conditionnelle qui caractérisent l'espace fonctionnel de notre modèle. Tandis que les autres hypothèses sont des conditions téchniques qui garantissent le bon comportement des estimateurs $\hat{F}^z$ et $\hat{f}^z$.*

2.2.1 Résultat

Théorème 2.2.1 *Sous les hypothèses (H1)-(H8), on a*

$$\sup_{x \in \mathbf{S}} |\widehat{h}^z(x) - h^z(x)| = o(h_K^{b1}) + o(h_H^{b2}) + o_{p.co}\left(\sqrt{\frac{\log n}{n h_H \phi_z(h_K)}}\right) \tag{2.1}$$

tel que

$$\widehat{h}^z(x) = \frac{\widehat{f}^z(x)}{1-\widehat{F}^z(x)} \quad et \quad h^z(x) = \frac{f^z(x)}{1-F^z(x)}$$

Preuve du théorème (2.2.1)

La démonstration du théorème est basée sur la décomposition suivante :

$$\widehat{h}^z(x) - h^z(x) = \frac{1}{1-\widehat{F}^z(x)}\left(\left(\widehat{f}^z(x) - f^z(x)\right) + h^z(x)\left(\widehat{F}^z(x) - F^z(x)\right)\right)$$

D'où, pour tout $x \in \mathcal{S}$ et pour une constante $c < \infty$

$$\sup_{x\in\mathbf{S}} |\widehat{h}^z(x) - h^z(x)| \leq c \left(\frac{\sup_{x\in\mathbf{S}} |\widehat{f}^z(x) - f^z(x)| + \sup_{x\in\mathbf{S}} |\widehat{F}^z(x) - F^z(x)|}{\inf_{x\in\mathbf{S}} |1-\widehat{F}^z(x)|}\right) \tag{2.2}$$

La démonstration du théorème repose sur les résultats suivants

$$\sup_{x\in\mathbf{S}} |\widehat{F}^z(x) - F^z(x)| = o(h_K^{b1}) + o(h_H^{b2}) + o_{p.co}\left(\sqrt{\frac{\log n}{n\phi_z(h_K)}}\right)$$

$$\sup_{x\in\mathbf{S}} |\widehat{f}^z(x) - f^z(x)| = o(h_K^{b1}) + o(h_H^{b2}) + o_{p.co}\left(\sqrt{\frac{\log n}{nh_H\phi_z(h_K)}}\right)$$

$$\exists\eta > 0 \qquad \text{tel que} \qquad \sum_n \mathbb{P}\left(\inf_{x\in\mathbf{S}} |1-\widehat{F}^z(x)|\right) < \infty$$

dont les démonstration sont basées, respectivement, sur les décompositions suivantes

$$\hat{F}^x(z) - F^x(z) = \frac{1}{\hat{F}_D^x}\left\{\left(\hat{F}_N^z(x) - \mathbb{E}\hat{F}_N^z(x)\right) - \left(F^z(x) - \mathbb{E}\hat{F}_N^z(x)\right)\right\} + \frac{F^z(x)}{\hat{F}_D^z}(\mathbb{E}\hat{F}_D^z - \hat{F}_D^z) \tag{2.3}$$

$$\begin{aligned} \hat{f}^z(x) - f^z(x) &= \frac{1}{\hat{f}_D^z}\left\{\left(\hat{f}_N^z(x) - \mathbb{E}\hat{f}_N^z(x)\right) - \left(f^z(x) - \mathbb{E}\hat{f}_N^z(x)\right)\right\} \\ &+ \frac{f^z(x)}{\hat{f}_D^z}(\mathbb{E}\hat{f}_D^z - \hat{f}_D^z) \end{aligned}$$

Finalement, le théorème est une conséquance des résultats suivants.

Lemme 2.2.1 *sous les hypothèses* $(H1), (H5)$ *et* $(H6)$ *on a*

$$|\widehat{F}_D^z - \mathbb{E}\widehat{F}_D^z| = o_{p.co}\left(\sqrt{\frac{\log n}{n\phi_z(h_K)}}\right) \tag{2.4}$$

Lemme 2.2.2 *sous les hypothéses* $(H1)$ *et* $(H2)$ *on a*

$$\sup_{x\in\mathbf{S}}|F^z(x)-\mathbb{E}\widehat{F}_N^z(x)| = o(h_K^{b1})+o(h_H^{b2}) \tag{2.5}$$

Lemme 2.2.3 *sous les hypothéses* $(H1)$ *et* $(H8)$ *on a*

$$\sup_{x\in\mathbf{S}}|\widehat{F}_N^z(x)-\mathbb{E}\widehat{F}_N^z(x)| = O_{p.co}\left(\sqrt{\frac{\log n}{nh_H\phi_z(h_K)}}\right) \tag{2.6}$$

Corollaire 2.2.1 *sous les hypothéses de théorème on a*

$$\exists\eta>0, \sum_n \mathbb{P}(\widehat{F}_D^z<\eta)<\infty \tag{2.7}$$

Dans les preuves suivants nous noterons pour tout $i\in\mathbb{N}$

$$K_i(x)=K(h_K^{-1}d(z,Z_1)), \qquad H_i(x)=H(h_H^{-1}(x-X_i))$$

2.2.2 Démonstration

Preuve du lemme 2.2.1

on a :

$$\widehat{F}_D^z-\mathbb{E}\widehat{F}_D^z = \frac{1}{n\mathbb{E}K_1}\sum_{i=1}^n\left(K\left(\frac{d(z,Z_i)}{h_K}\right)-\mathbb{E}K\left(\frac{d(z,Z_i}{h_K}\right)\right)$$

Où

$$\Delta_i=\frac{1}{n\mathbb{E}K_1}\left(K\left(\frac{d(z,Z_i)}{h_K}\right)-\mathbb{E}K\left(\frac{d(z,Z_i}{h_K}\right)\right)$$

d'aprés l'hypothèse $(H1)$, on a

$$C\mathbb{E}\mathbb{I}_{B(z,h_K)}(Z_i)<\mathbb{E}K_1<C\mathbb{E}\mathbb{I}_{B(z,h_K)}(Z_i)$$

alors

$$C\phi_x(h_K)<\mathbb{E}K_1<C\phi_x(h_K)$$

et puisque la fonction K est bornée, on peut majorer diretement

$$|\Delta_i|<\frac{c}{\phi_z(h_k)}, \quad \mathbb{E}(\Delta_i^2)<\frac{c'}{\phi_z(h_k)}$$

On applique l'inégalité de Hoeffdings (1.1.1, [4], 1963) aux variables Δ_i

$$\mathbb{P}\left(|\widehat{F}_D^z-\mathbb{E}\widehat{F}_D^z|>\epsilon\right)=\mathbb{P}\left(\frac{1}{n}\left|\sum_n\Delta_i\right|>\epsilon\right)\leq 2\exp\left(\frac{-n\epsilon^2C}{4\phi_z(h_K)}\right)$$

En prenant

$$\epsilon = \epsilon_0 \sqrt{\frac{\log n}{n\phi_z(h_K)}}$$

On arrive, pour tout $\epsilon_0 > 0$ à

$$\mathbb{P}\left(|\widehat{F}_D^z - \mathbb{E}\widehat{F}_D^z| > \epsilon \sqrt{\frac{\log n}{n\phi_z(h_k)}} \right) \leq 2n^{-c\epsilon_0^2}$$

Pour un choix convenable de ϵ, la sèrie de terme gènèral $n^{-c\epsilon^2}$ converge. On peut écrire

$$\sum_{n=1}^{+\infty} \mathbb{P}\left(|\widehat{F}_D^z - \mathbb{E}\widehat{F}_D^z| > \epsilon \sqrt{\frac{\log n}{n\phi_z(h_k)}} \right) < +\infty$$

Il suffit de choisir $\epsilon_0 > \dfrac{1}{\sqrt{C}}$ pour que la série converge.

Preuve du lemme 2.2.2

Du fait que les variables aléatoires sont identiquement distribuées, on a

$$\begin{aligned} \mathbb{E}\widehat{F}_N^z(x) - F^z(x) &= (\mathbb{E}K_1(z))^{-1}\left[\mathbb{E}K_1(z)H_1(t) - F^z(x)\right] \\ &= (\mathbb{E}K_1(z))^{-1}\mathbb{E}\left(K_1(z)\left[\mathbb{E}(H_1/Z) - F^z(x)\right]\right) \end{aligned}$$

En intégrant par parties, on montre que

$$\mathbb{E}(H_1(t)/Z) = \int_{\mathbb{R}} H\left(h_H^{-1}(x-z)\right) f^z(\mu)d\mu = h_H^{-1}\int_{\mathbb{R}} H^{'}(h_H^{-1}(x-z))F^x(z)dz$$

On considérons le changement de variable $t = \frac{x-\mu}{h_H}$ on obtient

$$\mathbb{E}\left(H_1(t)/Z\right) = \int_{\mathbb{R}} H^{(1)}(t)F^z(x - h_H t)dt$$

donc

$$|\mathbb{E}(H_1(t)/Z) - F^z(x)| \leq \int_{\mathbb{R}} H^{'}(t)|F^z(x - h_H t) - F^z(x)|dt$$

Sous (H2), on a

$$\forall x \in S, \quad \mathbb{I}_{B(z,h_K)}|\mathbb{E}(H_1(t)/Z) - F^z(x)| \leq \int_{\mathbb{R}} H^{'}(t)(h_K^{b_1} + |t|^{b_2}h_H^{b_2})dt$$

Comme H est une densité de probabilité, alors, l'hypothèse (H5) achève la démonstration du lemme.

Preuve de lemme (2.2.3)

On utilise le recouvrement suivant

$$S \subset \bigcup_{k=1}^{s_n} (\tau_k - l_n, \tau_k + l_n)$$

avec

$$l_n = C\tau_n^{-1} = Cn^{-c}$$

tel que $c > 0, C > 0$ et $\tau_1, \cdots, \tau_{s_n}$ sont des points de S.

Ainsi, on utilise la décomposition suivante :

$$\mathbb{P}\left(\sup_{x\in\mathbf{S}}\left|\widehat{F}_N^z(x) - \mathbb{E}\widehat{F}_N^z(x)\right| > \epsilon\sqrt{\frac{\log n}{n\phi_z(h_K)}}\right) \leq \underbrace{\mathbb{P}\left(\sup_{x\in\mathbf{S}}\left|\widehat{F}_N^z(x) - \widehat{F}_N^z(\tau_x)\right| > \epsilon/3\sqrt{\frac{\log n}{n\phi_z(h_K)}}\right)}_{A_1}$$

$$+\underbrace{\mathbb{P}\left(\sup_{x\in\mathbf{S}}\left|\widehat{F}_N^z(\tau_x) - \mathbb{E}\widehat{F}_N^z(\tau_x)\right| > \epsilon/3\sqrt{\frac{\log n}{n\phi_z(h_K)}}\right)}_{A_2}$$

$$+\underbrace{\mathbb{P}\left(\sup_{x\in\mathbf{S}}\left|\mathbb{E}\widehat{F}_N^z(\tau_x) - \mathbb{E}\widehat{F}_N^z(x)\right| > \epsilon/3\sqrt{\frac{\log n}{n\phi_z(h_K)}}\right)}_{A_3}$$

En ce qui Concerne le terme A_1, on a

$$\begin{aligned}
\frac{1}{\widehat{F}_D^z}\left|\widehat{F}_N^z(x) - \widehat{F}_N^z(\tau_x)\right| &\leq \frac{1}{n\widehat{F}_D^z\mathbb{E}K_1(z)}\sum_{i=1}^{n} K_i(z)\left|H_i(x) - H_i(\tau_x)\right| \\
&\leq \frac{c|x-\tau_x|}{h_H}\left(\frac{1}{n\mathbb{E}K_1}\sum_{i=1}^{n} K_i\right) \\
&\leq \frac{l_n}{h_H}\widehat{F}_D^z \\
&\leq c\frac{l_n}{h_H}
\end{aligned}$$

Par la définition de l_n et sous l'hypothèse $(H8)$ on obtient,

$$\frac{l_n}{h_H} = o\left(\sqrt{\log n(n\phi_z(h_k))^{-1}}\right)$$

Pour le terme A_2, on a

$$\mathbb{P}\left(\sup_{x\in S}|\hat{F}_N^z(\tau_x) - \mathbb{E}\hat{F}_N^z(\tau_x)| > \eta\sqrt{\frac{\log n}{n\phi_z(h_{K)}}}\right)$$

$$= \mathbb{P}\left(\max_{k\in(1,\cdots,\tau_n)}\left|\hat{F}_N^z(\tau_x) - \mathbb{E}\hat{F}_N^z(\tau_x)\right| > \eta\sqrt{\frac{\log n}{n\phi_z(h_K}}\right)$$

$$\leq s_n \max_{k\in(1,\cdots,\tau_n)}\mathbb{P}\left(|\hat{F}_N^z(\tau_x) - \mathbb{E}\hat{F}_N^z(\tau_x)| > \eta\sqrt{\frac{\log n}{n\phi_z(h_K)}}\right)$$

et

$$\widehat{F}_N^z(\tau_x) - \mathbb{E}\widehat{F}_N^z(\tau_x) = \frac{1}{n}\sum_{i=1}^n \Delta_i$$

avec

$$\Delta_i = \frac{1}{\mathbb{E}K_1}\left(H_iK_i - \mathbb{E}H_iK_i\right)$$

Pour appliquer l'inégalité de Hoeffding (1.1.1, [4]) aux variables Δ_i, il faut majorer les deux termes $|\Delta_i|$ et $\mathbb{E}\Delta_i$. En effet, pour le premier terme $|\Delta_i|$,
on a

$$|\Delta_i| \leq C\phi_z(h_K)^{-1}$$

Pour le deuxième terme $\mathbb{E}\Delta_i^2$, on a

$$\begin{aligned}\mathbb{E}\Delta_i^2 &= \frac{\mathbb{E}H_i^2K_i^2}{(\mathbb{E}K_1)^2} - \frac{\mathbb{E}^2(H_iK_i)}{(\mathbb{E}K_1)^2}\\ &\leq \frac{\mathbb{E}(K_i^2\mathbb{E}(H_i^2|X))}{(\mathbb{E}K_1)^2}\\ &\leq C\phi_z(h_K)^{-1}\end{aligned}$$

Donc

$$\mathbb{P}\left(|\widehat{F}_N^z(\tau_x) - \mathbb{E}\widehat{F}_N^z(\tau_x)| > \epsilon\right) = \mathbb{P}\left(\frac{1}{n}\sum_{i=1}^n|\Delta_i| > \epsilon\right) \leq 2\exp\left(\frac{-n\epsilon^2C}{4\phi_z(h_K)}\right)$$

En prenant

$$\epsilon = \epsilon_0 \sqrt{\frac{\log n}{n\phi_z(h_K)}}$$

On arrive, pour tout $\epsilon_0 > 0$ à

$$\mathbb{P}\left(|\widehat{F}_N^z(\tau_x) - \mathbb{E}\widehat{F}_N^z(\tau_x)| > \epsilon_0 \sqrt{\frac{\log n}{n\phi_z(h_K)}}\right) < 2n^{-C\epsilon_0^2}$$

donc

$$\sum_n \mathbb{P}\left(|\widehat{F}_N^z(\tau_x) - \mathbb{E}\widehat{F}_N^z(\tau_x)| > \epsilon_0 \sqrt{\frac{\log n}{n\phi_z(h_K)}}\right) < \sum_n 2n^{-C\epsilon_0^2}$$

Il suffit de choisir $\epsilon_0 > \dfrac{1}{\sqrt{C}}$ pour que la série converge.
Concernant le terme A_3, d'aprés A_1, on a

$$\sup_{x\in\mathcal{S}} |\mathbb{E}\widehat{F}_N^z(\tau_x) - \mathbb{E}\widehat{F}_N^z(x)| \leq c\frac{l_n}{h_H}$$

Preuve du corrolaire (2.2.1)

On a

$$\mathbb{E}\hat{F}_D^z = 1$$

alors

$$\mathbb{P}\left(|\hat{F}_D^z| \leq 1/2\right) \leq \mathbb{P}\left(|\hat{F}_D^z - \mathbb{E}\hat{F}_D^z| > 1/2\right)$$

Donc,

$$\sum_{i=1}^n \mathbb{P}\left(|\hat{F}_D^z| \leq 1/2\right) \leq \sum_{i=1}^n \left(\mathbb{P}|\hat{F}_D^z - \mathbb{E}\hat{F}_D^z| > 1/2\right) < \infty$$

Ainsi, le résultat est une conséquence du lemme précedent.
Pour la deuxième décomposition, on a

$$\begin{aligned} \widehat{f}^z(x) - f^z(x) &= \frac{1}{\widehat{f}_D^z}\left\{(\widehat{f}_N^z(x) - \mathbb{E}\widehat{f}_N^z(x)) - (f^z(x) - \mathbb{E}\widehat{f}_N^z(x))\right\} \\ &+ \frac{f^z(x)}{\widehat{f}_D^z}[\mathbb{E}\widehat{f}_D^z(x) - \widehat{f}_D^z(x)] \end{aligned}$$

La preuve est trés voisine de celle du théorème précédente avec une une légère modification dans la définition de l'estmateur en remplaçant $\hat{F}^z(x)$ par $\hat{f}^z(x)$. La preuve est basée sur un décomposition similaire. Ainsi, il suffit de montrer les équations suivantes :

Lemme 2.2.4 *sous les hypothèses* $(H1)$ *et* $(H2)$ *on a*

$$\sup_{x\in\mathbf{S}} |f^z(x) - \mathbb{E}\widehat{f}_N^z(x)| = o(h_K^{b1}) + o(h_H^{b2}) \tag{2.8}$$

Lemme 2.2.5 *sous les hypothèses* $(H1)$ *et* $(H8)$ *on a*

$$\sup_{x\in\mathbf{S}} |(\widehat{f}_N^z(x) - \mathbb{E}\widehat{f}_N^z(x)| = o_{p.co}\left(\sqrt{\frac{\log n}{nh_H\phi_z(h_K)}}\right) \tag{2.9}$$

Lemme 2.2.6 *sous les hypothèses* $(H1)$ *et* $(H8)$ *on a*

$$\sup_{x\in\mathbf{S}} |(\widehat{f}_D^z(x) - \mathbb{E}\widehat{f}_D^z(x)| = o_{p.co}\left(\sqrt{\frac{\log n}{nh_H\phi_z(h_K)}}\right) \tag{2.10}$$

Corollaire 2.2.2 *sous les hypothèses* $(H1)$ *et* $(H8)$ *on a*

$$\sum_{i=1}^{n} \mathbb{P}\left(\inf_{x\in\mathbf{S}} |\widehat{f}_D^z(x)| > \epsilon\right) < \infty \tag{2.11}$$

Preuve du lemme 2.2.4

On a, par équidistribution des observations

$$\begin{aligned}
\mathbb{E}\widehat{f}_N^z(x) - \widehat{f}^z(x) &= \frac{1}{h_H\mathbb{E}K_1}\left[\mathbb{E}K_1H^{(1)}\left(\frac{x-X_1}{h_H}\right) - h_Hf^z(x)\right] \\
&= \frac{1}{h_H\mathbb{E}K_1}\mathbb{E}\left(K_1\left[\mathbb{E}(H^{(1)}(h_H^{-1}(x-X_1))/Z) - h_Hf^z(x)\right]\right)
\end{aligned}$$

on a :

$$\mathbb{E}(H^{(1)}(h_H^{-1}(x-X_1))/Z) = \int_{\mathbb{R}} H^{(1)}\left(\frac{x-u}{h_H}\right) f^z(u)du$$

On fait un changement de variable $t = \frac{x-u}{h_H}$

$$\begin{aligned}
|\mathbb{E}(H_1^{(1)}(h_H^{-1}(x-X_1))/Z) - h_Hf^z(x)| &= h_H\int_{\mathbb{R}} H_1^{(1)}(t)|f^z(x-h_Ht) - f^z(x)|dt \\
&\leq h_H\int_{\mathbb{R}} H^{(1)}(t)|f^z(x-h_Ht)dt - f^z(x)|
\end{aligned}$$

on a l'hypothése $(H3)$ pour obtenir

$$|\mathbb{E}(H_1^{(1)}(h_H^{-1}(x-X_1))/Z) - h_Hf^z(x)| \leq ch_H\int_{\mathbb{R}} H^{(1)}(h_k^{b_1} + |t|^{b_2}h_H^{b_2})dt$$

Comme $H^{(1)}$ est une densité de probabilité, alors l'hypothèse $(H3)$ achève la preuve du lemme.

Preuve du lemme 2.2.5

On montre que

$$\begin{aligned}
\sup_{x\in\mathbf{S}} |\widehat{f}_N^z(x) - \mathbb{E}\widehat{f}_N^z(x)| &\leq \underbrace{\sup_{x\in\mathbf{S}} |\widehat{f}_N^z(x) - \widehat{f}_N^z(\tau_x)|}_{B_1} + \underbrace{\sup_{x\in\mathbf{S}} |\widehat{f}_N^z(\tau_x) - \mathbb{E}\widehat{f}_N^z(\tau_x)|}_{B_2} \\
&+ \underbrace{\sup_{x\in\mathbf{S}} |\mathbb{E}\widehat{f}_N^z(\tau_x) - \mathbb{E}\widehat{f}_N^z(x)|}_{B_3}
\end{aligned}$$

Concerne le terme B_1

$$\sup_{x\in S}\left|\hat{f}_N^z(x) - \hat{f}_N^z(\tau_x)\right| \leq \sup_{x\in S}\left(\frac{1}{nh_H\mathbb{E}K_1}\sum_{i=1}^n K_i|H_i^{(1)}(x) - H_i^{(1)}(\tau_x)|\right)$$

D'aprés l'hypothèse $(H5)$ on a

$$\begin{aligned}
\frac{1}{\hat{f}_D^z}\sup_{x\in S}\left(\frac{1}{nh_H\mathbb{E}K_1}\sum_{i=1}^n K_i|H_i^{(1)}(x) - H_i^{(1)}(\tau_x)|\right) &= \sup_{x\in S}\frac{C|x-\tau_x|}{nh_H^2\mathbb{E}K_1}\sum_{i=1}^n K_i \\
&\leq \sup_{x\in S}\frac{C|x-\tau_x|}{h_H^2} \\
&\leq C\frac{l_n}{h_H^2}
\end{aligned}$$

Comme $l_n = Cn^{-c}$, alors

$$\begin{aligned}
\sup_{x\in S}\left|\hat{f}_N^z(x) - \hat{f}_N^z(\tau_x)\right| &\leq \frac{Cn^{-c}}{h_H^2} \\
&\leq \frac{C}{n^c h_H^2}
\end{aligned}$$

Par conséquent,

$$\sup_{x\in S}\left|\hat{f}_N^z(x) - \hat{f}_N^z(\tau_x)\right| = o\left(\sqrt{\frac{\log n}{nh_H\phi_z h_H}}\right)$$

Pour le terme B_2

$$\mathbb{P}\left(\sup_{x\in\mathbf{S}} |\widehat{f}_N^z(\tau_x) - \mathbb{E}\widehat{f}_N^z(\tau_x)| > \eta/3\sqrt{\frac{\log n}{nh_H\phi_z(h_K)}}\right)$$

$$= \mathbb{P}\left(\max_{x\in(1,\ldots,s_n)} |\widehat{f}_N^z(x) - \mathbb{E}\widehat{f}_N^z(\tau_x)| > \eta/3\sqrt{\frac{\log n}{nh_H\phi_z(h_K)}} \right)$$

$$\leq S_n \max_{x\in(1,\ldots,s_n)} \mathbb{P}\left(|\widehat{f}_N^z(\tau_x) - \mathbb{E}\widehat{f}_N^z(\tau_x)| > \eta/3\sqrt{\frac{\log n}{nh_H\phi_z(h_k)}} \right)$$

on a aussi

$$\begin{aligned} |\widehat{f}_N^z(\tau_x) - \mathbb{E}\widehat{f}_N^z(\tau_x)| &= \frac{1}{n}\sum_{i=1}^n \left(\frac{H_1^{(1)}(\tau_x)K_i}{h_H\mathbb{E}K_1} - \frac{\mathbb{E}(H_1^{(1)}(\tau_m)K_i)}{h_H\mathbb{E}K_1} \right) \\ &= \frac{1}{n}\sum_{i=1}^n \Delta_i^* \end{aligned}$$

On applique maintenant l'inégalité d'Hoeffding (lemme (1.1.1, [4]). Du fait que le noyau K est borné, alors, on peut dire

$$|\Delta_i^*| \leq \frac{c}{nh_H\phi_z(h_H)}$$

Il suffit d'évaluer,

$$Var\Delta_i \leq \mathbb{E}(\Gamma_i^2)$$

alors

$$\begin{aligned} \mathbb{E}\Gamma_i^2 &= \frac{\mathbb{E}(K_i^2 H_i^1(\tau_x^2))}{(h_H\mathbb{E}K_1)^2} \\ &= \frac{\mathbb{E}K_i^2\mathbb{E}(H_i^1(\tau_x^2)/Z)}{(h_H\mathbb{E}K_1)^2} \end{aligned}$$

donc

$$\underbrace{\left| \frac{1}{h_H}\mathbb{E}(H_1^{(1)}(\tau_x)^2)/Z - f^z(\tau_x)\int_{\mathbb{R}} H^{(1)2}(x)dx \right|}_{I} = \left| \frac{1}{h_H}\mathbb{E}H^{(1)2}\left(\frac{\tau_x - x}{h_H}\right) - f^z(\tau_x)\int_{\mathbb{R}} H^{(1)2}(x)dx \right|$$

$$= \left| \frac{1}{h_H}\int_{\mathbb{R}} H^{(1)2}\left(\frac{\tau_x - z}{h_H}\right) f^z(t)dt - f^z(\tau_x)\int_{\mathbb{R}} H^{(1)2}(t)dt \right|$$

par un changement des variables $\tau_k - z = u$, on obtient

$$I = \left| \int_{\mathbb{R}} \frac{1}{h_H} H^{(1)2}\left(\frac{u}{h_H}\right) f^z(\tau_x - u)du - f^z(\tau_x)\int_{\mathbb{R}} H^{(1)2}(t)\,dt \right|$$

On pose, $t = \dfrac{u}{h_H}$, on obtient

$$\begin{aligned}
&= \left|\int_{\mathbb{R}} \frac{1}{h_H} H^{(1)2}\left(\frac{u}{h_H}\right)\left(f^z(\tau_x - u) - f^z(\tau_x)\right) du\right| \\
&\leq \left|\int_{|u|\leq M} \frac{1}{h_H} H^{(1)2}\left(\frac{u}{h_H}\right)\left(f^z(\tau_x - u) - f^z(\tau_x)\right) du\right| \\
&+ \left|\int_{|u|>M} \frac{1}{h_H} H^{(1)2}\left(\frac{u}{h_H}\right)\left(f^z(\tau_x - u)\right) du\right| \\
&\leq C \sup_{|u|\leq M} |f^z(\tau_x - u) - f^z(\tau_x)| \int_{|u|\leq M} H^{(1)^2}\left(\frac{u}{h_H}\right) du \\
&+ \left|\int_{|u|\leq M} H^{(1)^2}\left(\frac{u}{h_H}\right)\left(f^z(\tau_x - u) - f^z(\tau_x)\right|
\end{aligned}$$

On prend $t = \dfrac{u}{h_H} \Rightarrow du = h_H dt$ et on a $|u| > M \Rightarrow |u| > A/h_H$

$$\begin{aligned}
&\leq C \sup_{|u|\leq M} |f^z(\tau_x - u) - f^x(\tau_x)| + \left|\int_{|u|>A/h_H} H^{(1)^2}(t)\left(f^z(\tau_x - th_H) - f^{\tau_x}(\tau_x)\right) h_H dt\right| \\
&\leq C \sup_{|u|\leq M} |f^z(\tau_x - u) - f^x(\tau_x)| + \left|\int_{|u|>A/h_H} H^{(1)^2}(t)\left(f^z(\tau_x - th\right)\right| \\
&+ \left|\int_{|u|>A/h_H} H^{(1)^2}(t) f^z(\tau_x)\right| \\
&\leq \underbrace{C \sup_{|u|\leq M} |f^z(\tau_x - u) - f^x(\tau_x)|}_{B_1} + \underbrace{\left|\int_{|u|>M/h_H} tH^{(1)^2}(t)\frac{f^z(\tau_x - th_H)}{t}\right|}_{B_2} \\
&+ \underbrace{f^z(\tau_x)\int_{|u|>M/h_H} H^{(1)^2}(u) dt}_{B_3}
\end{aligned}$$

D'aprés l'hypothèse $(H5, iii)$ on a

$$\forall \epsilon > 0, \forall M > 0, \exists n_{A,\epsilon}, B_2 + B_3 < \epsilon$$

et puisque f^x est continue, alors

$$\forall \epsilon > 0, \exists n_{M,\epsilon}, \forall M \leq M_\epsilon, B_1 < \epsilon$$

Ainsi

$$\lim_{n\to\infty} \mathbb{E}(H_1^{(1)}(\tau_x)^2/Z) = f^z \int_{\mathbb{R}} H^{(1)^2}(t) dt$$

et puisque

$$0 < \mathbb{E}K_1 < C\phi_x(h_K), \quad \text{et} \quad \mathbb{E}K_1^2 < C'\phi_x(h_K)$$

Alors

$$\mathbb{E}\Gamma_i^2 = O(h_H^{-1}\phi_x(h_K)^{-1})$$

On applique maintenant l'inégalité de Hoeffdings (1.1.1, [4]), on obtient

$$\begin{aligned}
\mathbb{P}\left(|\widehat{f}_N^z(\tau_x) - \mathbb{E}\widehat{f}_N^z(\tau_x)| > \eta\sqrt{\frac{\log n}{nh_H\phi_z(h_K)}}\right) &= \mathbb{P}\left(\frac{1}{n}\sum_{n=1}^{\infty}|\Delta_i| > \eta\sqrt{\frac{\log n}{nh_H\phi_z(h_K)}}\right) \\
&\leq 2\exp\left(-cn^{\frac{\eta^2\log n}{nh_H\phi_z(h_K)}}h_H\phi_z(h_K)\right) \\
&\leq cn^{-c\eta^2}
\end{aligned}$$

en choisisant η de façon que $c\eta^2 = \frac{2}{3}\alpha + \frac{2}{3}$. on obtient

$$\begin{aligned}
\mathbb{P}\left(\sup_{x\in}|\widehat{f}_N^z(\tau_x) - \mathbb{E}\widehat{f}_N^z(\tau_x)| > \eta\sqrt{\frac{\log n}{nh_H\phi_z(h_K)}}\right) &\leq cS_n n^{-c\eta^2} \\
&\leq n^{-c\eta^2}c/l_n \\
&\leq n^{-1-\alpha}
\end{aligned}$$

D'où

$$\sup_{x\in\mathbf{S}}\left|\widehat{f}^z(\tau_x) - \mathbb{E}\widehat{f}_N^z(\tau_x)\right| = o_{p.co}\left(\sqrt{\frac{\log n}{nh_H\phi_z(h_K)}}\right)$$

Pour le terme B_3

d'aprés B_1 on a

$$\frac{1}{f_D^z}\sup_{x\in S}\left|\widehat{f}^z(x) - \mathbb{E}\widehat{f}_N^z(\tau_x)\right| \leq C\frac{l_n}{h_K^2}$$

Donc

$$\frac{1}{f_D^z}\sup_{x\in S}\left|\mathbb{E}\widehat{f}^z(\tau_x) - \mathbb{E}\widehat{f}_N^z(x)\right| \leq C\frac{l_n}{h_K^2}$$

Preuve du corrolaire (2.2.2)

On a

$$\mathbb{P}\left(|\hat{f}_D^z| \leq 1/2\right) \leq \mathbb{P}\left(|\hat{f}_D^z - \mathbb{E}\hat{f}_D^z| > 1/2\right)$$

Donc

$$\sum_{i=1}^{n}\mathbb{P}\left(|\hat{f}_D^z| \leq 1/2\right) \leq \sum_{i=1}^{n}\left(\mathbb{P}|\hat{f}_D^z - \mathbb{E}\hat{f}_D^z| > 1/2\right)$$

Donc il reste de montrer que

$$\sum_{i=1}^{n}\mathbb{P}\left(\inf_{x\in\mathcal{S}}|1 - \widehat{F}^z(x)|\right) < \infty$$

On a

$$\inf_{x \in \mathcal{S}} |1 - \widehat{F}^z(x)| \leq \left(1 - \sup_{x \in \mathcal{S}} \widehat{F}^z(x)\right)/2$$

$$\Longrightarrow \sup_{x \in \mathcal{S}} |\widehat{F}^z(x) - F^z(x)| \geq \left(1 - \sup_{x \in \mathcal{S}} \widehat{F}^z(x)\right)/2$$

En terme de probabilité, on obtient

$$\sum_{i=1}^{n} \mathbb{P}\left(\inf_{x \in \mathcal{S}} |1 - \widehat{F}^z(x)| \leq \left(1 - \sup_{x \in \mathcal{S}} \widehat{F}^z(x)\right)/2\right)$$

$$\leq \sum_{i=1}^{n} \mathbb{P}\left(\sup_{x \in \mathcal{S}} |\widehat{F}^z(x) - F^z(x)| \geq \left(1 - \sup_{x \in \mathcal{S}} \widehat{F}^z(x)\right)/2\right) < \infty$$

2.3 Cas dépendant

Le but est consacré à la généralisation du résultats donné dans le cas précèdent à des observations mélageantes. On fixe comme objectif la convergence presque complète uniforme de l'estimateur à noyau de la fonction de hasard conditionnelle défini dans la première section. Notons que les autres propriétés asymptotique se généralisent de la même manière.

2.3.1 Hypothèses

Dans cette section on garde les mêmes notations ainsi que les mêmes hypothèses du cas indépendant et ajoute aussi les hypothèses suivantes :

(H'1) La suite $(X_i, Z_i)_{i \in \mathbb{N}}$ est α-mélangeante tels que

$$\exists a, c \in \mathbb{R}_+^*, \quad \forall n \in \mathbb{N}, \quad \alpha(n) \leq cn^{-a}$$

(H'2) La densité jointe de (X_i, X_j) sachant (Z_i, Z_j) existe et est bornée, et

$$\exists \epsilon_1 \in]0, 1], 0 < \sup_{i \neq j} \mathbb{P}((Z_i, Z_j) \in B(z, h) \times B(z, h)) = o\left(\phi_z(h))^{1+\epsilon_1}\right)$$

(H'3)

$$\exists \epsilon_1 \in]0, 1[, \quad a > \frac{1 + \epsilon_1}{\epsilon_1 \epsilon_2}, \quad \text{et} \quad h_H \phi_z(h_K) = o(n^{-\epsilon_2})$$

Remarque 2.3.1 *Les hypothèses $(H'1) - (H'3)$ sont ajoutés pour éviter l'expression de covariance dans la vitesse de convergence. Autrement dit, on peut démontrer la convergence presque complète sans ces hypothèses. Cependant, la vitesse de converence* *sera donné en fonction de covariance des observations et elle sera lente par rapport à la vitesse du cas indépendant. Ainsi, nous établissons la convergence presque complète avec la même précision, mais, sous des conditions un plus fort que le cas indépendant.*

Théorème 2.3.1 *sous les hypothèses du théorème (2.2.1) et* $(H'1) - (H'3)$ *on a*

$$\sup_{x\in\mathbf{S}} \left|\widehat{h}^z(x) - h^z(x)\right| = o(h_K^{b1}) + o(h_H^{b2}) + o_{p.co}\left(\sqrt{\frac{\log n}{nh_H\phi_z(h_K)}}\right) \tag{2.12}$$

2.3.2 Démonstration

preuve du théorème (2.3.1)

La démonstartion est basé sur les mêmes aruments analytiques utilisés dans preuve du théorème (2.2.1), on peut dire que la propriété de l'indépendance des observations n'aucune influence sur la partie biais de la vitesse. Ainsi, nous allons évalué la partie de dispersion laquelle est basée sue le lemme suivant :

Lemme 2.3.1 *Sous les hypothèses* $(H1) - (H8)$ *et* $(H'.1) - (H'.3)$ *on a*

$$S_n^{2*} = o(nh_K) + o(n^2\alpha(h_K \log n)^{-1}) \tag{2.13}$$

où

$$S_{n*}^2 = \sum_{i=1}^{n}\sum_{i\neq j} |cov(\Delta_i, \Delta_j)|$$

et

$$\Delta_i = K\left(\frac{z-Z_i}{h_K}\right) H\left(\frac{x-X_i}{h_H}\right) - \mathbb{E}\left(K\left(\frac{z-Z_i}{h_K}\right) H\left(\frac{x-X_i}{h_H}\right)\right)$$

Lemme 2.3.2 *Sous les hypothèses* $(H1) - (H8)$ *et* $(H'.1) - (H'.3)$ *on a*

$$\sup_{x\in\mathbf{S}} |\widehat{F}_D^z - \mathbb{E}F_D^z| = o\left(\sqrt{\frac{\log n}{n\phi_z(h_K)}}\right), \qquad p.co \tag{2.14}$$

Lemme 2.3.3 *Sous les hypothèses* $(H1) - (H8)$*, et* $(H'1 - H'3)$ *on a*

$$\sup_{x\in\mathbf{S}} |\widehat{F}_N^z(x) - \mathbb{E}\widehat{F}_N^z(x)| = o\left(\sqrt{\frac{\log n}{n\phi_z(h_K)}}\right), \qquad p.co \tag{2.15}$$

Lemme 2.3.4 *Sous les hypothèses du théorème, on a*

$$\sup_{x\in\mathbf{S}} |\widehat{f}_N^z(x) - \mathbb{E}\widehat{f}_N^z(x)| = o\left(\sqrt{\frac{\log n}{nh_H\phi_z(h_K)}}\right), \qquad p.co \tag{2.16}$$

Preuve du lemme 2.3.1

Par définition de Δ_i, on a

d'une part

$$\begin{aligned} cov(\Delta_i,\Delta_j) &= \mathbb{E}(\Delta_i\Delta_j) \\ &= \mathbb{E}\left(K\left(\frac{z-Z_i}{h_K}\right)H\left(\frac{x-X_i}{h_H}\right)K\left(\frac{z-Z_j}{h_K}\right)H\left(\frac{x-X_j}{h_H}\right)\right) \\ &+ \mathbb{E}\left(K\left(\frac{z-Z_j}{h_K}\right)H\left(\frac{x-X_j}{h_H}\right)\right)\mathbb{E}\left(K\left(\frac{z-Z_j}{h_K}\right)H\left(\frac{x-X_j}{h_H}\right)\right) \\ &\leq C\left|K\left(\frac{z-Z_j}{h_K}\right)K\left(\frac{z-Z_j}{h_H}\right)\right| + \mathbb{E}\left|K\left(\frac{z-Z_j}{h_K}\right)\right|\mathbb{E}\left|K\left(\frac{z-Z_j}{h_H}\right)\right| \end{aligned}$$

Par ailleurs,

$$\begin{aligned} |cov(\Delta_i,\Delta_j)| &\leq C\left(\mathbb{E}(\mathbb{I}_{B(z,h_K)}(Z_i)\mathbb{I}_{B(z,h_K)}(Z_j)) + \mathbb{E}(\mathbb{I}_{B(z,h_K)}(Z_i))\mathbb{E}(\mathbb{I}_{B(z,h_K)}(Z_j)\right) \\ &\leq C(\mathbb{P}(Z_i,Z_j)\in B(z,h_K)\times B(z,h_K) + \mathbb{P}(Z_i\in B(z,h_K)\mathbb{P}(Z_j\in B(z,h_K)) \\ &= o(\phi_z(h_K)^{1+\epsilon} + o(\phi_z^2(h_K)) \end{aligned}$$

D'autre part, nous avons d'aprés l'inégalité de covariance (1.1.5, [4])

$$|cov(\Delta_i,\Delta_j)| \leq C\alpha(|i-j|)$$

A l'aide de ces équations on montre que

$$S_n^{2*} = (n\phi_z(h_K))$$

En effet, on utilise les techniques de Masry (1.1.5, [4] 1986) et on partage la somme sur les deux ensembles

$$\begin{aligned} S_1 &= \{(i,j) \quad \text{tel que} \quad 1\leq j-i\leq m_n\} \\ S_2 &= \{(i,j), \quad \text{tel que } m_n+1\leq j-i\leq n-1\} \end{aligned}$$

Par ailleurs

$$\begin{aligned} S_n^{*2} &= \sum_{i=1}^{n}\sum_{i\neq j=1}^{n}|cov(\Delta_i,\Delta_j)| \\ &= \sum_{S_1}|cov(\Delta_i,\Delta_j)| + \sum_{S_2}|cov(\Delta_i,\Delta_j)| \\ &\leq C\sum_{S_1}(\phi_z(h_K)^{1+\epsilon} + \phi_z^2(h_K) + C\sum_{S_2}\alpha(|i-j|>u_n) \end{aligned}$$

$$
\begin{aligned}
&\leq C\sum_{i=1}^{n}\sum_{1\leq|j-i|\leq m_n}\phi_z(h_K)^{1+\epsilon}+\sum_{i=1}\sum_{|j-i|>m_n}\alpha(|i-j|>u_n)\\
&\leq Cnm_n(\phi_z(h_K)^{1+\epsilon}+(\phi_z^2(h_K)+n^2\alpha(u_n)\\
&\leq Cnm_n(\phi_z(h_K)^{1+\epsilon}+(\phi_z^2(h_K)))+n^2(u_n)^{-a}
\end{aligned}
$$

On prend $u_n=\phi_z(h_K)^{-\epsilon}$, on obtient, alors

$$S_n^{*2}=o(n\phi_z(h_K))$$

et

$$S_n^2=S_n^{*2}+nVar(\Delta_1)$$

et

$$Var(\Delta_i)\leq C\phi_z(h_K)$$

d'où,

$$S_n^2=o(n\phi_z(h_K)$$

Preuve du lemme 2.3.3

On applique l'inégalité de Fuk-Nagaev (Rio 1990, [4], 1.1.2) aux variables Δ_i, on obtient,

$$\mathbb{P}\left(\left|\widehat{F}_N^z(\tau_x)-\mathbb{E}\widehat{F}_N^z(\tau_x)\right|>4\epsilon\right)\leq C\left\{\frac{n}{r}\left(\frac{r}{\epsilon n\mathbb{E}K_1}\right)^{a+1}+\left(1+\frac{\epsilon^2n^2(\mathbb{E}K_1)^2}{rS_n^2}\right)^{-\frac{r}{2}}\right\}$$

Si on prend $\epsilon=\epsilon_0\sqrt{\dfrac{\log n}{n\phi_z(h_K)}}$, on obtient

$$
\begin{aligned}
&\mathbb{P}\left(\left|\widehat{F}_N^z(\tau_x)-\mathbb{E}\widehat{F}_N^z(\tau_x)\right|>4\epsilon_0\sqrt{\frac{\log n}{n\phi_z(h_K)}}\right)\\
&\leq 4\left(\frac{\eta^2\log n}{16r}\right)^{-r/2}+2ncr^{-1}\left(\frac{8r\sqrt{n\phi_z(h_K)}}{n\phi_z(h_K)\epsilon_0\sqrt{\log n}}\right)+2ncr^{-1}\left(\frac{8r}{\epsilon_0}\right)^{a+1}(n\log n\phi_z(h_K))^{-(a+1)/2}\\
&\leq 4n^{-\epsilon_0/32}+cn^{-(a+1)}n^{(1-(a+1)/2)}r^a n^{((a-3)/2-\beta(a+1)/2}(\log n)^{-(a+1)/2}
\end{aligned}
$$

On choisi $r=C(\log n)^2$, et $\tau_n=n^c$

$$\tau_n\max_{k\in(1,\cdots,s_n)}\mathbb{P}\left(\left|\widehat{F}_N^z(\tau_x)-\mathbb{E}\widehat{F}_N^z(\tau_x)\right|>4\epsilon_0\sqrt{\frac{\log n}{n\phi_z(h_K)}}\right)\leq Cn^{-\frac{\eta^2}{32}+c}+cn^{-1+c-(\frac{\beta(a+1)}{2})}(\log n)^{(3a-1)/2}$$

Pour ϵ_0 suffisamment grand, on a

$$\exists \nu > 0, \quad \tau_n \max_{k \in (1,\cdots,s_n)} \mathbb{P}\left(\left| \widehat{F}_N^z(\tau_x) - \mathbb{E}\widehat{F}_N^z(\tau_x) \right| > 4\epsilon_0 \sqrt{\frac{\log n}{n\phi_z(h_K)}} \right) \leq cn^{-1-\nu}$$

Finalement

$$\sum_{i=1}^{n} \left(\mathbb{P}\left(\sup_{x \in S} \left| \widehat{F}_N^z(\tau_x) - \mathbb{E}\widehat{F}_N^z(\tau_x) \right| > 4\epsilon \right) \right) < \infty$$

Preuve du lemme 2.3.4

On a

$$\mathbb{P}\left(\sup_{x \in \mathcal{S}} \left| \widehat{f}_N^z(\tau_x) - \mathbb{E}\widehat{f}_N^z(\tau_x) \right| > \epsilon \sqrt{\frac{\log n}{nh_H\phi_z(h_K)}} \right)$$

$$\leq \tfrac{c}{l_n} \max_{m_n \in (m_1,\cdots,m_n)} \mathbb{P}\left(\left| \widehat{f}_N^z(\tau_x) - \mathbb{E}\widehat{f}_N^z(\tau_x) \right| > \epsilon \sqrt{\frac{\log n}{nh_H\phi_z(h_K)}} \right)$$

on a aussi

$$\widehat{f}_N^z(\tau_x) - \mathbb{E}\widehat{f}_N^z(\tau_x) = \frac{1}{nh_H\mathbb{E}K_1} \sum_{i=1}^{n} \underbrace{H_i^{(1)}(\tau_x)K_i - \mathbb{E}(H_i^{(1)}(\tau_x)K_i)}_{\Lambda_i^*}$$

et

$$S_n'^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} \left| cov(\Lambda_i^*, \Lambda_j^*) \right|$$

Pour calcule $S_n'^2$ en utilisant la même méthode que dans S_n^2 et prenant

$$u_n = \frac{1}{\phi_z(h_K)^{\epsilon}},$$

on obtient

$$S_n'^2 = o(nh_H\phi_z(h_K)).$$

On applique l'inégalité de Fuk-Nagaev (Rio 1990, [4], 1.1.2) aux variables Δ_i, on obtient

$$\mathbb{P}\left(\sup_{x \in \mathcal{S}} \left| \widehat{f}_N^z(x_m) - \mathbb{E}\widehat{f}_N^z(x_m) \right| > \epsilon \sqrt{\frac{\log n}{nh_H\phi_z(h_K)}} \right) < Cl_n^{-1}(C_1' + C_2')$$

avec

$$C_1' = 4(1 + \frac{\epsilon^2}{rS_n'^2})^{\frac{-r}{2}} \quad et \quad C_2' = 4cnr^{-1}\left(\frac{r^{\epsilon+1}}{\epsilon} \right)$$

D'aprés l'hypothèse $H'3$ et avec un choix de $r = c(\log n)^2$ et $l_n = n^{-\frac{3}{2}a+\frac{1}{2}}$.
$\exists v > 0$, pour n assez grand, on a

$$l_n^{-1}(C_1' + C_2') \leq Cn^{-1-v}$$

Finalement

$$\sup_{x \in \mathcal{S}} \left| \widehat{f}_N^z(\tau_x) - \mathbb{E}\widehat{f}_N^z(\tau_x) \right| = o\left(\sqrt{\frac{\log n}{nh_H \phi_z(h_K)}} \right)$$

Chapitre 3

Données censurées

Dans ce chapitre nous nous intéressons à la convergence presque complète de l'estimateur à noyau de la fonction de hasard conditionnelle avec une variable explicative fonctionnelle et incomplètes. De même que le chapitre précédent, la diffculté technique réside dans l'absence d'une mesure de référence classique. De plus, en statistique nonparamétrique vectorielle, l'étude de la normalité asymptotique nécessite un calcul asymptotique exact des termes du biais et de la variance qui est basée sur la dérivabilité des modèles non-paramétriques. Cependant, la structure topologique de notre espace fonctionnel est trop faible pour définir cette notion. Ici nous considérons un espace fonctionnel semi-métrique. Ainsi, l'impact principal de cette contribution est la généralisation en dimension infinie des résultats déjà existants en statistique multi-variée mais sous des conditions moins restrictives. Dans la pratique, lors d'applications médicales en particulier, on peut se trouver en présence de variables censurées. Ce problème est habituellement modélisé en considérant une variable positive C dite de censure, et les variables aléatoires observées ne sont pas les couples (X_i, Z_i) mais seulement (T_i, Δ_i, Z_i) où $T_i = \min(X_i, C_i)$ et $\Delta_i = \mathbb{I}_{X_i \leq C_i}$.
L'objectif de ce chapitre est d'adapter ces idées au cadre de variable explicative Z fonctionnelle, et de construire un estimateur de type noyau de la fonction de hasard conditionnelle h^z adapté aux échantillons censurés.

3.1 Modèle

Soit (T_i, Δ_i, Z_i) un échantillon d'un couple de variable aléatoire où $T_i = \min(X_i, C_i)$ et $\Delta_i = \mathbb{I}_{X_i \leq C_i}$, X à prend ses valeurs dans $\mathbb{R}$ et Z à valeurs dans un espace semi-

métrique $(F, d(,;,))$ de dimension infinie.
Le probleme général dans ce chapitre est l'estimation de la fonction de hasard conditionnelle, d'une variable aléatoire réelle X sachant une variable aléatoire à valeurs dans un espace fonctionnel, pour des données censurées dans le cas indépendantes et α-mélangeantes. On peut définir l'estimateur de la fonction de hasard coditionnelle à partir des estimateurs suivantes :

$$\widehat{\varphi}^z(t) = \frac{\sum_{i=1}^{n} K(h_K^{-1}d(z,Z_i))H'(h_H^{-1}(t-T_i))\Delta_i}{h_H \mathbb{E}K(h_K^{-1}d(z,Z_i))}$$

K est un noyau, H' est la dérivé de H, h_K, (resp h_H)) est une suite de nombre réel positif. et

$$\widehat{L}^z(t) = \frac{\sum_{i=1}^{n} K(h_K^{-1}d(z,Z_i))H(h_H^{-1}(t-T_i))}{\mathbb{E}K(h_K^{-1}d(z,Z_i))}$$

Finalement, l'estimation de la fonction de la fonction hasard conditionnelle dans le cas imcomplète est définie par :

$$\widehat{h}^z(t) = \frac{\widehat{\varphi}^z(t)}{1-\widehat{L}^z(t)}$$

3.2 Cas indépendant

3.2.1 Hypothèses

On garde les hypothèse du cas précédent et on ajoute les conditions suivantes :

H1a Les triplet (X_i, C_i, Z_i) sont i.i.d[1].

H2a Conditionnellement à Z_i, les variables X et C sont indépendantes

H3a $\exists A_z < \infty, \exists b_1, b_2 > 0, \forall (t_1,t_2) \in S^2, \forall (z_1,z_2) \in N_z^2$:

$$|L^{z_1}(t_1) - L^{z_2}(t_2)| \leq A_z(d(z_1,z_2)^{b_1} + |t_1-t_2|^{b_2})$$

$$|\varphi^{z_1}(t_1) - \varphi^{z_2}(t_2)| \leq A_z(d(z_1,z_2)^{b_1} + |t_1-t_2|^{b_2})$$

H4a $\exists \mu < \infty, \quad \varphi^{z'}(t) < \mu, \quad \forall (t,z') \in \mathbb{R}_+ \times N_z$

H5a $\exists \mu < \infty, \quad L^{z'}(t) < 1-\eta, \quad \forall (t,z') \in \mathbb{R}_+ \times N_z$

1. indépendant et identiquement distribuée

3.2.2 Propriétés Asymptotique

Théorème 3.2.1 *Sous les hypothése (H1)-(H8) et (H1a)-(H5a) on a*

$$\sup_{t\in\mathcal{S}}\left|\widehat{h}^z(t)-h^z(t)\right| = o(h_K^{b1})+o(h_H^{b2})+O_{p.co}\left(\sqrt{\frac{\log n}{nh_H\phi_z(h_K)}}\right)$$

Preuve du théorème(3.2.1)

La démonstration du théorème est une consequence de la décomposition suivante

$$\widehat{h}^z(t)-h^z(t) = \frac{1}{1-\widehat{L}^z(t)}\left(\widehat{\varphi}^z(t)-\varphi^z(t)\right)+\frac{h^z(t)}{1-\widehat{L}^z(t)}\left(\widehat{L}^z(t)-L^z(t)\right)$$

D'ou pour tout $x\in\mathcal{S}$ et $C>0$

$$\sup_{t\in\mathbf{S}}|\widehat{h}^z(t)-h^z(t)|\leq C\left[\frac{\sup_{x\in\mathbf{S}}|\widehat{\varphi}^z(t)-\varphi^z(t)|+\sup_{x\in\mathbf{S}}\left|\widehat{L}^z(t)-L^z(t)\right|}{\inf_{x\in\mathbf{S}}|1-\widehat{L}^z(t)|}\right] \tag{3.1}$$

D'aprés la décomposition (3.1), il suffit de montrer que

$$\sup_{x\in\mathbf{S}}\left|\widehat{L}^z(t)-L^z(t)\right| = o(h_K^{b_1})+o(h_H^{b_2})+o_{p.co}\left(\sqrt{\frac{\log n}{n\phi_z(h_K)}}\right) \tag{3.2}$$

$$\sup_{x\in\mathbf{S}}|\widehat{\varphi}^z(t)-\varphi^z(t)| = o(h_K^{b_1})+o(h_H^{b_2})+o_{p.co}\left(\sqrt{\frac{\log n}{nh_H\phi_z(h_K)}}\right) \tag{3.3}$$

On peut dire que (3.3)écrire sous la forme

$$\begin{aligned}\widehat{\varphi}^z(t)-\varphi^z(t) &= \frac{1}{\widehat{\varphi}_D^z}\left\{(\widehat{\varphi}_N^z(t)-\mathbb{E}\widehat{\varphi}_N^z(t))-(\varphi^z(t)-\mathbb{E}\widehat{\varphi}_N^z(t))\right\}\\ &+ \frac{\varphi^z(t)}{\widehat{\varphi}_D^z}\left[\widehat{\varphi}_D^z-\mathbb{E}\widehat{\varphi}_D^z\right]\end{aligned}$$

tel que

$$\widehat{\varphi}_N^z(t)=\frac{\sum_{i=1}^n K_iH_i^{'}\Delta_i}{h_H\mathbb{E}K_1},\qquad \widehat{\varphi}_D^z(t)=\frac{\sum_{i=1}^n K_i}{h_H\mathbb{E}K_1}.$$

La preuve est bassée sur les lemmes suivantes :

Lemme 3.2.1 *Sous les hypothèses $(H1)$, $(H2)$ et $(H3a)$ on a*

$$\sup_{t\in\mathbf{S}}|\varphi^z(t)-\mathbb{E}\hat{\varphi}_N^z(t)| = o(h_K^{b1})+o(h_H^{b2}) \tag{3.4}$$

Lemme 3.2.2 *Sous les hypothèses* $(H1), (H5), (H6)$ *et* $(H4a)$ *on a*

$$\sup_{t\in\mathbf{S}} |\widehat{\varphi}_N^z(t) - \mathbb{E}\widehat{\varphi}_N^z(t)| = o_{p.co}\left(\sqrt{\frac{\log n}{n\phi_z(h_K)}}\right) \tag{3.5}$$

Lemme 3.2.3 *Sous les hypothèses* $(H1)$, $(H8)$ *et* $(H4a)$ *on a*

$$|\widehat{\varphi}_D^z(x) - \mathbb{E}\widehat{\varphi}_D^z(x)| = o_{p.co}\left(\sqrt{\frac{\log n}{nh_H\phi_z(h_K)}}\right) \tag{3.6}$$

Preuve de (3.2.1) :

Les observations sont indépendant et identiquement distribuée.

$$\begin{aligned}
\mathbb{E}\widehat{\varphi}_N^z(t) &= \mathbb{E}\left(\frac{1}{nh_H\mathbb{E}K_1(z)}\sum_{i=1}^n K_i(z)H_i(t)\Delta_i\right)\\
&= \frac{1}{h_H\mathbb{E}K_1(z)}\mathbb{E}\left(K_1(z)\mathbb{E}\left(H_1(t)\mathbb{I}_{(X_1\leq C_1)}/Z_1\right)\right)\\
&= \frac{1}{h_H\mathbb{E}K_1(z)}\mathbb{E}\left(K_1(z)\mathbb{E}\left(H_1(t)S_1^{Z_1}(X_1)/Z_1\right)\right)
\end{aligned}$$

d'autre part on a

$$\begin{aligned}
\mathbb{E}(H_1(t)S_1^z/Z_1) &= \int H^{'}\left(\frac{t-u}{h_H}\right)S_1^{Z_1}(u)f^{Z_1}(u)du\\
&= h_H\int H^{'}(v)\varphi^{Z_1}(t-vh_H)dv\\
&= h_H\left(\varphi^z(t) + o\left(h_H^{b_2} + h_K^{b^1}\right)\right)
\end{aligned}$$

Preuve de (3.2.2) ;

Soit $\epsilon > 0$. On utilise la même démarche que lors de la preuve du théorème précédente et on recouvre S par un nombre fini d'intervalles

$$S \subset \bigcup_{k=1}^{\tau_n}]t_k - l_n, t_k + l_n[$$

$$\mathbb{P}\left(\sup_{x\in\mathbf{S}} |\widehat{\varphi}_N^z(t) - \mathbb{E}\widehat{\varphi}_N^z(t)| > \epsilon\sqrt{\frac{\log n}{n\phi_z(h_K)}}\right)$$

$$\leq \underbrace{\mathbb{P}\left(\sup_{x\in\mathbf{S}} |\widehat{\varphi}_N^z(t) - \widehat{\varphi}_N^z(\tau_t)| > \epsilon/3\sqrt{\frac{\log n}{n\phi_z(h_K)}}\right)}_{A_1}$$

$$+ \underbrace{\mathbb{P}\left(\sup_{x\in\mathbf{S}}|\widehat{\varphi}_N^z(\tau_t) - \mathbb{E}\widehat{\varphi}_N^z(\tau_t)| > \epsilon/3\sqrt{\frac{\log n}{n\phi_z(h_K)}}\right)}_{A_2}$$

$$+ \underbrace{\mathbb{P}\left(\sup_{x\in\mathbf{S}}|\mathbb{E}\widehat{\varphi}_N^z(\tau_t) - \mathbb{E}\widehat{\varphi}_N^z(t)| > \epsilon/3\sqrt{\frac{\log n}{n\phi_z(h_K)}}\right)}_{A_3}$$

Pour le terme A_1 et pour tout $t \in S$, on a

$$\begin{aligned}
|\hat{\varphi}_N^z(t) - \hat{\varphi}_N^x(\tau_t)| &= \left|\frac{1}{nh_H\mathbb{E}K_1(z)}\sum_{i=1}^n \Delta_i K_i(z)(H_i(t) - H_i(\tau_t)\right| \\
&\leq \frac{C}{nh_H\mathbb{E}K_1(z)}\sum_{i=1}^n K_i(z)\frac{|t-\tau_t|}{h_H} \\
&\leq C\hat{\varphi}_D^z l_n h_H^{-2}
\end{aligned}$$

En prenant $l_n = Cu_n = Cn^{-c}$, on arrive à

$$\frac{1}{\hat{\varphi}_D^z}|\hat{\varphi}_N^z(t) - \varphi_N^z(\tau_t)| \leq \frac{C}{h_H^2 n^c}$$

Implique qu'il existe

$$\frac{C}{h_H^2 n^c} = o\left(\sqrt{\frac{\log n}{n\phi_z(h_K)}}\right)$$

Par conséquent

$$\sup_{t\in S}|\hat{\varphi}_N^z(t) - \hat{\varphi}_N^x(\tau_t)| = o\left(\sqrt{\frac{\log n}{n\phi_z(h_K)}}\right)$$

Pour le deuxième terme

$$\hat{\varphi}_N^z(\tau_t) - \mathbb{E}\hat{\varphi}_N^z(\tau_t) = \frac{1}{n}\sum_{i=1}^n \Delta_i^*$$

Pour cela il est nécessaire de trouver des majorants pour $|\Delta_i|$ et pour $\mathbb{E}\Delta_i^2$, pour appliquer l'inégalité de Hoeffdings (1.1.1, [4], 1963), en reprenant les calculs et on arrive à

$$|\Delta_i \leq \frac{C}{\phi_z(h)}$$

Pour le moments d'ordre deux de Δ_i, on pose

$$W_i = \frac{1}{h_H^2(\mathbb{E}K_1(z))^2}\mathbb{E}K_i^2(z)H_i^2(\tau_t)\Delta_i^2$$

Tel que

$$\begin{aligned}
\mathbb{E}W_i^2 &= \frac{1}{h_H^2(\mathbb{E}K_1(z))^2}\mathbb{E}K_i^2(z)H_i^2(\tau_t)\Delta_i^2 \\
&\leq C\frac{1}{h_H^2(\mathbb{E}K_1(z))^2}\mathbb{E}K_i^2(z)\mathbb{E}(H_i^2(\tau_t)/Z_i)) \\
&\leq C\frac{1}{h_H\phi_z(h)}\mathbb{E}\left(K_i^2(z)\int\frac{1}{h_H}H'\left(\frac{\tau_t-u}{h_H}\right)^2 f^{Z_i}(u)du\right) \\
&= O\left(\frac{1}{h_H\phi_z(h)}\right)
\end{aligned}$$

Il est évident que

$$\mathbb{E}\Delta_i^2 = Var(\tau_i) \leq \mathbb{E}\tau_i^2 \leq \frac{C}{h_H\phi_z(h_K)}$$

Pour tout $\epsilon > 0$, on a

$$\sup_{x\in\mathbf{S}}\left(|\widehat{\varphi}_N^z(\tau_x) - \mathbb{E}\widehat{\varphi}_N^z(\tau_x)| > \epsilon/3\sqrt{\frac{\log n}{n\phi_z(h_K)}}\right)$$

$$= \sum_{k=1}^{\tau_n}\max_{\tau_k}\left(\mathbb{P}\left(|\widehat{\varphi}_N^z(\tau_t) - \mathbb{E}\widehat{\varphi}_N^z(\tau_t)| > \epsilon/3\sqrt{\frac{\log n}{n\phi_z(h_k)}}\right)\right) \leq 2\sum_{k=1}^{\tau_t} n^{-c-(\epsilon_0^2/\beta)}$$

Un choix convenable de ϵ_0 permet de conclure que,

$$\sup_{x\in\mathbf{S}}|\widehat{\varphi}_N^z(\tau_t) - \mathbb{E}\widehat{\varphi}_N^z(\tau_t)| = 0\left(\frac{\log n}{n\phi_z(h_K)}\right), \qquad p.co$$

En ce qui concerne le dernier terme

$$|\mathbb{E}\hat{\varphi}_N(\tau_t) - \mathbb{E}\hat{\varphi}_N(t)| \leq \frac{Cl_n}{\phi_z(h_H)h_H}$$

On obtient

$$\begin{aligned}
|\mathbb{E}\hat{\varphi}_N(\tau_t) - \mathbb{E}\hat{\varphi}_N(t)| &\leq \frac{Cn^{-c}}{h_H^2\phi_z(h_K)} \\
&\leq \frac{C}{n^c h_H^2\phi_z(h_K)}
\end{aligned}$$

On montre l'existence de

$$\sup_{x\in S}|\mathbb{E}\hat{\varphi}_N(\tau_t) - \mathbb{E}\hat{\varphi}_N(t)| = o\left(\sqrt{\frac{\log n}{nh_H^2\phi_z(h_H)}}\right)$$

La démonstration de (3.2) est identique à celle de (2.3). Puisque l'estimateur $\hat{L}$ est le même que la fonction de répartition conditionnelle, donc la démonstration est similaire au cas complète.

3.3 Cas dépendant

Dans cette section on généralise les résultats du cas précédent à des observations mélangeantes. on étudie la convergence presque complète uniforme.

3.3.1 Hypothèses

On garde les mêmes notations, ainsi les mêmes hypothèses et on ajoute les conditions ci-dessous permettent de trouver la même vitesse de convergence que le cas indépendant.

– Hb1 Sur la corrélation des observations

  Les observations sont algébriquement $\alpha-$ mélangeante, c'est à dire qu'il existe deux constantes $c \in \mathbb{R}^{*+}$ et $a \in \mathbb{R}*+$ telles que les coefficients de mélange vérifiant,

$$\alpha(n) \leq cn^{-a}$$

– Hb2 Sur la concentration mixte des variables fonctionnelles

$$\exists \epsilon_1 \in]0,1[, 0 < \sup_{i \neq j} \mathbb{P}((Z_i, Z_j) \in B(z,h) \times B(z,h)) = o(\phi_z(h_K))^{1+\epsilon})$$

– H3b Sur le paramètre de lissage

$$\exists \epsilon_2 \in]0,1[, a > \frac{1+\epsilon_1}{\epsilon_1 \epsilon_2}, \quad \text{et} \quad h_H \phi_z(h_K) = o(n^{-\epsilon_2}$$

3.3.2 Propriétés Asymptotique

Théorème 3.3.1 *Sous les hypothése de théorème (3.2.1) et (Hb1,Hb2,Hb3) on a*

$$\sup_{t \in \mathcal{S}} \left| \widehat{h}^z(t) - h^z(t) \right| \quad = \quad o(h_K^{b1}) + o(h_H^{b2}) + o_{p.co}\left(\sqrt{\frac{\log n}{n h_H \phi_z(h_K)}} \right)$$

Démonstration

Le caracrère de l'indépendance des observations n'intervient pas dans la partie biais. Autrementdit, les vitesses de convergence de la partie biais serons les mêmes dans le cas des mélanges. Cependant, la partie dispersions est basé sur les deux lemmes suivants :

Lemme 3.3.1 *Sous les hypothèses du théorème, on a*

$$\sup_{t\in\mathbf{S}} |\widehat{\varphi}_N^z(t) - \mathbb{E}\widehat{\varphi}_N^z(t)| = o_{p.co}\left(\sqrt{\frac{\log n}{n\phi_z(h_K)}}\right) \tag{3.7}$$

Lemme 3.3.2 *Si les conditions sont satisfait, on a*

$$|\widehat{\varphi}_D^z(t) - \mathbb{E}\widehat{\varphi}_D^z(t)| = o_{p.co}\left(\sqrt{\frac{\log n}{nh_H\phi_z(h_K)}}\right) \tag{3.8}$$

Démonstration du lemme 3.3.1

On garde les mêmes arguments utilisés dans la démonstration du lemme du cas précédent, ainsi que le même recouvrement du compact S. Il est clair que le choix de coefficient de mélange nous permet de montrer que le premier et le troisième terme de la décomposition sont de type $o\left(\frac{\log n}{nh_H\phi_z(h_K)}\right)$ et on démontre seulement le deuxième terme

$$\sup_{t\in S} |\hat{\varphi}_N^z(\tau_t) - \mathbb{E}\widehat{\varphi}_N^z(\tau_t)| = o_{p.co}\left(\sqrt{\frac{\log n}{nh_H\phi_z(h_K)}}\right)$$

En effet

$$\hat{\varphi}_N^z(\tau_t) - \mathbb{E}\hat{\varphi}_N^z(\tau_t) = \frac{1}{n\phi_z(h_K)}\sum_{i=1}^{n} W_i$$

tel que

$$W_i = K_iH_i\Delta_i - \mathbb{E}(K_iH_i\Delta_i)$$

En utilisant l'inégalité exponentielle de Fuk-Nagaev (lemme(1.1.2), [4]), on obtient pour tout $\epsilon > 0$ et pour tout $r > 1$

$$\mathbb{P}\left(\left|\frac{1}{n\phi_z(h_K)}\sum_{i=1}^{n} W_i\right| > 4\epsilon\right) \leq 4\left(1 + \frac{\epsilon^2}{rS_n^2}\right)^{-r/2} + 2ncr^{-1}\left(\frac{2r}{\epsilon}\right)^{a+1}$$

où

$$S_n^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} |cov(W_{i1}, W_{i2})|$$

et par suite

$$|cov(W_{i1}, W_{i2})| \leq |\mathbb{E}(W_{i1}W_{i2})|$$

En vertu de l'hypothèse $(Hb2)$ on a

$$|\mathbb{E}W_{i1}W_{i2}| \leq \frac{C}{(h_H^2\mathbb{E}K_1(z))^2}\phi_z(h_K)^{1+\epsilon_1} \tag{3.9}$$

$$= 0(h_H^{-2}(\max(\phi_z(h_K)^{-1+\epsilon_1}, 1))) \tag{3.10}$$

et

$$cov(W_{i1}, W_{i2}) \leq C h_H^{-2} \phi_z(h_K)^{-2} \alpha(|i_1 - i_2|)$$

Puisque

$$S_n^2 = \sum_{i=1}^{n} var(W_i) + \sum_{0<|i_1-i_2|\leq v_n} cov(W_{i1}, W_{i2}) + \sum_{|i_1-i_2|>v_n} cov(W_{i1}, W_{i2})$$

Finalement

$$\begin{aligned} S_n^2 &= 0\left(\frac{n}{h_H \phi_z(h_K)}\right) + 0\left(n v_n h_H^{-2} \max(\phi_z(h_K)^{-1+\epsilon_1}, 1\right) \\ &+ o\left(h_H^{-2} \phi_z(h_K)^{-2} \sum_{|i_1-i_2|>v_n} \alpha(|i_1 - i_2|)\right) \end{aligned}$$

Il sufit de prendre $v_n = \phi_z(h_K)^{-\epsilon_1}$ pour obtenir

$$S_n^2 = O\left(\frac{n}{h_H \phi_z(h_K)}\right)$$

$$\hat{\varphi}_N^z(\tau_t) - \mathbb{E}\hat{\varphi}_N^z(\tau_t) = o\left(n^{-1}\sqrt{\log n S_n^2}\right), \qquad \text{p.co} \tag{3.11}$$

alors

$$\mathbb{P}\left(\left|\frac{1}{n h_H \phi_z(h_K)} \sum_{i=1}^{n} W_i\right| > 4\epsilon\right) \leq 4\left(1 + \frac{\epsilon^2}{r S_n^2}\right)^{-r/2} + 2ncr^{-1}\left(\frac{2r}{\epsilon}\right)^{a+1}$$

On revient à l'inégalit de Fuk nagaev (lemme(1.1.2), [4]), on remplace S_n^2 par son approximation asymptotique et on prend $r = C(logn)^2$
(resp. $\epsilon = \epsilon_0 \left(n h_H E K_1\right)^{-1} \sqrt{n h_H \phi_z(h_K) \log n}$) on obtient

$$P\left(|\hat{\varphi}_N^z(\tau_t) - \mathbb{E}\hat{\varphi}_N^z(\tau_t)| > 4\epsilon_0 \frac{\sqrt{S_n^2 \log n}}{n h_H \mathbb{E} K_1}\right) \leq C \frac{n}{(\log n)^2} \left(\frac{(\log n)^2}{\epsilon_0 \sqrt{n h_H \phi_z(h_K) \log n}}\right)^{a+1}$$

$$+ C exp\left(\frac{-C(logn)^2}{2} log\left(1 + \frac{\epsilon_0^2 n h_H \phi_z(h_H) \log n}{C S_n^2 (\log n)^2}\right)\right).$$

Le dernier terme de la partie gauche est tel que

$$\exp\left(-\frac{C(\log n)^2}{2} log\left(1 + \frac{\epsilon_0^2}{16 C \log n}\right)\right) \leq \exp\left(-(\epsilon_0^2/32)(\log n) \frac{log\left(1 + \frac{\epsilon_0^2}{16Clogn}\right)}{\frac{\epsilon_0^2}{16C\log n}}\right).$$

En raison que $\lim \log(1+x)/x = 0$ quad x tend vers 0, on peut trouver, sous un choix convenable de ϵ_0 un $\nu > 0$ vérifiant

$$exp\left(\frac{-C(logn)^2}{2} log\left(1+\frac{\epsilon_0^2 n h_H \phi_z(h_K)\log n}{CS_n^2(\log n)^2}\right)\right) \leq n^{-1-\nu}.$$

Concernant le premier terme, on utilise l'hypothèse $(Hb3)$, laquelle nous permet d'écrire on montre que

$$\phi_z(h_K) < n^{-\epsilon_2} h_H^{-1}$$

Ainsi,

$$h_H^{-(a+1)/2} n^{1-\frac{(a+1)}{2}} \phi_z(h_K)^{-(a+1)/2} < n^{-1-\nu}$$

En regroupant les inégalités précédentes, on peut trouver un $\nu > 0$ tel que

$$P\left(\hat{\varphi}_N(\tau_t) - \mathbb{E}\hat{\varphi}_N(\tau_t)| > 4\epsilon_0 \frac{\sqrt{S_n^2 \log n}}{n h_H \mathbb{E} K_1}\right)) \leq C n^{-1-\nu}$$

Finalement, et comme

$$EK_1(x) = O(\phi_z(h_K)0$$

on obtient

$$\left(\hat{\varphi}_N^z(\tau_t) - \mathbb{E}\hat{\varphi}_N^z(\tau_t)| > 4\epsilon_0 \frac{\sqrt{S_n^2 \log n}}{n h_H \mathbb{E} K_1}\right) = o\left(\sqrt{\frac{logn}{n h_H \phi_x(h_H)}}\right).$$

Conclusion

On peut monsionnée quelques remarques,

1. Dans le cas infin dimensionnel, la dégradation de la vitesse est liée à la faible concentration de la mesure de probabilité de la variables explicative fonctionnelle, il s'agit d'un problème de concentration.
2. L'utlisation d'une semi-métrique qui donne une forte concentration à la mesure de probabilité de la variable fonctionnelle, est alors une solution originale de ce problème de dégradation de la vitesse.
3. Les vitesses des convergences obtenues sont insensibles aux corrélations des observations. Cependant, certains hypothèses supplémentaires ont été introduites afin de prendre en compte la dépendance des observations.
4. Les données censurées sont des observations pour lesquelles la valeur exacte d'un évènement n'est pas toujours connue.
5. Les données censurées proviennent du fait qu'on n'a pas accés à toute information : au d'observer des réalisations indépendentes et identiquement distribuées de durées X, on observe la réalisation de la variable X soumis à diverse perturbations, indépendants ou non du phénomène étudié. Dans la suite, la variable durée de survie T est définie comme le délai écoulé entre la date d'origine T_0 et la date de survenue de l'événement. Donc, $\Delta = 1$ si l'événement est observer d'où $X_i = T_i$. On observe les vraies durées ou les durées complètes. $\Delta = 0$ si l'individu est censuré d'où $X_i = C_i$. On observe les durée imcomplètes (censurées).

Bibliographie

[1] G. Estévez-Pèrez, A. Quintela-del-Rio and P. Vieu, Convergence rate for cross-validatory bandwidth in kernel hazard estimation from dependent samples. J. Statist. Plann. Inference 104 (2002), 1-30.

[2] G. Estévez, On convergence rates for quadratic errors in kernel hazard estimation. Statist. Probab. Lett. 57 (2002), 231-241.

[3] Ezzahrioui, M.(2007). Prévision dans les modèles conditionnels en dimension infinie. PhD Thesis. Univ. du Littoral Côte d'Opale.

[4] F. Ferraty, A. Laksaci and P. Vieu, Estimating some characteristics of the conditional distribution in nonparametric functional models. Statist. Inference Stochastic Process. 9 (2006), 47-76.

[5] F. Ferraty et Vieu, ph Modèle de régresion pour variable aléatoire uni, multi et ∞-dimensionnées (2004). Cours de DEA

[6] F. Ferraty ,A. Rabhi,et P. Vieu,(2008).Estimation non-paramétrique de la fonction de hasard avec variable explicative fonctionnelle. Rev. Roumaine Math. Pures Appl.53,1-18.

[7] F. Ferraty and P. Vieu, Nonparametric Functional Data Analysis. Theory and Practice. Springer Series in Statistics. Springer, New York, (2006).

[8] J-P. Lecoutre and E. Ould-Said, Hazard rate estimation for strong mixing and censored processes. J. Nonparametr. Statist. 5 (1995), 83-89.

[9] P. J. Lecoutre, E. Ould-Saïd, (1993). Estimation de la fonction de hasard pour un processus fortement mélangeant avec censure. Publ. Inst. Statist. Univ. Paris. 37 , No.1-2, 59-69.

[10] J. Li, and L.T. Tran, (2007). Hazard rate estimation on random fields. J. Multivariate Anal. 98 ,1337-1355

[11] A. Laksaci, M. Mechab(2010). Estimation non parametrique de la fonction de hasard avec variable explicative fonctionnelle cas des donnees spatiales. Rev : Roumaine, Math Pures Appl. 55 , 35-51.

[12] G. G. Roussas, (1968). On some properties of nonparametric estimates of probability density functions. Bull. Soc. Math. Greece (N.S.) 9 29-43.

[13] A. Quintela-del-Rio, (2010). On non-parametric techniques for area-characteristic seismic hazard parameters. Geophys. J. Int. 180, 339-346.

[14] G.S. Watson, Leadbetter, M. R. (1964). Hazard analysis. I. Biometrika, 51, 175-184.