ESTIMATION NON-PARAMÉTRIQUE DE LA FONCTION DE HASARD AVEC VARIABLE EXPLICATIVE FONCTIONNELLE

FRÉDÉRIC FERRATY, ABBES RABHI et PHILIPPE VIEU

We introduce a nonparametric estimate of the conditional hazard function, when the covariate is functional. We prove consistency properties (with rates) in various situations, including censored and/or dependent variables. The rates of convergence emphasize the crucial role played by the small ball probabilities with respect to the distribution of the explanatory functional variable.

AMS 2000 Subject Classification: 62G05, 62G20, 62N02.

Key words: Données censurées, estimation non-paramétrique, fonction de hasard conditionnelle, probabilités de petites boules, processus α-mélangeant, variable fonctionnelle.

1. INTRODUCTION

L'estimation du taux de hasard, de part la variété de ses possibilités d'application, est une question importante en statistique. Ce sujet peut (et doit) être abordé sous plusieurs angles selon la complexité du problème posé: présence éventuelle de censure dans l'échantillon observé (phénomène courant dans les applications médicales par exemple), présence éventuelle de dépendance entre les variables observées (phénomène courant dans les applications sismologiques ou économétriques par exemple) ou bien présence de variables explicatives. De nombreuses techniques ont été étudiées dans la littérature pour traiter de ces différentes situations mais toutes ne traitent que de variables aléatoires explicatives réelles ou multi-dimensionnelles.

Les progrès techniques réalisés en matière de recueil et de stockage des données permettent de disposer de plus en plus souvent de données statistiques fonctionnelles: courbes, images, tableaux, ... Ces données sont modélisées comme étant des réalisations d'une variable aléatoire prenant ses valeurs dans un espace abstrait de dimension infinie, et la communauté scientifique s'est naturellement intéressée depuis quelques années au développement d'outils statistiques capables de traiter ce type d'échantillon.

REV. ROUMAINE MATH. PURES APPL., **53** (2008), *1*, 1–18

Ainsi, l'estimation d'un taux de hasard en présence de variable explicative fonctionnelle est une question d'actualité à laquelle cet article propose d'apporter un premier élément de réponse. Après un rapide survol bibliographique présenté au Paragraphe 2.1, le modèle de taux de hasard pour variable explicative fonctionnelle est présenté au Paragraphe 2.2. Les estimateurs que nous définissons sont basés sur les techniques de noyau de convolution. Ils sont tout d'abord présentés au Paragraphe 2.3 dans le cadre classique, avant d'être généralisés à des variables censurées dans le Paragraphe 2.4.

Ces estimateurs sont à notre connaissance les premiers à traiter de taux de hasard en présence de variable fonctionnelle. Nous proposons dans cet article d'étudier leurs comportements asymptotiques au travers de divers résultats de convergence presque complète (avec explicitation des vitesses de convergence). Pour fixer les idées, nous commençons dans le Paragraphe 3.1 par donner des résultats dans le cadre simple de données indépendantes et non censurées. Les extensions au cadre censuré sont présentées au Paragraphe 4.1. Afin de compléter l'éventail de nos résultats, nous étendons ceux-ci au cadre de variables aléatoires non nécessairement indépendantes. Plus précisément nous donnons dans les Paragraphes 3.2 et 4.2, des résultats de convergence lorsque les variables sont issues d'un processus uniformément fortement mélangeant.

Dans le cadre non censuré les propriétés de l'estimateur de la fonction de hasard conditionnelle s'obtiennent relativement facilement à partir de la littérature connue en matière d'estimation de fonction de répartition et de densité conditionnelles. Ainsi, les preuves des résultats du Paragraphe 3 seront présentées de manière synthétique en utilisant au maximum la littérature existante. Par contre, dans le cadre plus intéressant de variables censurées, ces propriétés asymptotiques ne s'obtiennent pas aussi directement, et pour améliorer la lisibilité du Paragraphe 4 les détails techniques des preuves qu'il contient sont reportés en fin d'article.

2. POSITION DU PROBLÈME

2.1. Le contexte bibliographique

Si X est une variable aléatoire associée à une durée de vie (c'est à dire, une variable aléatoire à valeurs dans $\mathbb{R}^+$, le taux de hasard de X (parfois appelé aussi fonction de hasard, taux de défaillance ou taux de survie) est défini au point x comme étant la probabilité instantanée que cette durée de vie se termine à l'instant x. Précisément, on a

$$h(x) = \lim_{\Delta x \to 0} \frac{P\left(X \leq x + \Delta x | X \geq x\right)}{\Delta x}, \quad x > 0. \tag{1}$$

Lorsque la variable X possède une densité f par rapport à la mesure de Lebesgue il est aisé de voir que ce taux de hasard peut être écrit

$$h(x) = \frac{f(x)}{S(x)}, \tag{2}$$

pour tout x tel que $F(x) < 1$, où F désigne la fonction de répartition de X et $S = 1 - F$ la fonction de survie de X.

Dans de nombreuses situations pratiques, on peut disposer d'une variable explicative Z et la question devient celle de l'estimation du taux de hasard conditionnel défini pour $x > 0$ par

$$h^Z(x) = \lim_{\Delta x \to 0} \frac{P\left(X \leq x + \Delta x | X > x,\ Z\right)}{\Delta x}, \tag{3}$$

qui s'écrit lui aussi naturellemet à partir de la densité conditionnelle $f^Z(\cdot)$ et de la fonction de répartition conditionnelle $F^Z(\cdot)$ (au travers de la fonction de survie conditionnelle $S^Z(\cdot) = 1 - F^Z(\cdot)$) de X sachant Z, sous la forme

$$h^Z(x) = \frac{f^Z(x)}{S^Z(x)} \tag{4}$$

dès que $F^Z(x) < 1$. L'étude des fonctions h et h^Z est d'un intérêt évident dans de nombreux domaines scientifiques (biologie, médecine, fiabilité, sismologie, économétrie, ...), et de nombreux auteurs se sont intéréssés à la construction d'estimateurs non-paramétriques de h. Une des techniques les plus courantes pour construire des estimateurs de h (respectivement de h^Z) est basée sur le résultat (2) (respectivement le résultat (4)) et consiste à étudier un quotient entre un estimateur de f (respectivement de f^Z) et un estimateur de S (respectivement de S^Z). L'article de Patil *et al.* [13] fait une présentation générale de ces techniques d'estimation. Les méthodes non-paramétriques basées sur les idées de noyau de convolution, qui sont connues pour leur bon comportement dans les problèmes d'estimation de densité (conditionnelle ou non), sont ainsi abondamment utilisées en estimation non-paramétrique de fonction de hasard. Un large éventail de la littérature dans ce domaine est fourni par les revues bibliographiques de Singpurwalla and Wong [15], Hassani *et al.* [7], Izenman [8], Gefeller and Michels [6] et Pascu and Vaduva [12].

2.2. **Hasard conditionnel pour explicative fonctionnelle**

Les progrès des procédés de recueil de données ont pour conséquence immédiate d'offrir la possibilité aux statisticiens de disposer de plus en plus souvent d'observations de variables fonctionnelles. Les ouvrages de Ramsay and Silverman [14] et Ferraty and Vieu [5] proposent un large éventail de méthodes statistiques, paramétriques ou non, récemment développées pour traiter divers problèmes d'estimation dans lesquels interviennent des variables aléatoires fonctionnelles (c'est à dire à valeurs dans un espace de dimension infinie). Jusqu'à présent de tels développements statistiques pour variables fonctionnelles n'existent pas dans le contexte de l'estimation d'une fonction de hasard, et ce malgré le potentiel évident en matière d'applications.

L'objectif de cet article est d'étudier un modèle de hasard conditionnel dans lequel la variable explicative Z n'est pas nécessairement réelle ou multidimensionnelle mais seulement supposée être à valeurs dans un espace abstrait $\mathcal{F}$ muni d'une semi-métrique d. Comme dans tout problème d'estimation non-paramétrique, la dimension de l'espace $\mathcal{F}$ joue un rôle important dans les propriétés de concentration de la variable X. Ainsi, lorsque cette dimension n'est pas nécessairement finie, les fonctions de probabilité de petites boules définies par

$$\phi_z(h) = P\left(Z \in B(z,h)\right) = P\left(Z \in \{z' \in \mathcal{F}, d(z,z') < h\}\right),$$

interviennent de manière directe dans le comportement asymptotique de tout estimateur non-paramétrique fonctionnel (voir Ferraty and Vieu [5]). Les résultats asymptotiques que nous présentons dans la suite de cet article sur l'estimation de la fonction h^Z n'échapperont pas à cette règle.

Dorénavant, z désigne un élément fixé de l'espace fonctionnel $\mathcal{F}$, N_z désigne un voisinage fixé de z et $\mathcal{S}$ est un compact fixé de $\mathbb{R}^+$. Nous serons amenés à faire l'hypothèse ci-dessous concernant la fonction de concentration ϕ_z:

(H1) $\forall h > 0,\ \phi_z(h) > 0$.

Le modèle non-paramétrique sur la fonction h^Z à estimer sera déterminé par des conditions de régularité portant sur la loi conditionnelle de X sachant Z. Ces conditions sont les suivantes:

(H2) $\exists A_z < \infty,\ \exists b_1, b_2 > 0,\ \forall (x_1,x_2) \in \mathcal{S}^2,\ \forall (z_1,z_2) \in N_z^2$:

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

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

(H4) $\exists \beta > 0,\ \forall (x,z') \in \mathcal{S} \times N_z,\ F^{z'}(x) \leq 1 - \beta$.

2.3. Construction de l'estimateur pour échantillon complet

Soient $(X_i, Z_i)_{1 \leq i \leq n}$ des variables aléatoires suivant chacune la même loi qu'un couple (X, Z) où X est à valeurs dans $\mathbb{R}$ et Z à valeurs dans l'espace semi-métrique $(\mathcal{F}, d(\cdot\,;\cdot))$. Dans ce paragraphe, nous nous plaçons dans le cas le plus simple pour lequel toutes les variables X_i et Z_i ont été observées.

Les avancées récentes en statistique non-paramétrique pour variables fonctionnelles, telles que présentées dans Ferraty and Vieu [5], font apparaître que les techniques basées sur les noyaux de convolution sont facilement transposables au cadre de variables fonctionnelles. Par ailleurs ces techniques à noyau possèdent de bonnes propriétés dans les problèmes d'estimation de fonction de hasard pour variables de dimension finie. Le lecteur pourra consulter le travail de Singpurwalla and Wong [15] qui est un des papiers pionniers en la matière et celui de Estévez [2] pour les résultats les plus récents dans ce domaine.

Il est donc tout naturel d'essayer de construire un estimateur de la fonction h^Z en s'inspirant de ces idées. Pour estimer la fonction de répartition conditionnelle et la densité conditionnelle en présence de variable Z fonctionnelle, Ferraty *et al.* [4] ont proposé les estimateurs à noyau fonctionnels suivants:

$$\widehat{F}^z(x) = \frac{\sum_{i=1}^{n} K\left(h_K^{-1} d(z, Z_i)\right) H\left(h_H^{-1}(x - X_i)\right)}{\sum_{i=1}^{n} K\left(h_K^{-1} d(z, Z_i)\right)} \tag{5}$$

et

$$\widehat{f}^z(x) = \frac{\sum_{i=1}^{n} K\left(h_K^{-1} d(z, Z_i)\right) H'\left(h_H^{-1}(x - X_i)\right)}{h_H \sum_{i=1}^{n} K\left(h_K^{-1} d(z, Z_i)\right)}, \tag{6}$$

où K est un noyau, H une fonction de répartition et $h_K = h_{K,n}$ (resp. $h_H = h_{H,n}$) est une suite de nombres réels positifs. Le bon comportement de ces estimateurs, à la fois du point de vue asymptotique et sur le plan appliqué, est mis en évidence dans Ferraty and Vieu [5].

L'estimateur à noyau de la fonction de hasard conditionnelle fonctionnelle h^Z peut donc se construire de la manière suivante:

$$\widehat{h}^Z(x) = \frac{\widehat{f}^Z(x)}{1 - \widehat{F}^Z(x)}. \tag{7}$$

Les hypothèses dont nous aurons besoin ultérieurement concernant les paramètres de cet estimateur, c'est à dire concernant K, H, h_H et h_K, sont peu restrictives. En effet, d'une part elles ne sont pas propres au problème d'estimation de h^Z (mais plutôt inhérentes aux problèmes d'estimation de F^Z et f^Z), et d'autre part elles correspondent aux hypothèses faites habituellement dans le cadre de variables non fonctionnelles. Plus précisément, nous introduirons les conditions suivantes qui garantissent le bon comportement des estimateurs $\widehat{F}^z$ et $\widehat{f}^z$ (voir Ferraty and Vieu [5]):

(H5) Le noyau cummulatif H est dérivable et tel que

i) $\exists A < \infty,\ \forall (x_1, x_2) \in \mathbb{R}^2,\ |H'(x_1) - H'(x_2)| \leq A|x_1 - x_2|$;
ii) H' est de support compact $[-1, 1]$ et $H'(t) > 0,\ \forall t \in [-1, 1]$.

(H6) Le noyau fonctionnel K vérifie les conditions
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 largeur de fenêtre h_K vérifie les conditions
$$\lim_{n\to\infty} h_K = 0 \text{ et } \lim_{n\to\infty} \frac{\log n}{n h_H \phi_x(h_K)} = 0.$$

(H8) La largeur de fenêtre h_H vérifie les conditions
$$\lim_{n\to\infty} h_H = 0 \text{ et } \exists a > 0,\ \lim_{n\to\infty} n^a\, h_H = \infty.$$

Sous ces conditions très générales, nous établirons dans le Paragraphe 3.1 la convergence de l'estimateur à noyau $\widehat{h}^z$ de la fonction de hasard conditionnelle fonctionnelle h^z lorsque les couples de variables $(X_i, Z_i)_{i=1,\ldots,n}$ sont indépendants. Dans le Paragraphe 3.2, ces résultats seront généralisés en s'affranchissant de la condition d'indépendance.

2.4. **Estimation dans le cas censuré**

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 les (T_i, Δ_i, Z_i) où $T_i = \min(X_i,\ C_i)$ et $\Delta_i = I_{X_i \leq C_i}$. Dans la suite nous utiliserons les notations F_1^Z et f_1^Z pour désigner la fonction de répartition et la densité conditionnelles de C sachant Z, et nous utiliserons la notation $S_1^Z = 1 - F_1^Z$. De tels modèles à censure ont été abondamment étudiés dans la littérature pour des variables aléatoires réelles ou multi-dimensionnelles, et dans des cadres non-paramétriques les techniques à noyau sont particulièrement utilisées (voir Tanner and Wong [16], Padgett [11], Lecoutre and Ould-Said [9] et van Keilegom and Veraverbeke [17], pour un échantillon nécessairement non-exhaustif de la littérature dans ce domaine).

L'objectif de ce paragraphe 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. Si l'on introduit les notations $L^Z = 1 - S_1^Z S^Z$ et $\varphi^Z = f^Z S_1^Z$, on peut reformuler l'expression (4) comme

$$h^Z(t) = \frac{\varphi^Z(t)}{1 - L^Z(t)},\ \forall t,\ L^Z(t) < 1. \tag{8}$$

On peut alors définir des estimateurs des fonctions φ^Z et L^Z en posant

$$\widehat{L}^Z(t) = \frac{\sum_{i=1}^{n} K\left(h_K^{-1} d(z, Z_i)\right) H\left(h_H^{-1}(t - T_i)\right)}{\sum_{i=1}^{n} K\left(h_K^{-1} d(z, Z_i)\right)} \tag{9}$$

et

$$\widehat{\varphi}^Z(t) = \frac{\sum_{i=1}^{n} K\left(h_K^{-1} d(z, Z_i)\right) \Delta_i H'\left(h_H^{-1}(t - T_i)\right)}{h_H \sum_{i=1}^{n} K\left(h_K^{-1} d(z, Z_i)\right)}. \tag{10}$$

Finalement l'estimateur de la fonction de hasard est donné par

$$\tilde{h}^Z(t) = \frac{\widehat{\varphi}^Z(t)}{1 - \widehat{L}^Z(t)}. \tag{11}$$

Outre les hypothèses introduites dans le Paragraphe 2.3, nous aurons besoin de conditions supplémentaires. Ces hypothèses, qui sont identiques à celles que l'on retrouve dans la littérature classique pour variables non-fonctionnelles (voir références précédentes), sont les suivantes:

(H9) Conditionnellement à Z, les variables X et C sont indépendantes;

(H10) $\exists A_z < \infty,\ \exists b_1, b_2 > 0,\ \forall (t_1, t_2) \in \mathcal{S}^2,\ \forall (z_1, z_2) \in N_z^2$:

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

(H11) $\exists \mu < \infty,\ \varphi^{z'}(t) < \mu,\ \forall (t, z') \in \mathbb{R}_+ \times N_z$,

(H12) $\exists \eta > 0,\ L^{z'}(t) \leq 1 - \eta,\ \forall (t, z') \in \mathbb{R}_+ \times N_z$.

Sous ces conditions très générales, nous établirons dans le Paragraphe 4.1 les vitesses de convergence de l'estimateur à noyau $\tilde{h}^z$ de la fonction de hasard conditionnelle fonctionnelle h^z lorsque les couples de variables $(X_i, Z_i)_{i=1,\ldots,n}$ sont indépendants. Dans le Paragraphe 4.2 ces résultats seront généralisés en s'affranchissant de la condition d'indépendance.

3. CONVERGENCE DE L'ESTIMATEUR POUR ÉCHANTILLON COMPLET

L'objectif de ce paragraphe est d'établir la convergence presque complète de l'estimateur à noyau $\widehat{h}^Z$ de la fonction de hasard conditionnelle fonctionnelle h^Z lorsque l'échantillon observé n'est pas censuré. Les résultats présentés sont accompagnés par la donnée des vitesses de convergence. Comme nous allons le voir, ces résultats seront relativement faciles à obtenir en ce sens qu'ils vont découler quasi-directement de résultats déjà connus en estimation non-paramétrique de densité ou de fonction de répartition conditionnelle fonctionnelles. Les preuves seront succintes et renverront abondamment à la littérature existante, mais elles seront présentées de sorte à préparer le terrain pour l'obtention des résultats plus délicats dans le cadre d'échantillons censurés (voir Paragraphe 4).

3.1. **Cas d'échantillon i.i.d.**

Nous commençons par l'étude d'échantillons statistiques vérifiant une hypothèse classique d'indépendance, à savoir,

(H13a) Les couples (X_i, Z_i) sont i.i.d.

Théorème 3.1. *Sous les hypothèses* (H1)–(H8) *et* (H13a) *on a*

$$\sup_{x\in\mathcal{S}}|\widehat{h}^z(x)-h^z(x)| = \mathcal{O}\left(h_K^{b_1}\right)+\mathcal{O}\left(h_H^{b_2}\right)+\mathcal{O}\left(\sqrt{\frac{\log n}{n\,h_H\,\phi_z(h_K)}}\right), \quad p.co. \tag{12}$$

Preuve. La preuve est basée sur la décomposition

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

valable pour tout $x\in\mathcal{S}$, ce qui, pour une constante $C<\infty$, amène

$$\sup_{x\in\mathcal{S}}\left|\widehat{h}^z(x)-h^z(x)\right| \le C\,\frac{\left\{\sup_{x\in\mathcal{S}}\left|\widehat{f}^z(x)-f^z(x)\right|+\left|F^z(x)-\widehat{F}^z(x)\right|\right\}}{\inf_{x\in\mathcal{S}}|1-\widehat{F}^z(x)|}. \tag{13}$$

Ainsi de manière classique (voir par exemple la Proposition A6ii de Ferraty and Vieu [5]) le résultat annoncé découle directement des propriétés

$$\sup_{x\in\mathcal{S}}|F^z(x)-\widehat{F}^z(x)| = \mathcal{O}\left(h_K^{b_1}\right)+\mathcal{O}\left(h_H^{b_2}\right)+\mathcal{O}\left(\sqrt{\frac{\log n}{n\,\phi_z(h_K)}}\right), \quad \text{p.co.} \tag{14}$$

et

$$\sup_{x\in\mathcal{S}}|f^z(x)-\widehat{f}^z(x)| = \mathcal{O}\left(h_K^{b_1}\right)+\mathcal{O}\left(h_H^{b_2}\right)+\mathcal{O}\left(\sqrt{\frac{\log n}{n\,h_H\,\phi_z(h_K)}}\right), \quad \text{p.co.} \tag{15}$$

Les résultats (14) et (15) sont des résultats connus (voir par exemple Ferraty and Vieu [5], Propositions 6.19 et 6.20). □

3.2. **Cas d'échantillon dépendant**

Dans de nombreux domaines d'applications, comme en particulier en économétrie (voir par exemple les problèmes traités par Engle and Russell [1] ou Nielson and Linton [10]) ou en sismologie (voir par exemple les données étudiées dans Estévez-Pérez *et al.* [3]), l'hypothèse d'indépendance entre les variables observées n'est pas réaliste. Pour pouvoir étendre les résultats obtenus ci-dessus, il est nécessaire d'introduire une structure probabiliste qui permette de contrôler la dépendance entre les variables constituant l'échantillon statistique. Une manière naturelle de faire consiste à introduire une condition d'indépendance asymptotique. Nous ferons ici l'hypothèse de mélange

(H13b) La suite $(X_i, Z_i)_{i\in\mathbb{N}}$ est α-mélangeante et ses coefficients de mélange $\alpha(n)$ sont tels que $\exists a,\ c \in \mathbb{R}_+^* :\ \forall n \in \mathbb{N},\ \alpha(n) \leq cn^{-a}$.

Afin de contrôler les lois jointes, nous aurons aussi besoin d'introduire les conditions ci-dessous.

(H13c) La densité jointe de (Y_i, Y_j) sachant (Z_i, Z_j) existe et est bornée, et

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

L'hypothèse

$$\text{(H13d)} \qquad \exists \epsilon_2 \in]0,1[,\ a > \frac{1+\epsilon_1}{\epsilon_1\epsilon_2} \quad \text{et} \quad h_H\phi_z(h_K) = \mathcal{O}(n^{-\epsilon_2})$$

permet de maîtriser les liens existant entre les paramètres de lissage h_H, h_K et les coefficients de mélange $\alpha(n)$. Il faut noter que ces hypothèses sont classiques dans les problèmes d'estimation non-paramétrique avec variables dépendantes, fonctionnelles ou non (voir Ferraty and Vieu [5], Chapitre 11). On a le résultat suivant.

THÉORÈME 3.2. *Sous les hypothèses* (H1)–(H8) *et* (H13b)–(H13d), *on a*

$$\text{(16)} \quad \sup_{x\in\mathcal{S}} |\widehat{h}^z(x) - h^z(x)| = \mathcal{O}\left(h_K^{b_1}\right) + \mathcal{O}\left(h_H^{b_2}\right) + \mathcal{O}\left(\sqrt{\frac{\log n}{n\, h_H\, \phi_z(h_K)}}\right), \quad \text{p.co.}$$

Preuve. Les propriétés (14) et (15) restent vraies sous les hypothèses de mélange puisque ce sont des cas particuliers des Propositions 11.22.ii et 11.23.ii de Ferraty and Vieu [5]. Le résultat (16) est donc obtenu directement à partir de la décomposition (13) et des résultats (14) et (15). □

4. **ESTIMATION AVEC DONNÉES CENSURÉES**

L'objectif maintenant est de reprendre ces propriétés asymptotiques dans le cadre plus général d'un échantillon censuré tel que décrit dans le Paragraphe 2.4. Nous allons commençer dans le Paragrapahe 4.1 par traiter le cas d'échantillon i.i.d., puis nous étendrons nos résultats à des données issues d'un processus mélangeant. De toute évidence, l'obtention de ces résultats nécessitera des développements techniques plus sophistiqués que ceux présentés dans le cadre non censuré. Pour assurer une bonne lisibilité de ce Paragraphe 4, la présentation de ces détails techniques se fera ultérieurement lors du Paragraphe 5.

4.1. **Le cas indépendant**

Nous commençons par l'étude d'échantillons statistiques vérifiant une hypothèse classique d'indépendance, à savoir,

(H14a) Les triplets (X_i, C_i, Z_i) sont i.i.d.

THÉORÈME 4.1. *Sous les hypothèses* (H1)–(H11) *et* (H14a), *on a*

$$\sup_{t\in\mathcal{S}} |\widehat{h}^z(t) - h^z(t)| = \mathcal{O}\left(h_K^{b_1}\right) + \mathcal{O}\left(h_H^{b_2}\right) + \mathcal{O}\left(\sqrt{\frac{\log n}{n\, h_H\, \phi_z(h_K)}}\right), \quad p.co. \tag{17}$$

Preuve. Le résultat est basé sur la décomposition ci-dessous, dans laquelle C est une constante réelle strictement positive:

$$\sup_{t\in\mathcal{S}} \left|\widehat{h}^z(t) - h^z(t)\right| \leq C\ \frac{\left\{\sup_{t\in\mathcal{S}} \left|\widehat{\varphi}^z(t) - \varphi^z(t)\right| + \left|L^z(t) - \widehat{L}^z(t)\right|\right\}}{\inf_{t\in\mathcal{S}} |1 - \widehat{L}^z(t)|}, \tag{18}$$

qui s'obtient à partir de (4) et (8) en procédant comme pour établir (13). Puisque $\widehat{L}^Z$ n'est autre que l'estimateur à noyau de la fonction de répartition conditionnelle de T sachant Z, on obtient directement (voir Ferraty and Vieu [5], Proposition 6.19) que

$$\sup_{t\in\mathcal{S}} \left|\widehat{L}^Z(t) - L^Z(t)\right| = \mathcal{O}\left(h_K^{b_1}\right) + \mathcal{O}\left(h_H^{b_2}\right) + \mathcal{O}\left(\sqrt{\frac{\log n}{n\phi_z(h_K)}}\right), \quad \text{p.co.} \tag{19}$$

Les propriétés de l'estimateur $\widehat{\varphi}^Z$ sont données dans le Lemme 4.1 ci-dessous. Finalement, le résultat recherché est obtenu directement à partir de (18), (19) et (20). Le Lemme 4.1 sera prouvé dans le Paragraphe 5. □

LEMME 4.1. *Sous les conditions du Théorème* 4.1, *on a*

$$\sup_{t\in\mathcal{S}} \left|\widehat{\varphi}^Z(t) - \varphi^Z(t)\right| = \mathcal{O}\left(h_K^{b_1}\right) + \mathcal{O}\left(h_H^{b_2}\right) + \mathcal{O}\left(\sqrt{\frac{\log n}{n\, h_H\, \phi_z(h_K)}}\right), \quad p.co. \tag{20}$$

4.2. **Le cas dépendant**

Ces résultats s'étendent aux cas d'échantillons censurés non nécessairement indépendants en introduisant des conditions similaires à celles présentées dans le Paragraphe 3.2. Ces conditions sont les suivantes:

(H14b) la suite $(X_i, C_i, Z_i)_{i\in\mathbb{N}}$ est α-mélangeante et de coefficients de mélange $\alpha(n)$ tels que $\exists a,\ c \in \mathbb{R}_+^*:\ n \in \mathbb{N}\ \alpha(n) \leq cn^{-a}$;

(H14c) la densité jointe de (Y_i, Y_j) sachant (Z_i, Z_j) existe et est bornée, et

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

(H14d) $$\exists \epsilon_2 \in]0,1[,\ a > \tfrac{1+\epsilon_1}{\epsilon_1\epsilon_2} \quad \text{et} \quad h_H\phi_z(h_K) = \mathcal{O}(n^{-\epsilon_2}).$$

Théorème 4.2. *Sous les hypothèses* (H1)–(H12) *et* (H14b)–(H14d)*, on a*

$$\sup_{t\in\mathcal{S}} |\widehat{h}^z(t) - h^z(t)| = \mathcal{O}\left(h_K^{b_1}\right) + \mathcal{O}\left(h_H^{b_2}\right) + \mathcal{O}\left(\sqrt{\frac{\log n}{n\, h_H\, \phi_z(h_K)}}\right), \quad p.co. \tag{21}$$

Preuve. Puisque le résultat (19) reste valable sous condition d'α-mélange (voir Ferraty and Vieu [5], Proposition 11.22.ii), le résultat du Théorème 4.1 s'obtient directement à partir de la décomposition (18) et du Lemme 4.2 ci-dessous qui sera prouvé dans le Paragraphe 5. □

Lemme 4.2. *Sous les conditions du Théorème* 4.2*, on a*

$$\sup_{t\in\mathcal{S}} \left|\widehat{\varphi}^Z(t) - \varphi^Z(t)\right| = \mathcal{O}\left(h_K^{b_1}\right) + \mathcal{O}\left(h_H^{b_2}\right) + \mathcal{O}\left(\sqrt{\frac{\log n}{n\, h_H\, \phi_z(h_K)}}\right), \quad p.co. \tag{22}$$

5. **PREUVES DES LEMMES TECHNIQUES**

Dans ce qui suit C et c désigneront des constantes réelles strictement positives génériques. Par ailleurs, on introduit les notations suivantes:

$$K_i(z) = K\left(h_K^{-1} d(z, Z_i)\right), \quad H_i(t) = H'\left(h_H^{-1}(t - T_i)\right),$$

$$\widehat{\varphi}_N^z(t) = \frac{1}{nh_H EK_1(z)} \sum_{i=1}^n K_i(z) H_i(t) \Delta_i, \quad \widehat{\varphi}_D(z) = \frac{1}{nEK_1(z)} \sum_{i=1}^n K_i(z),$$

$$V_i = \frac{1}{EK_1(z)} K_i(z), \quad W_i = \frac{1}{h_H EK_1(z)} K_i(z) H_i(t) \Delta_i,$$

$$s_n^2 = \sum_{i_1=1}^n \sum_{i_2=1}^n \operatorname{cov}(V_{i_1}, V_{i_2}), \quad S_n^2 = \sum_{i_1=1}^n \sum_{i_2=1}^n \operatorname{cov}(W_{i_1}, W_{i_2}).$$

Preuve du Lemme 4.1. En utilisant la décomposition

$$\widehat{\varphi}^z(t) - \varphi^z(t) = \frac{(\widehat{\varphi}_N^z(t) - \varphi_N^z(t))\,\varphi_D(z) - (\widehat{\varphi}_D(z) - \varphi_D(z))\,\varphi_N^z(t)}{\widehat{\varphi}_D(z)\varphi_D(z)}, \tag{23}$$

et en vertu de la Proposition A6ii de Ferraty and Vieu [5], le résultat du Lemme 4.1 découlera directement des trois propriétés suivantes:

$$|\widehat{\varphi}_D(z) - 1| = \mathcal{O}\left(\sqrt{\frac{\log n}{n\,h_H\,\phi_z(h_K)}}\right), \quad \text{p.co.}, \tag{24}$$

$$\sup_{t\in\mathcal{S}} |E\widehat{\varphi}_N^z(t) - \varphi^z(t)| = \mathcal{O}(h_K^{b_1}) + \mathcal{O}(h_H^{b_2}), \tag{25}$$

et

$$\frac{1}{\widehat{\varphi}_D(z)} \sup_{t\in\mathcal{S}} |\widehat{\varphi}_N^z(t) - E\widehat{\varphi}_N^z(t)| = \mathcal{O}\left(\sqrt{\frac{\log n}{n\,h_H\,\phi_z(h_K)}}\right), \quad \text{p.co.} \tag{26}$$

- *Preuve de* (24). Il suffit de remarquer que l'on peut écrire

$$\widehat{\varphi}_D(z) = \frac{1}{n}\sum_{i=1}^n V_i,$$

avec

$$|V_i| = \mathcal{O}\left(\frac{1}{\phi_z(h)}\right), \tag{27}$$

et

$$EV_i^2 = \mathcal{O}\left(\frac{1}{\phi_z(h)}\right). \tag{28}$$

En appliquant une inégalité exponentielle pour variables bornées (par exemple le Corollaire A9i de Ferraty and Vieu [5]) et en tenant compte des résultats (27) et (28), on arrive à

$$P\left[|\widehat{\varphi}_D(z) - E\widehat{\varphi}_D(z)| > \varepsilon\sqrt{\frac{\log n}{n\,\phi_z(h_K)}}\right] = \mathcal{O}(n^{-C\varepsilon^2}).$$

Il suffit maintenant de choisir ϵ suffisament grand pour obtenir le résultat (24).

- *Preuve de* (25). Pour tout $x \in \mathcal{S}$ on a

$$\begin{aligned} E\widehat{\varphi}_N^z(t) &= \frac{1}{h_H E K_1(z)} E\left(K_1(z)H_1(t)\Delta_1\right) \\ &= \frac{1}{h_H E K_1(z)} E\left(K_1(z)E\left(H_1(t)I_{X_1\le C_1}\big|Z_1\right)\right) \\ &= \frac{1}{h_H E K_1(z)} E\left(K_1(z)E(H_1(t)S_1^{Z_1}(X_1)|Z_1)\right), \end{aligned} \tag{29}$$

la dernière égalité découlant de l'indépendance conditionnelle entre C_1 et X_1 introduite dans (H9). Par ailleurs nous avons

$$E(H_1(t)S_1^z(X_1)|Z_1) = \int H'\left(\frac{t-u}{h_H}\right) S_1^{Z_1}(u) f^{Z_1}(u)\mathrm{d}u \tag{30}$$
$$= h_H \int H'(v)\varphi^{Z_1}(t - vh_H)\mathrm{d}v = h_H \left(\varphi^z(t) + o(h_H^{b_2} + h_K^{b_1})\right),$$

la dernière égalité découlant de la propriété de Lipschitz de la fonction φ^z introduite dans (H10) et du fait que H' est une densité de probabilité. Il faut noter, toujours à cause de la condition (H10), que les $o(\cdot)$ intervenant dans le résultat (30) sont uniformes pour $t \in \mathcal{S}$. Ainsi, le résultat (25) est une conséquence immédiate de (29) et (30).

• *Preuve de* (26). La compacité de l'ensemble $\mathcal{S}$ permet de le recouvrir par u_n intervalles disjoints de sorte que

$$\mathcal{S} \subset \bigcup_{k=1}^{u_n} [\tau_k - l_n, \tau_k + l_n[,$$

où $\tau_1, \ldots, \tau_{u_n}$ sont des points de $\mathcal{S}$ et où l_n et u_n sont choisis tels que

$$\exists C > 0,\ \exists c > 0,\ l_n = C u_n^{-1} = n^{-c}. \tag{31}$$

Pour chaque $t \in \mathcal{S}$ on note τ_t l'unique τ_k tel que $t \in [\tau_k - l_n, \tau_k + l_n[$. Finalement, (26) découlera directement des résultats suivants:

$$\frac{1}{\widehat{\varphi}_D(z)} \sup_{t\in\mathcal{S}} |\widehat{\varphi}_N^z(t) - \widehat{\varphi}_N^z(\tau_x)| = \mathcal{O}\left(\sqrt{\frac{\log n}{n\, h_H\, \phi_z(h_K)}}\right), \tag{32}$$

$$\frac{1}{\widehat{\varphi}_D(z)} \sup_{t\in\mathcal{S}} |E\widehat{\varphi}_N^z(t) - E\widehat{\varphi}_N^z(\tau_x)| = \mathcal{O}\left(\sqrt{\frac{\log n}{n\, h_H\, \phi_z(h_K)}}\right), \quad \text{p.co.}, \tag{33}$$

et

$$\frac{1}{\widehat{\varphi}_D(z)} \sup_{t\in\mathcal{S}} |\widehat{\varphi}_N^z(\tau_x) - E\widehat{\varphi}_N^z(\tau_x)| = \mathcal{O}\left(\sqrt{\frac{\log n}{n\, h_H\, \phi_z(h_K)}}\right). \tag{34}$$

• *Preuve de* (32). A cause de la condition (H5i), il existe une constante finie C telle que pour tout $t \in \mathcal{S}$ on a

$$|\widehat{\varphi}_N^z(t) - \widehat{\varphi}_N^z(\tau_t)| = \frac{1}{nh_H EK_1(z)} \sum_{i=1}^n \Delta_i K_i(z)\left(H_i(t) - H_i(\tau_t)\right) \tag{35}$$
$$\leq \frac{C}{nh_H EK_1(z)} \sum_{i=1}^n K_i(z) \frac{|x - \tau_t|}{h_H} \leq C\widehat{\varphi}_D(z) l_n h_H^{-2}.$$

En utilisant (31) et en choisissant c assez grand, on obtient directement (32).

• *Preuve de* (33). Ce résultat s'obtient directement à partir de (24) et (35) en utilisant la Proposition A6ii de Ferraty and Vieu [5].

• *Preuve de* (34). L'obtention de (34) est basée sur l'utilisation d'une inégalité exponentielle. Plus précisément, il suffit de remarquer que l'on peut écrire

$$\widehat{\varphi}_N^z(t) = \frac{1}{n}\sum_{i=1}^n W_i,$$

avec

$$|W_i| = \mathcal{O}\left(\frac{1}{h_H\phi_z(h)}\right), \tag{36}$$

et

$$\begin{aligned} EW_i^2 &= \frac{1}{h_H^2(EK_1(z))^2}EK_i^2(z)H_i^2(t)\Delta_i^2 \\ &\leq C\frac{1}{h_H^2(EK_1(z)^2)}E\left(K_i^2(z)E(H_i^2(t)\big|Z_i)\right) \\ &\leq C\frac{1}{h_H\phi_z(h)^2}E\left(K_i^2(z)\int\frac{1}{h_H}H'\left(\frac{x-u}{h_H}\right)^2 f^{Z_i}(u)\mathrm{d}u\right) \\ &= \mathcal{O}\left(\frac{1}{h_H\phi_z(h)}\right). \end{aligned} \tag{37}$$

En utilisant la condition (31) on arrive à

$$\begin{aligned} &P\left[\sup_{x\in\mathcal{S}}|\widehat{\varphi}_N^z(\tau_t) - E\widehat{\varphi}_N^z(\tau_t)| > \varepsilon\sqrt{\frac{\log n}{n\,h_H\,\phi_z(h_K)}}\right] \\ &\qquad \leq n^c\max_{j=1,\ldots u_n}P\left[|\widehat{\varphi}_N^z(\tau_j) - E\widehat{\varphi}_N^z(\tau_j)| > \varepsilon\sqrt{\frac{\log n}{n\,h_H\,\phi_z(h_K)}}\right]. \end{aligned} \tag{38}$$

Par ailleurs, en appliquant une inégalité exponentielle pour variables bornées (par exemple le Corollaire A9i de Ferraty and Vieu [5]) et en tenant compte des résultats (36) et (37), on arrive à

$$P\left[|\widehat{\varphi}_N^z(\tau_j) - E\widehat{\varphi}_N^z(\tau_j)| > \varepsilon\sqrt{\frac{\log n}{n\,h_H\,\phi_z(h_K)}}\right] = \mathcal{O}(n^{-C\varepsilon^2}). \tag{39}$$

Il suffit maintenant de choisir ϵ suffisament grand pour obtenir directement le résultat recherché à partir de (38) et de (39).

• Les résultats (32), (33) et (34) suffisent pour conclure la preuve du résultat (26).

Finalement, le Lemme 4.1 découle directement des résultats (24), (25) et (26) et de la décomposition (23). $\square$

Preuve du Lemme 4.2. La preuve suit le même cheminement que la précédente, mais avec des difficultés supplémentaires liées à la non indépendance des variables constituant l'échantillon. Comme précédemment, au vu de la décomposition (23), il suffira de prouver que les résultats (24), (25) et (26) restent valables.

• *Preuve de* (24). La principale étape de la démonstration consiste à obtenir une évaluation de la somme des covariances s_n^2. Pour $i_1 \neq i_2$, d'aprés la condition (H14c) on a

$$|EV_{i_1}V_{i_2}| \leq \frac{C}{(EK_1(z))^2}\phi_z(h_K)^{1+\varepsilon_1} = \mathcal{O}(\phi_z(h_K)^{-1+\varepsilon_1})$$

et, par suite,

$$|\text{cov}(V_{i_1}, V_{i_2})| \leq \frac{C}{(EK_1(z))^2}\phi_z(h_K)^{1+\varepsilon_1} = \mathcal{O}(\max\{\phi_z(h_K)^{-1+\varepsilon_1}, 1\}). \tag{40}$$

D'un autre côté, en utilisant une inégalité de covariance pour processus mélangeant (voir par exemple la Proposition A10i de Ferraty and Vieu [5]) on peut écrire

$$\text{cov}(V_{i_1}, V_{i_2}) \leq C\phi_z(h_K)^{-2}\alpha(|i_1 - i_2|). \tag{41}$$

Finalement, pour toute suite v_n positive on peut ecrire

$$s_n^2 = \sum_{i=1}^{n} \text{var}(V_i) + \sum_{0<|i_1-i_2|\leq v_n} \text{cov}(V_{i_1}, V_{i_2}) + \sum_{|i_1-i_2|>v_n} \text{cov}(V_{i_1}, V_{i_2}), \tag{42}$$

et en utilisant respectivement (28), (40) et (41) pour traiter chacun des trois termes de (42) on arrive à

$$\begin{aligned} s_n^2 =& \mathcal{O}\left(\frac{n}{\phi_z(h_K)}\right) + \mathcal{O}(nv_n \max\{\phi_z(h_K)^{-1+\varepsilon_1}, 1\}) + \\ &+ \mathcal{O}(\phi_z(h_K)^{-2} \sum_{|i_1-i_2|>v_n} \alpha(|i_1 - i_2|)). \end{aligned}$$

Il suffit maintenant de choisir $v_n = \phi_z(h_K)^{-\varepsilon_1}$ pour arriver à

$$\begin{aligned} s_n^2 &= \mathcal{O}\left(\frac{n}{\phi_z(h_K)}\right) + \mathcal{O}(\phi_z(h_K)^{-2}n(n - v_n)\alpha(v_n)) \\ &= \mathcal{O}\left(\frac{n}{\phi_z(h_K)}\right) + \mathcal{O}(\phi_z(h_K)^{-2}n^2\phi_z(h_K)^{a\varepsilon_1}) = \mathcal{O}\left(\frac{n}{\phi_z(h_K)}\right), \end{aligned} \tag{43}$$

la dernière égalité découlant directement de la condition (H14d).

En utilisant les bornes données par (27) et (28), et en appliquant une inégalité exponentielle pour variables bornées mélangeantes (par exemple le Corollaire A13ii de Ferraty and Vieu [5]), on arrive à

$$\widehat{\varphi}_D(z) - E\widehat{\varphi}_D(z) = \mathcal{O}(n^{-1}\sqrt{\log n s_n^2}), \quad \text{p.co.} \tag{44}$$

Le résultat (24) découle directement de (43) et (44).

• *Preuve de* (25). Pour traiter ce terme déterministe, les calculs effectués lors de la preuve du Lemme 4.1 n'utilisaient pas l'indépendance des variables. Le résultat (25) reste donc valable sous conditions de mélange.

• *Preuve de* (26). En reprenant la démarche et les notations introduites lors de la preuve du Lemme 4.1, il suffit de vérifier que les résultats (32), (33) et (34) restent vrais.

• *Preuve de* (32) *et* (33). Les calculs effectués lors de la preuve du Lemme 4.1 n'utilisaient pas l'indépendance des variables. Ainsi, le résultat (35) reste vrai sous condition de dépendance. Par conséquent (32) et (33) sont obtenus directement à partir de (24) et (35) et en utilisant la Proposition A6ii de Ferraty and Vieu [5].

• *Preuve de* (34). La principale étape de la démonstration consiste à obtenir une évaluation de la somme des covariances S_n^2. En reprenant la même démarche que celle utilisée pour prouver (24), pour $i_1 \neq i_2$ on a

$$|EW_{i_1}W_{i_2}| \leq \frac{C}{(h_H^2 EK_1(z))^2}\phi_z(h_K)^{1+\varepsilon_1} = \mathcal{O}(h_H^{-2}\phi_z(h_K)^{-1+\varepsilon_1})$$

et, par suite,

$$\begin{aligned} |\mathrm{cov}(W_{i_1}, W_{i_2})| &\leq \frac{C}{h_H^{-2}(EK_1(z))^2}\phi_z(h_K)^{1+\varepsilon_1} \\ &= \mathcal{O}\left(h_H^{-2}(\max\{\phi_z(h_K)^{-1+\varepsilon_1}, 1\})\right). \end{aligned} \tag{45}$$

D'un autre côté, en utilisant une inégalité de covariance pour processus mélangeant (voir par exemple la Proposition A10i de Ferraty and Vieu [5]) on a

$$\mathrm{cov}(W_{i_1}, W_{i_2}) \leq C h_H^{-2}\phi_z(h_K)^{-2}\alpha(|i_1 - i_2|). \tag{46}$$

Finalement, pour toute suite v_n positive on peut écrire

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

et en utilisant respectivement (28), (45) et (46) on arrive à

$$S_n^2 = \mathcal{O}\left(\frac{n}{h_H\phi_z(h_K)}\right) + \mathcal{O}\big(nv_nh_H^{-2}\max\{\phi_z(h_K)^{-1+\varepsilon_1}, 1\}\big) + \mathcal{O}\Bigg(h_H^{-2}\phi_z(h_K)^{-2}\sum_{|i_1-i_2|>v_n}\alpha(|i_1-i_2|)\Bigg).$$

Il suffit maintenant de choisir $v_n = \phi_z(h_K)^{-\varepsilon_1}$ et d'utiliser la condition (H14d) pour arriver à

$$S_n^2 = \mathcal{O}\left(\frac{n}{h_H\phi_z(h_K)}\right). \tag{47}$$

En utilisant les bornes données par (36) et (37), et en appliquant une inégalité exponentielle pour variables bornées mélangeantes (par exemple le Corollaire A13ii de Ferraty and Vieu [5]), on arrive à

$$\widehat{\varphi}_N^z(\tau_j) - E\widehat{\varphi}_N^z(\tau_j) = \mathcal{O}\left(n^{-1}\sqrt{\log nS_n^2}\right), \quad \text{p.co.} \tag{48}$$

Le résultat (34) découle directement de (38), (47) et (48).

• Ainsi, il suffit d'utiliser (32), (33) et (34) pour terminer la preuve de (26).

Finalement, le Lemme 4.2 découle directement des résultats (24), (25) et (26) et de la décomposition (23). □

Remerciements. Nous tenons à remercier les participants au groupe de travail STAPH en Statistique Fonctionnelle et Opératorielle du Laboratoire de Statistique et Probabilités de Toulouse. Les activités de ce groupe de travail sont accessibles sur la page `http://www.lsp.ups-tlse.fr/staph`.

RÉFÉRENCES

[1] R.F. Engle and J. Russell, *Autoregressive conditional duration: A new model for irregulary spaced transaction data.* Econometrica **66** (1988), 1127–1162.

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

[3] G. Estévez-Pérez, 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.

[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 and P. Vieu, *Nonparametric Functional Data Analysis. Theory and Practice.* Springer Series in Statistics. Springer, New York, 2006.

[6] O. Gefeller and P. Michels, *A review on smoothing methods for the estimation of the hazard rate based on kernel functions.* In: Y. Dodge and J. Whittaker (Eds.), *Computational Statistics*, pp. 459–464. Physica-Verlag, 1992.

[7] S. Hassani, P. Sarda et P. Vieu, *Approche non-paramétrique en théorie de la fiabilité: revue bibliographique.* Rev. Statist. Appl. **35** (1986), *4*, 27–41.

[8] A. Izenman, *Developments in nonparametric density estimation.* J. Amer. Statist. Assoc. **86** (1991), 205–224.

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

[10] J.P. Nielson and O. Linton, *Kernel estimation in a nonparametric marker dependent hazard model.* Ann. Statist. **23** (1995), 1735–1748.

[11] W.J. Padgett, *Nonparametric estimation of density and hazard rate functions when samples are censored.* In P.R. Krishnaiah and C.R. Rao (Eds.), *Handbook of Statistics* **7**, pp. 313–331. Elsevier Science Publishers, 1988.

[12] M. Pascu and I. Vaduva, *Nonparametric estimation of the hazard rate: a survey.* Rev. Roumaine Math. Pures Appl. **48** (2003), 173–191.

[13] P.N. Patil, M.T. Wells and J.S. Maron, *Some heuristics of kernel based estimators of ratio functions.* J. Nonparametr. Statist. **4** (1994), 203–209.

[14] J. Ramsay and B. Silverman, *Functional Data Analysis,* 2nd Ed. Springer Series in Statistics. Springer, New York, 2005.

[15] N. Singpurwalla and M.Y. Wong, *Estimation of the failure rate – a survey of nonparametric models. Part I: Non-Bayesian methods.* Comm. Statist. Theory Methods **12** (1983), 559-588.

[16] M. Tanner and W.H. Wong, *The estimation of the hazard function from randomly censored data by the kernel methods.* Ann. Statist. **11** (1983), 989–993.

[17] I. Van Keilegom and N. Veraverbeke, *Hazard rate estimation in nonparametric regression with censored data.* Ann. Inst. Statist. Math. **53** (2001), 730–745.

Reçu le 2 février 2007

Frédéric Ferraty, Philippe Vieu
Université Paul Sabatier, Institut de Mathématiques
Laboratoire de Statistique et Probabilités
31062 Toulouse, France
ferraty@cict.fr

et

Abbes Rabhi
Université Djillali Liabes
Departement de Mathématiques
Sidi Bel Abbes, Algérie