Remerciements

Ce modeste travail de thèse ne pouvait avoir lieu sans l'aide et la collaboration des gens qui ont été présents durant toute la période de mes études et même avant. Leur soutien m'était d'une importance capitale.

Je tiens à remercier spécialement mon directeur de thèse le Professeur ***Attouch Mohamed Kadi*** d'avoir bien assuré la direction et l'encadrement de mes travaux de thèse. Je vous remercie Mr. pour votre confiance en mes capacités, pour le temps et la patience que vous m'avez accordés tout au long de ces années en me fournissant d'excellentes conditions logistiques. Je garderai dans mon coeur votre générosité, votre compréhension et votre efficacité. Pour tout ce que vous m'avez donné, je vous remercie très sincèrement.

Je voudrais aussi remercier chaleureusement chacun des membres du jury qui me fond le grand honneur d'y participer.

Mes sincères remerciements et ma gratitude vont aussi au Pr. ***Laksaci Ali*** pour avoir accepté de juger ce travail et de présider le jury de soutenance. Soyez assurés de mon entière reconnaissance.

Je remercie vivement monsieur ***Gheriballah Abd El Kader*** pour la confiance dont il me fait preuve en faisant partie de ce jury.

Merci également Messieurs ***Madani Fethi, Benchick Toufik, Chikher Mezzouar Zouaoui*** qui ont accepté de juger mon travail en tant qu'examinateurs. Je leur adresse mes sentiments les plus respectueux.

Je tiens aussi à exprimer ma reconnaissance à tous les membres du laboratoire de Statistique et Processus Stochastiques (LSPS) ainsi les enseignants du département de Probabilité et Statistiques.

Dédicace

Afin d'être reconnaissante envers ce qui m'ont appuyée et encouragée à effectuer ce travail de recherches, je dédie ce mémoire :

A mes très chers parents " Mama et Papa"

Affables, aimables, honorables : vous présentez pour moi le symbole de la bonté par excellence, la source de tendresse et l'exemple du dévouement, vous qui n'avez pas cessé de m'encourager et de prièr pour moi.

Aucune dédicace ne serait assez éloquente pour exprimer ce que vous méritez pour tous les sacrifices que vous n'avez cessé de me donner depuis ma naissance, durant mon enfance et même à l'âge adulte.

Je vous dédie ce travail en témoignage de mon profond amour. Puisse Dieu, le tout puissant, vous préserver et vous accorder santé, longue vie et bonheur.

A mon très cher mari "Taher"

Quand je t'ai connu, j'ai trouvé l'homme de ma vie, mon âme soeur et la lumière de mon chemin.

Tes sacrifices, ton soutien moral et matériel, ta gentillesse sans égal, ton profond attachement m'ont permis de réussir mes études.

Que dieu réunisse nos chemins pour un long commun serein et que ce travail soit témoignage de ma reconnaissance et de mon amour sincère et fidèle.

A mon très cher frère "Yacine" et à mes très chères soeurs "Akila" et "Maroua"

Mon très cher frère et mes très chères soeurs présents dans tous mes moments d'examens par leur soutien moral et matériel.

Je vous souhaite un avenir plein de joie, de bonheur, de réussite et de sérénité.

Je vous exprime à travers ce travail mes sentiments de fraternité et d'amour.

A tous ceux qui ont une relation de prés ou de loin avec la réalisation du présent rapport.

Table des matières

Résumé

Dans cette thèse, nous nous intéressons à l'estimation non paramétrique de la fonction de hasard conditionnelle par la méthode des k plus proches voisins (k-NN) d'une variable aléatoire à réponse réelle conditionnée par une variable aléatoire explicative fonctionnelle.

Dans un premier temps ; on considère une sequence (X_i, Y_i) indépendantes et identiquement distribuées à valeur dans $\mathbb{E} \times \mathbb{R}$, ou $\mathbb{E}$ est un espace semi métrique de dimension infini. Nous donnons un estimateur de notre fonction conditionnelle et nous montrons, sous des hypothèses générales, sa convergence presque complète (a.com.) avec taux de convergence, aussi nous étudions la normalité asymptotique de ce modèle. L'efficacité de cet estimateur est donnée par comparaison avec l'estimateur à noyau introduit par FERRATY et al.[12] et LAKSACI ET MECHAB [15].

La deuxième partie du travail est une généralisation des résultats obtenus dans la première partie au cas des données α-mélangeantes.

Notre étude est illustrée par quelques applications sur des données simulées et réelles.

Summary

In this thesis, we are interested in the nonparametric estimate of the conditional hazard function with the k nearest neighbors (k-NN) method for a scalar response variable given a random variable taking values in a semi-metric space.

In the First part, we consider (X_i, Y_i) an independent sequence identically distributed which is a random pair valued in $\mathbb{E} \times \mathbb{R}$ where $\mathbb{E}$ is a semi-metric space. We give an estimator for the conditional hazard function. Under general assumptions, we prove its almost complete convergence (a.com.) with rate of convergence and the asymptotic normality of this model is given. Effectiveness of this estimator is studied by the comparison with the kernel estimator given in Ferraty et al.[12] and Laksaci and Mechab [15].

In the second part of work, we propose the same preceding results with α mixing data.

Our study is illustrated by some applications on simulated and real data.

Chapitre 1

Introduction générale

1.1 Note bibliographique sur la statistique non paramétrique fonctionnelle

La branche de la statistique fonctionnelle constitue un champs de recherches d'actualité ; à la fois diversifié par ses aspects fondamentaux et par les différents domaines qu'elle recoupe ; statistique paramétrique dans laquelle on suppose que la distribution suit un certain modèle décrit par un nombre fini de paramètres, et la statistique non paramétrique, qui est basée sur l'idée de ne pas faire des hypothèses sur la distribution.

En effet, la statistique non paramétrique connaît un grand essor chez de nombreux auteurs et dans des différents domaines. La preuve de ce succès les plusieurs publications scientifiques données dans ce sujet. Les premiers travaux dans ce domaine datent des années 50, par Rao et Tucker (1958) qui ont étudié l'estimation de maximum de vraissemblance pour la distribution multinomiale.

Notons que le modèle le plus fréquemment rencontré en statistique non paramétrique est le modèle de régression qui décrit la relation entre deux ou plusieurs variables aléatoires. Ce modèle a été déjà pris en considération par plusieurs auteurs. Les propriétés et les résultats obtenus de l'estimateur de la fonction de régression dans le cas des variables indépendantes et identiquement distribuées (i.i.d) par Nadaraya (1964). Watson (1964) a établi la convergence uniforme de cet estimateur, la convergence uniforme presque sure est obtenue par Devroye (1978) et la normalité asymptotique du même estimateur a été établie par Roussas (1989). Dans la même année Gyorfie a obtenu les résultats asymptotiques de l'estimateur de la fonction de régression sur des processus α mélangeants, Vieu (1991) a donné les termes asymptotiques exactes de l'erreur quadratique de l'estimateur à noyau de la fonction de régression.

Par ailleurs, Ferraty et Vieu (2000) ont donné les premiers résultats dans le cas fonctionnel. Ces résultats ont été développés en (2002) par Ferraty et Laksaci et al. en traitant le problème de prévision sur les processus à temps continu. Plusieurs auteurs ont généralisé ces résultats ; Masri (2005) a étudié la normalité asymptotique de l'estimateur de la fonction de régression sous la condition de dépendance, la convergence en moyenne quadratique du même estimateur a été montrée par Laksaci (2007).

Signalons aussi que nombreux auteurs ont traité d'autres cas des variables dans l'estimation de la fonction de régression. Citons Ould-Said (2012) qui a étudié la convergence uniforme et la normalité asymptotique avec des variables censurées. Sous les conditions ergodiques, Laib et Louani (2010-2011) ont étudié l'estimation de la fonction de régression fonctionnelle.

Les modèles conditionnels ont pris leurs places dans la statistique ; la fonction de répartition conditionnelle, la densité conditionnelle, le quantile conditionnel, le mode conditionnel et la fonction de hasard conditionnelle. A ces sujets, les premiers résultats ont été réalisés au debut des années soixante par Roussas (1968) qui a étudié les propriétés asymptotiques de l'estimateur à noyau de la fonction de répartition conditionnelle où il a montré la convergence en probabilité. Youndjé (1993) a donné l'estimateur de la fonction de densité conditionnelle avec des données indépendantes et dépendantes. Ainsi l'intérêt que revêt le mode conditionnel dans le domaine de la prévision, plusieurs auteurs s'en sont intéressés, nous pouvons citer Samanta et Thavaneswaren (1990) qui ont étudié l'estimation non paramétrique du mode conditionnel et ils ont établi la normalité asymptotique de l'estimateur de ce modèle avec des données indépendantes et identiquement distribuées, le cas α mélange a été étudié par Ould-Said (1993), puis il a étendu ce dernier travail en (1997) pour des données ergodiques. Louani et Ould-Said (1998-1999) ont montré la normalité asymptotique de l'estimateur du mode conditionnel avec des données censurées et des données fortement mélangeantes.

Dans ces dernières années, l'analyse des données fonctionnelles consiste à traiter les données sous forme de courbes. cet outil informatique nous a permis de traiter de grands ensembles de données. Dans ce contexte, on trouve plusieurs exemples de types des données, les données biométriques utilisés Altman et Back (1994) et Presnell (1998). On trouve aussi les données des ondes radar dans Poskitt et Presnell (2001). Les données les plus utilisées récemment ; les données spectrométriques en chimiométrie lesquelles ont été utilisées par Frank et Friedman (1993), Ferraty et Vieu (2002), Abraham Matzner et Molinari (2003).D'autres exemples peuvent être trouvés dans les livres ; Ramsay et Silverman (2002), Dabo Niang et Ferraty (2002).

La statistique non paramétrique fonctionnelle des fonctions conditionnelles (mode conditionnel, densité conditionnelle ...) trouve auprès des chercheurs un intérêt grandissant cela est traduit par les nombreuses publications et applications données, ainsi les ouvrages qui sont devenus des références dans ce domaine. Nous citons par exemple : les monographies de Ramsay et Selvarman (2002 et 2005), Ferraty et Vieu (2006) qui ont présenté l'estimation non paramétrique fonctionnelle à noyau de la fonction de répartition conditionnelle, la fonction de densité conditionnelle, le mode et le quantile conditionnel avec des données indépendantes et identiquement distribuées et α mélangeantes dont ils ont établi la convergence presque complète avec le taux de convergence. La normalité asymptotique de ces estimateurs est donnée par Ezzahraoui et Ould Said (2008).

Quelques auteurs ont généralisé les travaux de Ferraty et Vieu (2006) afin d'obtenir d'autres propriétés asymptotiques des différents estimateurs sous d'autres conditions et hypothèses. Par exemple Laksaci et Maref (2009-2011) ont donné l'estimation non paramétrique fonctionnelle du quantile conditionnel avec des variables spacialement dépendantes et α mélangeantes. Récemment en (2010), Ferraty et Tadj et al. ont présenté la version uniforme de la vitesse de convergence obtenu par Ferraty et Vieu (2006). Dans le même contexte, Dabo-Niang et Laksaci (2009) ont établi la convergence en norme L^P pour le mode et le quantile conditionnel.

1.2 Note bibliographique sur l'estimation par la méthode des k plus proches voisins

Pendant ces dernières années, un large éventail de littérature dans le domaine de l'estimation par la méthode des k plus proches voisins (k-NN) est fourni par les revues bibliographiques à cause de nombreux avantages qu'elle présente.
Le premier intérêt de cette méthode vient de la nature du paramètre de lissage. **En effet, dans la méthode à noyau traditionnelle le paramètre**

de lissage est la fenêtre h qui est un nombre réel positif. Cependant dans notre méthode, on remplace ce dernier par une variable aléatoire réelle (H_n) pour la variable explicative fonctionnelle (X). D'autre plusieurs aspects de cette méthode ; dans le cadre fonctionnel, elle respecte la structure locale des données, ce qui est primordial en dimension infini. Elle est couramment utilisée en pratique comme en Ferraty et Vieu (2006) et s'avère simple à manipuler parce que l'utilisateur n'a qu'un seul paramètre à contrôler (le nombre k de plus proche voisin), ce paramètre k prend ses valeurs dans un ensemble fini. De plus, cette méthode permet de construire en tout point un voisinage adapté aux données.
L'objectif de cette thèse est d'apporter une premiere justification théorique et pratique à l'utilisation courante de la méthode de k-NN en estimation non paramétrique fonctionnelle de la fonction de hasard conditionnelle.
La bibliographique de l'estimation par la méthode k-NN existe depuis Royall (1966) et Stone (1977) qui ont commencé par estimer la fonction de regression. D'autre auteurs avaient aussi l'intérêt d'étudier cet estimateur, trouvons par exemple Collomb (1980) qui a montré les différents types de convergence (probabilité et presque complète), Mack (1981) a étudié la convergence en L^2 et la distribution asymptotique, la convergence uniforme est donnée par Devroye (1978) et (1981). Pour le cas des données fonctionnelles, Ferraty et Vieu (2004) ont commencé par une introduction sur l'estimation des k-NN, Burba et al. (2008) ont obtenu la convergence presque complète de l'estimateur de la fonction de regression avec des données indépendantes et identiquement distribuées. Attouch et Benchikh (2012) ont établi la normalité asymptotique de la fonction de regression.

1.3 Note bibliographique de la fonction de hasard conditionnelle

1.3.1 Motivation et position du problème

La fonction de hasard conditionnelle demeure un outil indispensable dans l'analyse de survie et dans d'autres plusieurs domaines. Pour cela, l'estimation de cette fonction est une question très importante en statistique, ce sujet doit être abordé en plusieurs angles ; selon la complexité du problème posé et selon les hypothèses et les données proposées : les variables observées sont indépendantes et identiquement distribuées (Phénomène trouvé dans le domaine économique), les variables observées sont dépendantes (Phénomène trouvé dans le domaine seïsmologie) ou bien les variables sont proposées censurées(Phénomène trouvé dans le domaine médical).

Formellement, supposons Y une variables aléatoire à valeur dans $\mathbb{R}^+$ associée à une durée de vie. La fonction de hasard, appelée aussi le taux de hasard, taux de survie ou taux de défaillance, est définie au point y comme étant la probabilité instantanée que cette durée de vie se termine à l'instant y. Plus précisément, si h est la fonction de hasard, nous aurons :

$$h(y) = \lim_{\Delta y \to 0} \frac{\mathbb{P}(Y \leq y + \Delta y / Y \geq y)}{\Delta y}; \quad y > 0. \tag{1.1}$$

Lorsque la variable Y possède une densité f par apport à la mesure de Lebesgue, donc l'égalité (1.1) s'écrit :

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

avec : F est la fonction de répartition de Y.

Par la même manière, la fonction de hasard conditionnée à une variable explicative X est définie comme suite :

$$h^X(y) = \lim_{\Delta y \to 0} \frac{\mathbb{P}(Y \leq y + \Delta y / Y \geq y, X)}{\Delta y}; \quad y > 0 \tag{1.2}$$

à partir de la fonction de densité conditionnelle $f^X(y)$ et de la fonction de répartition conditionnelle $F^X(y)$, cette dernière s'écrit sous la forme :

$$h^X(y) = \frac{f^X(y)}{1 - F^X(y)}.$$

1.3.2 Contexte bibliographique

Le rôle de la fonction de hasard conditionnelle dans des différents domaines scientifiques (médecine, biologie, fiabilité...) c'est de laisser plusieurs auteurs s'y intéresser. Le premier travail dans ce contexte est donné par Watson et Leadbetter en (1964) qui ont proposé une introduction sur l'estimation de la fonction de hasard. Les résultats sur l'estimateur de la fonction de hasard ont été étudiés par nombreux auteurs ; nous citons Hassani et al. (1986), Roussas (1989) et Gefeller et Michels et récemment en (2007) Li et Tran. La littérature a prêté une certaine attention à l'estimation non paramétrique de la fonction de hasard conditionnelle avec des données fonctionnelles ; en (2008) Ferraty et al. ont établi la convergence presque complète de l'estimateur de la fonction de hasard conditionnelle dans le cas de l'indépendance des variables. Le cas de dépendance est donné dans Quintela-del-Río (2008), il a traité la convergence presque complète, la convergence en moyenne quadratique et la normalité asymptotique du même estimateur. Récemment la version uniforme de la convergence presque complète est obtenu par Ferraty et al. (2010). Laksaci et Mechab (2010) ont atteint les mêmes résultats dans le cas spacial.

De nombreuses techniques ont été étudiées dans la littérature pour donner les différents résultats sous différentes hypothèses proposées, mais toutes ne traitent que des variables explicatives réelles et multi-dimensionnelles et elles ne donnent que l'estimateur à noyau (kernel estimator). Contrairement, dans cette thèse, nous nous sommes intéressés à construire l'estimateur des k plus proches voisins (k-NN) de la fonction de hasard conditionnelle avec une variable explicative fonctionnelle et nous avons le but d'étudier les propriétés asymptotiques (la convergence presque complète et la normalité asymptotique) de cet estimateur, rappelons que la convergence presque complète de l'estimateur à noyau a été établi par Ferraty et al. (2010) et la normalité asymptotique a été traitée par Quintela-del-Río (2008).

1.4 Plan de la thèse

Notre but porte essentiellement sur l'estimateur de la fonction de hasard conditionnelle par une méthode utilisée récemment ; la méthode des k plus proches voisins en présence d'une variable explicative fonctionnelle.

Nous avons scinde l'ensemble de notre thèse en trois chapitres. Dans ce premier chapitre introductif, nous donnons des contextes bibliographiques sur l'estimation non paramétrique fonctionnelle et la fonction de hasard conditionnelle.

Dans le deuxième chapitre, nous considérons une suite d'observations indépendantes et identiquement distribuées et nous construisons un estimateur par la méthode des k plus proches voisins de la fonction de hasard conditionnelle à partir des estimateurs de la fonction de répartition et de la densité conditionnelles et nous étudions les propriétés asymptotiques de cet estimateur. A titre illustratif, nous donnons des exemples d'applications sur des données simulées et des données réelles.

Le troisième chapitre traite le problème de l'estimation non paramétrique de la même fonction sous des conditions de dépendence faible (mélange fort). Nous nous intéressons à étendre la convergence presque complète avec taux de convergence et la normalité asymptotique de l'estimateur des k plus proches voisins. **L'applicabilité de notre étude est discutée dans ce chapitre, où nous montrons que notre méthode est très utile et efficace pour l'utilisation en futur**.

1.5 Brève présentation de cadre général de la thèse

Soient $(X_i, Y_i)_{i=\overline{1,n}}$ n paires indépendantes et identiquement distribuées (i.i.d) (respect. dépendantes) comme (X, Y) et à valeur dans $\mathbb{E} \times \mathbb{R}$ où $(\mathbb{E}, d)$ est un espace semi métrique et non nécessairement de dimension fini. Nous nous intéressons à la fonction de hasard conditionnelle définie comme suit :

$$h^X(Y) = \frac{f^X(Y)}{1 - F^X(Y)}, \tag{1.3}$$

ou : $f^X(Y)$: est la fonction de densité conditionnelle.
$F^X(Y)$: est la fonction de répartition conditionnelle.

L'objectif ici est donc d'estimer la fonction $h^X(Y)$. Pour tout $x \in \mathcal{E}$, l'estimateur par la méthode des k plus proches voisins de la fonction 2.1 est donné par :

$$\widehat{h}^x_{k-NN}(Y = y) = \widehat{h}^x(y) = \frac{\widehat{f^x}(y)}{1 - \widehat{F}^x(y)} \tag{1.4}$$

avec : $\widehat{F}^x(y)$: est l'estimateur de la fonction de répartition conditionnelle défini par :

$$\widehat{F}^x(y) = \frac{\sum_{i=1}^{n} K(H_n^{-1} d(x, X_i)) R(g_n^{-1}(y - Y_i))}{\sum_{i=1}^{n} K(H_n^{-1} d(x, X_i))}; \forall y \in \mathbb{R} \tag{1.5}$$

$\widehat{f}^x(y)$: est l'estimateur de la fonction de densité conditionnelle conclu de (2.3)par :

$$\widehat{f}^x(y) = \frac{\sum_{i=1}^{n} K(H_n^{-1} d(x, X_i)) g_n^{-1} R'(g_n^{-1}(y - Y_i))}{\sum_{i=1}^{n} K(H_n^{-1} d(x, X_i))} : \tag{1.6}$$

où :
K : un noyau asymétrique.
R : est une fonction de distribution.
$(g_n)_{n\in\mathbb{N}}$: une sequence de nombres réels strictement positifs (depend de n).
H_n est une variable aléatoire positive définie par :

$$H_n(x) = \min\{h \in \mathbb{R}^+ / \sum_{i=1}^{n} \mathbb{1}_{B(x,h)}(X_i) = k\} \tag{1.7}$$

avec :

$$B(x, h) = \{x' \in \mathcal{E}; d(x, x') < h\}.$$

1.6 Brève présentation des résultats

Résultats : les propriétés asymptotiques de l'estimateur dans le cas d'indépendance des variables

Theorem 1.6.1 *Sous certaines conditions, nous aurons pour la convergence presque complète :*

$$\lim_{n\to\infty} \widehat{h}^x(y) = h^x(y). \qquad a.co.$$

Theorem 1.6.2 *Sous certaines conditions, nous aurons :*

$$\widehat{h}^x(y) - h^x(y) = O\left(\varphi_x^{-1}\left(\frac{k}{n}\right)^{\alpha}\right) + O\left(g_n^{\beta}\right) + O\left(\sqrt{\frac{\log n}{k_n g_n}}\right) \qquad a.co.$$

Theorem 1.6.3 *Nous aurons pour la normalité asymptotique de l'estimateur de k-NN, sous certaines hypothèses proposées :*

$$\left(\frac{k_n g_n}{\sigma_h^2(x,y)}\right)^{1/2}\left[\widehat{h}^x(y)-h^x(y)\right]\xrightarrow{\mathcal{D}}\mathcal{N}(0,1)\qquad as\ n\to\infty \tag{1.8}$$

où :

$$\sigma_h^2(x,y)=\frac{\alpha_2 h^x(y)}{\alpha_1^2\left(1-F^x(y)\right)}\quad (with:\alpha_j=K^j(1)-\int_0^1 (K^j)'(s)\lambda(s)\,ds\ \ for\ j=1,2) \tag{1.9}$$

$$\mathcal{A}=\{x\in\mathcal{E},\ f^x(y)[1-F^x(y)]\neq 0\}$$

$\xrightarrow{\mathcal{D}}$ *: la convergence en distribution.*

Les hypothèses imposées et les preuves des résultats ci-dessus seront données dans le chapitre 2.

Résultats ; la convergence presque complète et la normalité asymptotique où les données sont α mélangeantes

Theorem 1.6.4 *Sous les hypothèses détaillées dans le chapitre 3, on aura :*

$$\lim_{n\to\infty}\widehat{h}^x(y)=h^x(y).\qquad a.co.$$

Theorem 1.6.5 *Sous certaines conditions, nous aurons :*

$$\widehat{h}^x(y)-h^x(y)=O\left(\varphi_x^{-1}\left(\frac{k}{n}\right)^{\alpha}\right)+O\left(g_n^{\beta}\right)+O\left(\sqrt{\frac{\log n}{k_n g_n}}\right)\qquad a.co.$$

Theorem 1.6.6 *Nous aurons Sous les hypothèses détaillées dans le chapitre 3 :*

$$\left(\frac{k_n g_n}{\sigma_h^2(x,y)}\right)^{1/2}\left[\widehat{h}^x(y)-h^x(y)\right]\xrightarrow{\mathcal{D}}\mathcal{N}(0,1)\qquad as\ n\to\infty \tag{1.10}$$

avec :

$$\sigma_h^2(x,y)=\frac{\alpha_2 h^x(y)}{\alpha_1^2\left(1-F^x(y)\right)}\quad (with:\alpha_j=K^j(1)-\int_0^1 (K^j)'(s)\lambda(s)\,ds\ \ for\ j=1,2) \tag{1.11}$$

$$\mathcal{A}=\{x\in\mathcal{E},\ f^x(y)[1-F^x(y)]\neq 0\}$$

$\xrightarrow{\mathcal{D}}$ *: la convergence en distribution.*

Les démonstrations de ces théorèmes sont détaillées dans le chapitre 3.

Bibliographie

[1] Attouch, M. et Benchikh, T. Asymptotic distribution of robust k-nearest neighbour estimator for functional nonparametric models. *Mathematic Vesnic*, **64**, No. 4. (2012). pp. 275-285.

[2] Berlinet, A., Gannoun, A. et Matzner-lober, E. Asymptotic normality of convergent estimate of the conditionel quantile. *Statistics* 35 (2001), 139-169.

[3] Burba, F. ; Ferraty, F. et Vieu, P. Convergence de l'estimateur à noyau des k plus proches voisins en régression fonctionnelle non-paramétrique. *Preprint submitted to Elsevier Science.* (2008).

[4] Dabo-Niang, S. Estimation de la densite dans un espace de dimenssion infini : Application au diffusions. *Aca. Sci., Paris* (3) (2002), 213-216.

[5] Dabo-Niang, S. and Laksaci, A. Estimation non paramétrique de mode conditionnel pour variable explicative fonctionnelle. *Math Aca. Sci., Paris* (1) (2007), 49-52.

[6] Devroye, L.P. *The uniform convergence of nearest neighbour regression function estimators and their application in optimization*, IEEE Trans. Inform. Theory, **24**. (1978). 142-151.

[7] Eddy, W.F. The asymptotic distribution of kernel estimators of the mode. *Giebete* 59 (1982), 227-290.

[8] Ezzahrioui, M. Prévision dans les modèles conditionnels en dimission infinie. Thèse de doctorat, Université du Littoral Côte d'Opale.

[9] Ezzahrioui, M. and Ouled-Said, E. Asymptotic normality of a nonparametric estimator of the conditional mode function for functional data. *J. Nonparametric stat.* 20 (2008), 03-18.

[10] Ferraty, F. ; Rabhi, A. et Vieu, P. Estimation non parametrique de la fonction de hasard avec variable explicative fonctionnelle. *Roumaine Math*, (2008) **53**, 1–18.

[11] Ferraty, F., Tadj, A., Laksaci, A. and Vieu, P. Rate of uniforme consistency for nonparametric estimates wwith function variables. *J. Statist. Plan. and Inf.* 140 (2010), 335-352.

[12] Gannoun, A. Estimation non paramétrique de la médiane conditionnelle. *Pub. Inst. Stat. Univ. Paris* 35 (1990), 11-22.

[13] Gefeller, O. et Michels, P. A review on smoothing methods for the estimation of the hazard rate based on kernel function. *Computational Statistics* (1992), 459-464.

[14] Geffroy, J. Sur l'estimation d'une densité dans un espace métrique. *Aca. Sci., Paris* 278 (1974), 1449-1452.

[15] Hall, P. and Yao, Q. Methods for estimating a conditional distribution function. *J. Amer. Statist. Assoc.* 94 (1990), 154-163.

[16] Hssani, S. ; Sarda, P. et Vieu, P. Approche non paramétrique en théorie de la fiabilité. *Rev. Statist* 35 (1986), 27-41.

[17] Hyndman, R.J., and Yao, Q. Nonparametric estimation and symmetry tests for conditional density function. *J. Nonpametr. Stat.* 14 (2002), 259-278.

[18] Laksaci, A. Convergence en moyenne quadratique de l'estimateur à noyau de la densité conditionnelle avec variable explicative fonctionnelle. *Pub. Inst. Stat. Univ. Paris* 3 (2007), 69-80.

[19] Laksaci, A.et Mechab, B. Conditional hazard estimate for functional random fiels. *Communications Statistics-theory and methods.* (2010). Accepted paper.

[20] Li, J. et Tran, L.T. Hazard rate estimation on random fields. *J. Multivariate Anal.* 98 (2007). 1337Ű1355.

[21] Louani, D. On the asymptotic normality of the function and its derivatives under censoring. *Comm. Stat. Theory and methods* 27 (1998), 2909-2924.

[22] Louani, D. and Ouled-Said, E. Asymptotic normality of kernel estimators of the conditional mode under strong mixing hypothesis. *J. Nonparametric stat.* 11 (1999), 413-442.

[23] Mack, Y. P. Local properties of k-NN regression estimates. *SIAM J. Algebr Discrete Methods*, **2**. (1981). 311-323.

[24] Ouled-Said, E. Estimation non paramétrique du mode conditionnel. Application à la prévision. *Acad. Sci. Paris.* 36 (1993), 943-947.

[25] Ouled-Said, E. A note on ergodic processes prediction via estimation of the conditional mode function. *Scand. J. Stat.* 24 (1997), 231-239.

[26] Pascu, M. et Vaduva, I. Nonparametric estimation of the hazard rate. *Rev. Roumaine Math.* 48 (2003), 173-191.

[27] Quentela Del Rio, A., Vieu, P. A nonparametric conditional mode estimate. *J. Nonparam. Stat.* 8 (1997), 253-266.

[28] Quintela-del-rio (2008). Hasard function given a fonctional variable : nonparametric estimation under strong mixing conditions. *J. Nonparametric stat.*. 413-430.

[29] Ramsay, J. and Silverman, B. *Applied functional data analysis.* Springer-Verlag, New York (2002).

[30] Ramsay, J. and Silverman, B. *Applied functional data analysis*(Second Edition). Springer-Verlag, New York (2005).

[31] Roussas, G. On some properties of nonparametric estimates of probabilities density functions. *Soc. Math. Grèce* 9 (1968), 29-43.

[32] Roussas, G. Nonparametric estimation in mixing sequences of random variables. *J. Statist. Plann.*, 18 (1989) 135–149.

[33] R.M. Royall. (1966). *A class of nonparametric estimates of a smooth regression function*, Ph.D. Diss., Stanford University.

[34] Samanta, M. Nonparametric estimation of conditional quantiles. *Stat. Probability Letters* 7 (1989), 407-412.

[35] Samanta, M. and Thavaneswaran, A. Nonparametric estimation of conditional mode. *Stat. Theory and Math.* 16 (1990), 4515-4524.

[36] Singpurwalla, N. et Wong, M.Y. Estimation of the failure rate - a survey of nonparametric models. *Comm. Statist.* 12 (1983), 989-993.

[37] Stone, C. J. Consistent nonparametric regression. *Ann. Statist*, **5**. (1977). 595-645.

[38] Youndjé, E. Estimation non paramétrique de la densité conditionnelle par la méthode du noyau. PhD Thesis from the Rouen University.(1993).

[39] Youndjé, E. Propriétés de convergence de l'estimateur à noyau de la densité conditionnelle. *Rev. Roumaine Math. Pures Apppl.* 41 (1996), 535-566.

Chapitre 2

The k-nearest neighbors estimation of the conditional hazard function for functional data

Ce chapitre fait l'objet d'une publication dans *Journal REVSTAT*

The k-nearest neighbors estimation of the conditional hazard function for functional data

Mohammed Kadi Attouch , Fatima Zohra Belabed
Lab. de statistique et processus stochastique,
Univ. Djillali Liabès, Sidi Bel Abbès,
BP 89, Sidi Bel Abbès 22000, Algeria

abstract : In this paper, we study the nonparametric estimator of the conditional hazard function using the k nearest neighbors (k-NN) estimation method for a scalar response variable given a random variable taking values in a semi-metric space. We give the almost complete convergence (its corresponding rate) of this estimator and we establish the asymptotic normality. Then the effectiveness of this method is exhibited by a comparison with the kernel method estimation given in Ferraty et al. ([12]) and Laksaci and Mechab ([15])in both cases simulated data and real data.

Keywords : Functional data ; nonparametric regression ; k-NN estimator ; the conditional hazard function ; rate of convergence ; random bandwidth ; asymptotic normality.

Ams : 62G05 · 62G08 · 62G20 · 62G35.

2.1 Introduction

The conditional hazard function remains an indispensable tool in survival analysis and many other fields (medicine, reliability or seismology).
The nonparametric estimation of this function in the case of multivariate data is abundant. The first works date back to Waston and Leadbetter ([25]), they introduce the hazard estimate method, since, several results have been developed, see for example, Roussas ([26]) (for previous works), Li and Tran ([18]) (for recent references). The literature has paid quite some attention to nonparametric hazard rate estimation when the data are functional. The first work which deals with this question is Ferraty et al.([12]). They established the almost complete convergence of the kernel estimate of the conditional hazard function in the independent case. This result was extended to the dependent case by Quintela-del-Río ([21]), he treats the almost complete convergence, the mean quadratic convergence and the asymptotic normality

of this estimate. The uniform version of the almost complete convergence (with rate) in the i.i.d. cas was obtained by Ferraty et al. ([10]). Recently, Laksaci and Mechab ([16]) consider the spatial case. The almost complete convergence rate of an adapted estimate of this model are given.

Estimating the conditional hazard function is closely related to the conditional density, and for the last one, the bandwidth selection is very important for the performance of an estimate. The bandwidth must not be too large, so as prevent over-smoothing,

i.e. substantial bias, and must not be too small either, so as prevent detecting the underlying structure. Particularly, in nonparametric curve estimation, the smoothing parameter is critical for performance.

Starting from this point of view, this work deals with the nonparametric estimation with k nearest neighbors method k-NN, more precisely we consider a kernel estimator of the hazard function constructed from a local window to take into account the exact k nearest neighbors with real response variable Y and functional curves X.

The k nearest neighbor or k-NN estimator is a weighted average of response variables in the neighborhood of x. The existent bibliography of the k-NN method estimation dates bake to Royall ([27]) and Stone ([30]) and has received, since, continuous developments (Mack ([20]) derived the rates of convergence for the bias and variance as well as asymptotic normality in the multivariate case, Collomb ([5]) studied different types of convergence (probability, a.s, a.co) of the estimator of the regression function. Devroye ([6]) obtained the strong consistency and the uniform convergence. For the functional data studies, the k-NN kernel estimate was first introduced in the monograph of Ferraty and Vieu ([13]), Burba et al. ([3]) obtained the rate of almost complete convergence of the regression function using the k-NN method for independent data and the asymptotic normality of robust nonparametric regression function was established in Attouch and Benchikh ([1]).

This paper is organized as follows. In Section 2 we present the model and the k-NN estimator. Section 3, is dedicated to fix notations, hypotheses and the presentation of the main results, the almost complete convergence and the asymptotic normality. Section 4 is devoted to some applications in several problems of nonparametric statistics. Some technical auxiliary results are deployed in Section 5, subsequently, in Section 6, we show the proofs of our main result.

2.2 Models and estimators

Let $(X_i, Y_i)_{i=\overline{1,n}}$ be an independent sequence identically distributed (i.i.d.) as (X, Y) which is a random pair valued in $\mathcal{E} \times \mathbb{R}$. Here $(\mathcal{E}, d)$ is a semi-metric space. $\mathcal{E}$ is not necessarily of a finite dimension, and we do not suppose the existence of a density for the functional random variable X.
Our goal, in this article, is to estimate the conditional hazard function defined by :

$$h^X(Y) = \frac{f^X(Y)}{1 - F^X(Y)}, \tag{2.1}$$

where :
$f^X(Y)$: is the conditional density function of Y given X.
$F^X(Y)$: is the conditional distribution function of Y given X.
For a fixed $x \in \mathcal{E}$, the k-NN kernel estimator of $h^x(Y = y)$ is given by :

$$\widehat{h}^x_{k-NN}(Y = y) = \widehat{h}^x(y) = \frac{\widehat{f}^x(y)}{1 - \widehat{F}^x(y)} \tag{2.2}$$

with :

$$\begin{aligned} F^x(y) &= \mathbb{P}[Y \leq y / X = x] \\ &= \mathbb{E}[\mathbb{1}_{]-\infty, y]} / X = x] \\ &= r(\mathbb{1}_{]-\infty, y]}). \end{aligned}$$

Where ; $r(.)$ is the regression function defined in Ferraty and Vieu ([13]). Therefore :

$$\widetilde{F}^x(y) = \widehat{r}(\mathbb{1}_{]-\infty, y]}) = \frac{\sum_{i=1}^{n} \mathbb{1}_{]-\infty, y]} K(H_n^{-1} d(x, X_i))}{\sum_{i=1}^{n} K(H_n^{-1} d(x, X_i))}.$$

Finally, by Roussas ([23]), Samanta ([24]) and Ferraty and Vieu ([13]), the estimator of the conditional distribution function is given by :

$$\widehat{F}^x(y) = \frac{\sum_{i=1}^{n} K(H_n^{-1} d(x, X_i)) R(g_n^{-1}(y - Y_i))}{\sum_{i=1}^{n} K(H_n^{-1} d(x, X_i))} ; \forall y \in \mathbb{R} \tag{2.3}$$

where K is an asymmetrical kernel, H_n is a positive random variable, defined as follows :

$$H_n(x) = \min\{h \in \mathbb{R}^+ / \sum_{i=1}^{n} 1\!\!1_{B(x,h)}(X_i) = k\} \tag{2.4}$$

with :

$$B(x,h) = \{x' \in \mathcal{E}; d(x,x') < h\}.$$

R is a distribution function and $(g_n)_{n\in\mathbb{N}}$ is a sequence of strictly positive real numbers (depending on n). Under a differentiability assumption of $\widehat{F}^x(y)$, we can obtain the conditional density function by differentiating the conditional distribution function, then we have :

$$\widehat{f}^x(y) = \frac{\partial}{\partial y}\widehat{F}^x(y)$$

then :

$$\widehat{f}^x(y) = \frac{\sum_{i=1}^{n} K(H_n^{-1}d(x,X_i))g_n^{-1}R'(g_n^{-1}(y-Y_i))}{\sum_{i=1}^{n} K(H_n^{-1}d(x,X_i))}. \tag{2.5}$$

In parallel, in order to emphasize differences between the k-NN method and the traditional kernel approach, we define the estimator of the conditional hazard function FERRATY et al. ([12]) by :

$$\widehat{h}^x_{kernel}(y) = \frac{\widehat{f}^x_{kernel}(y)}{1-\widehat{F}^x_{kernel}(y)} \tag{2.6}$$

with :

$$\widehat{f}^x_{kernel}(y) = \frac{\sum_{i=1}^{n} K(h_n^{-1}d(x,X_i))g_n^{-1}R'(g_n^{-1}(y-Y_i))}{\sum_{i=1}^{n} K(h_n^{-1}d(x,X_i))} \tag{2.7}$$

and :

$$\widehat{F}^x_{kernel}(y) = \frac{\sum_{i=1}^{n} K(h_n^{-1}d(x,X_i))R(g_n^{-1}(y-Y_i))}{\sum_{i=1}^{n} K(h_n^{-1}d(x,X_i))}. \tag{2.8}$$

Where K is a kernel, R is a distribution function and $(h_n)_{n\in\mathbb{N}}, (g_n)_{n\in\mathbb{N}}$ are sequences of strictly positive numbers.

2.3 Asymptotic properties of the k-NN method

2.3.1 The almost complete convergence (a.co.)

We focus in the pointwise the almost complete convergence [1] and rate of convergence [2] of the k-NN estimator of the conditional hazard function $\widehat{h}^x(y)$ defined on (2.2).
Before giving the main asymptotic result, we need some assumptions. The first one is about the concentration function $\varphi_x(h)$ and can be interpreted as a small ball probability of the functional variable x given by :

(H1)

$$\begin{aligned}\varphi_x(h) &= \mathbb{P}\left(X \in B(x,h)\right)\\ &= \mathbb{P}\left[X \in \{x' \in \mathcal{E}, d(x,x') < h\}\right]\end{aligned}$$

with $\varphi_x(h)$ continuous and strictly increasing in a neighborhood of a 0 and $\varphi_x(0) = 0$.

(H2) We also need a kernel K :
The kernel K is a function from $\mathbb{R}$ into $\mathbb{R}^+$, we say that K is a kernel of type I, so that : there exist two real constants $C_1, C_2; 0 < C_1 < C_2 < \infty$; such that ;

$$C_1 1\!\!1_{[0,1]} < K < C_2 1\!\!1_{[0,1]}.$$

K is a kernel of type II, so that : the support of K is $[0,1]$ and if

1. Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of real random variables. We say that $(X_n)_{n\in\mathbb{N}}$ converges almost completely (a.co.) to some r.r.v. X if and only if :

$$\forall\epsilon > 0, \sum_{n=1}^{\infty} \mathbb{P}[|X_n - X| > \epsilon] < \infty.$$

2. Let $(u_n)_{n\in\mathbb{N}}$ be a sequence of positive real number ; We say that $X_n = O_{a.co.}(u_n)$ if and only if : $\exists\epsilon > 0$,so that, $\sum_{n=1}^{\infty} \mathbb{P}[|X_n| > \epsilon u_n] < \infty$. This kind of convergence implies both almost sure convergence and convergence in probability.

its derivative K' exists on $[0,1]$ and satisfies; for two real constants $-\infty < C_1 < C_2 < 0$;

$$C_1 < K' < C_2.$$

In this case, we also suppose that : $\exists C_3 > 0, \exists \epsilon_0$:

$$\forall \epsilon < \epsilon_0, \int_0^{\epsilon} \varphi_x(u)du > C_3 \epsilon \varphi_x(\epsilon).$$

(H3) R is a differentiable function such that :

$$\exists C < \infty; \forall (x_1, x_2) \in \mathbb{R}^2; |R'(x_1) - R'(x_2)| \leq C|x_1 - x_2|.$$

$$R' \textit{ is of the support compact } [-1,1].$$

(H4) $\exists \zeta > 0$

$$\begin{cases} \forall (x_1, x_2) \in \mathbb{R}^2; \mid R(x_1) - R(x_2) \mid \leq C \mid x_1 - x_2 \mid \\ \int \mid t \mid^{\zeta} R'(t)dt < \infty. \end{cases}$$

(H5) $(g_n)_{n \in \mathbb{N}}$ is a strictly positive sequence such that :

$$\begin{cases} \lim_{n\to\infty} g_n = 0, \ \exists a > 0; \lim_{n\to\infty} n^a g_n = \infty. \\ \lim_{n\to\infty} \dfrac{\log n}{n g_n \varphi_x(h)} = 0 \end{cases}$$

The nonparametric model of the function h^x will be determined by regularity conditions of the conditional distribution of Y given X. These conditions are :

(H6) N_x will denote a fixed neighboorhood of x, S will be a fixed compact subset of $\mathbb{R}$:

We will consider two kinds of nonparametric models, The first one is called the "Lipschitz-type" model that is defined :

$$Lip_{\mathcal{E}\times\mathbb{R}} : \begin{cases} f : \mathcal{E} \times \mathbb{R} \to \mathbb{R}, \forall (x_1, x_2) \in N_x^2, \forall (y_1, y_2) \in S^2, \exists C < \infty; \exists \alpha, \beta > 0 \\ |f(x_1, y_1) - f(x_2, y_2)| \leq C \left(d(x_1, x_2)^{\alpha} + |y_1 - y_2|^{\beta} \right). \end{cases}$$

(H7) The second one called the "Continuity type" model that is defined as :

$$C^0_{\mathcal{E}\times\mathbb{R}} = \{ f : \mathcal{E} \times \mathbb{R} \to \mathbb{R}; \forall x' \in N_x, \lim_{d(x,x')\to 0} f(x', y) = f(x, y) \}.$$

(H8) Finally, we will consider the conditional moments of the response random variable Y :

$$\forall m \geq 2; \mathbb{E}[\mid Y \mid^m /X = x] = \sigma_m(x) < \infty$$

with $\sigma_m(.)$ continuous on x.

Before studying the k-NN estimator, we remind asymptotic properties of $\widehat{h}^x_{kernel}$ defined by equation (2.6). FERRATY et al.([12]), showed the almost complete convergence of this estimator.

Theorem 2.3.1 – *In the "continuity type" model and under the assumptions (H1), (H2), (H6) and (H8) we have :*

$$\widehat{h}^x_{kernel}(y) \longrightarrow h^x(y). \qquad a.co.$$

– *Under the "Lipschitz type" model and the hypotheses (H1), (H2), (H3), (H5),(H8), we have :*

$$\widehat{h}^x_{kernel}(y) - h^x(y) = O(h_n^\alpha) + O(g_n^\beta) + O\left(\sqrt{\frac{\log n}{n\varphi_x(h)}}\right).$$

Now we state the almost complete convergence for the nonparametric k-NN method estimate, defined in (2.2).

Theorem 2.3.2 *In the "continuity type" model and under the hypotheses(H1), (H2), (H4), (H5) and (H6), suppose that $k = k_n$ is a sequence of positive real numbers such that $\frac{k_n}{n} \to 0$ and $\frac{\log n}{k_n} \to 0$, then we have :*

$$\lim_{n\to\infty} \widehat{h}^x(y) = h^x(y). \qquad a.co.$$

Proof. We consider the following decomposition :

$$\widehat{h}^x(y) - h^x(y) = \frac{1}{1-\widehat{F}^x(y)}[\widehat{f}^x(y) - f^x(y)] + h^x(y)\frac{1}{1-\widehat{F}^x(y)}[\widehat{F}^x(y) - F^x(y)]. \tag{2.9}$$

Then the proof of Theorem 2.3.2 can be deduced from the following intermediates results. ■

Lemma 2.3.1 *Under the hypotheses of Theorem 2.3.2, we have :*

$$\lim_{n\to\infty} \widehat{f}^x(y) = f^x(y). \qquad a.co. \tag{2.10}$$

And

$$\lim_{n\to\infty} \widehat{F}^x(y) = F^x(y). \qquad a.co. \tag{2.11}$$

Lemma 2.3.2 *Under the hypotheses of Theorem 2.3.2, we have :*

$$\exists \delta > 0, \sum_{n\in\mathbb{N}} \mathbb{P}[1 - \widehat{F}^x(y)) < \delta] < \infty. \tag{2.12}$$

Theorem 2.3.3 *The hypotheses (H1)-(H8) imply*

$$\widehat{h}^x(y) - h^x(y) = O\left(\varphi_x^{-1}\left(\frac{k}{n}\right)^{\alpha}\right) + O\left(g_n^{\beta}\right) + O\left(\sqrt{\frac{\log n}{k_n g_n}}\right) \qquad a.co.$$

Proof. We consider the decomposition (2.9), and the proof of this Theorem is a consequence of these results. ∎

Lemma 2.3.3 *Under the hypotheses of Theorem (2.3.3), we have :*

$$\widehat{f}^x(y) - f^x(y) = O\left(\varphi_x^{-1}\left(\frac{k_n}{n}\right)^{\alpha}\right) + O(g_n^{\beta}) + O\left(\sqrt{\frac{\log n}{k_n g_n}}\right) \qquad a.co. \tag{2.13}$$

Lemma 2.3.4 *Under the hypotheses of Theorem (2.3.3), we have :*

$$\widehat{F}^x(y) - F^x(y) = O\left(\varphi_x^{-1}\left(\frac{k_n}{n}\right)^{\alpha}\right) + O(g_n^{\beta}) + O\left(\sqrt{\frac{\log n}{k_n}}\right) \qquad a.co. \tag{2.14}$$

2.3.2 Asymptotic normality

This section contains results on the asymptotic normality of $\widehat{h}^x(y)$. For this, we have to add the followings assumptions :

(H9) For each sequence $U_n \downarrow 0$ as $n \to \infty$ of positive real numbers, there exists a function : $\lambda(.)$ such that :
$\forall t \in [0,1], \lim_{U_n\to\infty} \frac{\varphi_x(tU_n)}{\varphi_x(U_n)} = \lambda(t)$.

(H10) $\lim_{n\to\infty}\left(g_n^2-\varphi_x^{-1}\left(\frac{k}{n}\right)\right)\sqrt{k_n}=0$ and $\frac{1}{k_n g_n}=o(g_n^{\beta})$.

Theorem 2.3.4 *Assume that (H1), (H9), (H10) hold, then for any $x\in\mathcal{A}$, we have :*

$$\left(\frac{k_n g_n}{\sigma_h^2(x,y)}\right)^{1/2}\left[\widehat{h}^x(y)-h^x(y)\right]\xrightarrow{\mathcal{D}}\mathcal{N}(0,1)\qquad \text{as } n\to\infty \tag{2.15}$$

where

$$\sigma_h^2(x,y)=\frac{\alpha_2 h^x(y)}{\alpha_1^2\,(1-F^x(y))}\quad (with : \alpha_j=K^j(1)-\int_0^1 (K^j)'(s)\lambda(s)\,ds\ \ for\ j=1,2)$$

$$\mathcal{A}=\{x\in\mathcal{E},\ f^x(y)[1-F^x(y)]\neq 0\} \tag{2.16}$$

$\xrightarrow{\mathcal{D}}$ *means the convergence in distribution.*

Proof. We consider the decomposition (2.9) and we show that the proof of Theorem (2.3.4) is a consequence of the following results. ■

Lemma 2.3.5 *Under the hypotheses of Theorem (2.3.4), we have :*

$$\left(\frac{k_n g_n}{\sigma_f^2(x,y)}\right)^{1/2}\left[\widehat{f}^x(y)-f^x(y)\right]\xrightarrow{\mathcal{D}}\mathcal{N}(0,1)\quad \text{as } n\to\infty$$

where

$$\sigma_f^2(x,y)=f^x(y)\int R'^2(t)dt. \tag{2.17}$$

Lemma 2.3.6 *Under the hypotheses of Theorem 2.3.4, we have :*

$$\left(\frac{k_n g_n}{\sigma_F^2(x,y)}\right)^{1/2}\left[\widehat{F}^x(y)-F^x(y)\right]\xrightarrow{\mathcal{D}}\mathcal{N}(0,1)\quad \text{as } n\to\infty$$

where

$$\sigma_F^2(x,y)=F^x(y)\left[1-F^x(y)\right]. \tag{2.18}$$

Lemma 2.3.7 *Under the hypotheses of Theorem 2.3.4, we have :*

$$(1-\widehat{F}^x(y))\to(1-F^x(y))\quad \text{in probability.}$$

2.4 Applications

2.4.1 Conditional Confidence Interval

The main application of the Theorem (2.3.4) is the to build confidence interval for the true value of $h^x(y)$ given curve $X = x$. A plug-in estimate for the asymptotic standard deviation $\sigma(x, \theta_x)$ can be obtained using the estimators $\widehat{h}^x(y)$ and $\widehat{F}^x(y)$ of $h^x(y)$, $F^x(y)$ respectively. We get $\widehat{\sigma}(x, y) := \left(\frac{\widehat{\alpha_2}\widehat{h}^x(y)}{(\widehat{\alpha_1})^2(1-\widehat{F}^x(y))}\right)^{1/2}$. Then $\widehat{h}^x(y)$ can be used to get the following approximate $(1-\zeta)$ confidence interval for $h^x(y)$

$$\widehat{h}^x(y) \pm t_{1-\zeta/2} \times \left(\frac{\widehat{\sigma}_n^2(x, y)}{g_n k}\right)^{1/2}$$

where $t_{1-\zeta/2}$ denotes the $1-\zeta/2$ quantile of the standard normal distribution. We estimate empirically α_1 and α_2 by

$$\widehat{\alpha_1} = \frac{1}{kg(x)} \sum_{i=1}^{n} K_i \quad \text{and} \quad \widehat{\alpha_2} = \frac{1}{kg(x)} \sum_{i=1}^{n} K_i^2,$$

where $K_i = K\left(\frac{d(x, X_i)}{\phi^{-1}(k/n)}\right)$.

This last estimation is justified by the fact that, under (H1), (H5) and (H6), we have, (see Ferraty and Vieu ([13]) p. 44)

$$\frac{1}{kg(x)} \mathbb{E}[K_1^j] \to \alpha_j, \quad j = 1, 2.$$

2.4.2 A Simulation study

In this section we will show the effectiveness of k-NN method compared to the kernel estimation using simulated data. For this we considered a sample of a diffusion process on interval $[0, 1]$, $Z_1(t) = 2 - \cos(\pi t W)$ and $Z_2(t) = \cos(\pi t W)$, where W is the standard normal distribution and take $X(t) = AZ_1(t) + (1-A)Z_2(t)$, where A is random variable Bernoulli distributed. We carried out the simulation with 200-sample of the curve X which is represented by the following graph :

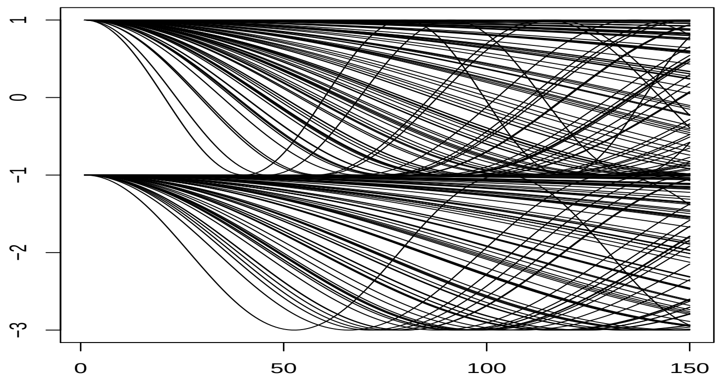


FIGURE 2.1 – The 200 curves X.

For the scalar response variable, we took $Y = Ar_1(X)+(1-A)r_2(X)$ where r_1 (resp. r_2) is the nonlinear regression model $r_1(X) = 0.25\times\left(\int_0^1 X'(t)dt\right)^2 +\epsilon$, with ϵ is $U([0, 0.5])$ (resp. $r_2(X)$ is the null function). We choose a quadratic kernel K defined by :

$$K(x) = \frac{3}{2}(1 - x^2)\mathbb{1}_{[0,1]}.$$

In practice, the semi-metric choice is based on the regularity of the curves X. For this we use the semi-metric defined by L_2-distance between the q^{th} derivatives of the curves. In order to evaluate the MSE (Mean Square Error) we proceed by the following algorithm :

- *Step 1.* We split our data into two subsets ; the first sample, of size n=120 corresponds to the learning sample which will be used, as a sample, to compute our conditional hazard function estimators for the 80 remaining curves (considered as the test sample).
  - $(X_j, Y_j)_{j\in J}$ learning sample,
  - $(X_i, Y_i)_{i\in I}$ test sample.
- *Step 2.*
  - We use the learning sample for computing the hazard function estimator $\widehat{h}_j$, for all $j \in J$.
  - We set : $i^* = \arg\min_{j\in J} d(X_i, X_j)$.

– We put : $\forall i \in I$

$$\widehat{T}_i = \widehat{h}^{X_{i^*}}(Y_i) \text{ for kernel method.}$$
$$\widehat{T}_i = \widehat{h}^{X_{k_{opt}}}(Y_i) \text{ for } k\text{-NN method,}$$

where :

$$X_{i^*} : \text{is the nearest curve to } X_j$$
$$k_{opt} : \arg\min_a(CV(a))$$

with :

$$CV(a) = \frac{1}{n}\left[\sum_{i\in J}\int (\widehat{f}^{-i}_{(a,b)}(X_i, y))^2 dy - 2\sum_{i\in J}\widehat{f}^{-i}_{(a,b)}(X_i, Y_i)\right]$$

and :

$$\widehat{f}^{-k}_{(a,b)}(x, y) = \frac{b^{-1}\sum_{i\in J, i\neq k} K\left(\frac{d(x, X_i)}{a}\right) R\left(\frac{y - Y_i}{b}\right)}{\sum_{i\in J} K\left(\frac{d(x, X_i)}{a}\right)}.$$

– *Step 3.* The error used to evaluate this comparison is the mean of square error (MSE) expressed by

$$\frac{1}{card(I)}\sum_{i\in I}\left|h(Y_i) - \widehat{T}(X_i, Y_i)\right|^2,$$

where $\widehat{T}$ designate the estimator used : kernel or k-NN method estimation and h is the true hazard function.

Consequently, the k-NN method gives slightly better results than the kernel method. This is confirmed by the MSE-k-NN$= 0.8227394$ and MSE-Kernel $= 1.347982$.

2.4.3 Real data application

To highlight the efficiency and robustness of the method of k nearest neighbors with respect to the kernel method in estimating the conditional hazard function, we will test these two methods in the presence or not of heterogeneous data.

To do this, based on the study of Burba et al. (2009) which emphasizes the effect of the nature of the data (homogeneous or heterogeneous) on the

quality of the estimate, especially the superiority of the k-nearest neighbors in the presence of very heterogeneous data.
For this purpose, we apply the described algorithm used in the simulation study to some chemiometrical real data available on the site[3], the original of these data (215 selected pieces of meat) comes from a quality control problem in the food industry that controls grease on a sample of finely chopped meat by chemical processes.
The sample of size 215 was splitted into learning sample of size 205 (with all data), 178 (without the heterogeneous data, 27 values) and testing sample of size 10. Figue 2.2 displays the curves of learning sample for all data and the curves of learning sample without the heterogeneous one.

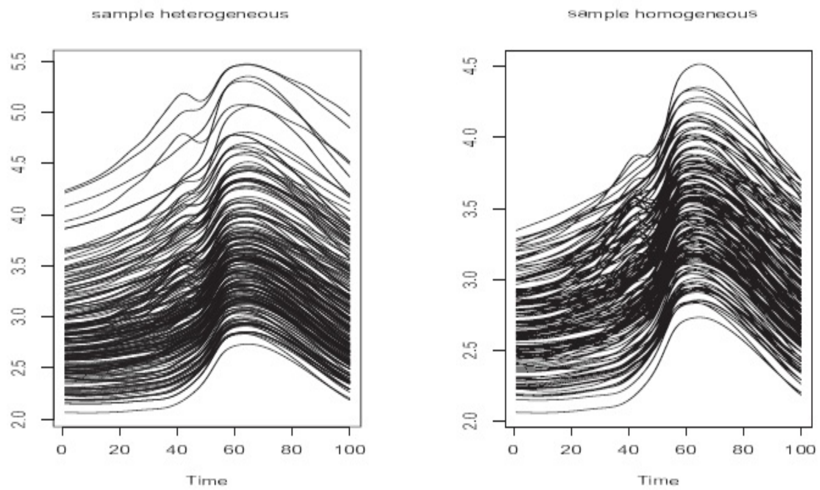


FIGURE 2.2 – The learning curves.

For our study, we use the standard L^2 semi-metric and a quadratic kernel function K.
We plot the conditional hazard function estimated for the first 3 values of testing sample, Figure 2.3 depicts that the k-NN method in presence of heterogeneous data give better estimation of the conditional hazard function prediction (regular function) than the kernel method estimation (non-regular function) and when the data are homogeneous the two method give the same result that can be easily seen in Figure 2.4.

3. http ://lib.stat.cmu.edu/datasets/tecator.

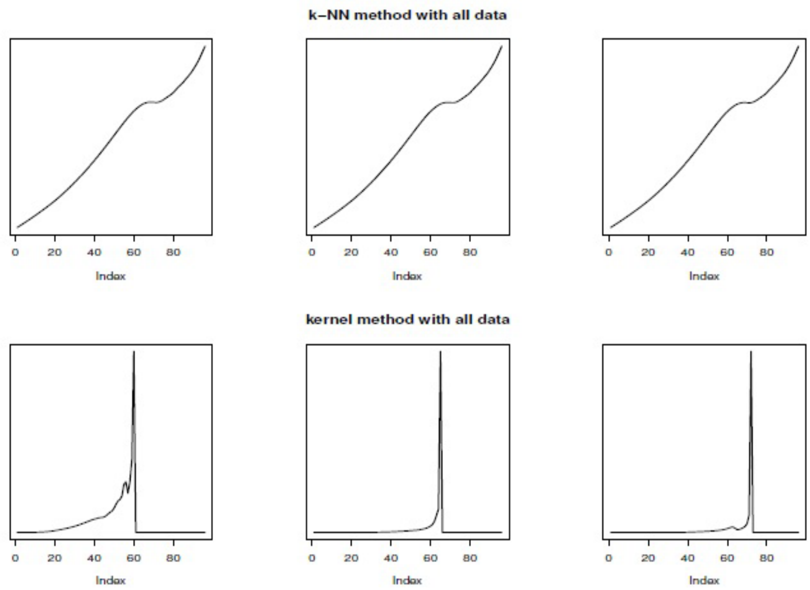


FIGURE 2.3 – k-NN method (upper panels) *vs* kernel method (lower panels) of conditional hazard function far all data.

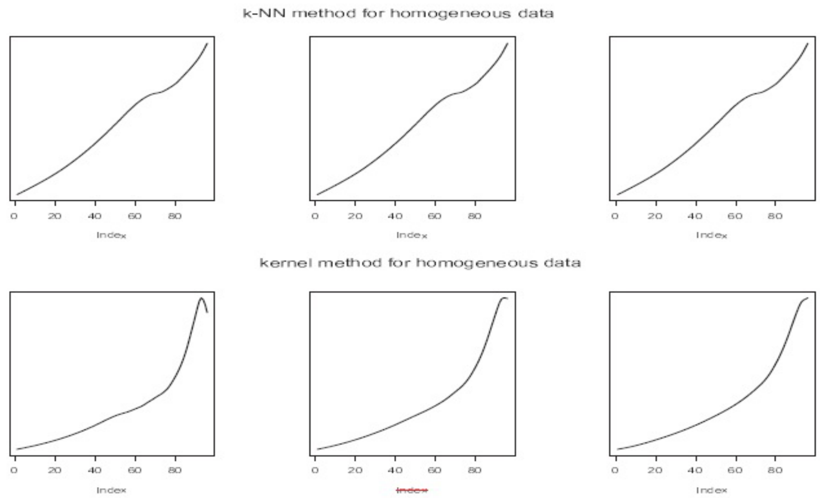


FIGURE 2.4 – k-NN method (upper panels) *vs* kernel method (lower panels) of conditional hazard function without heterogeneous data.

2.5 General technical tools

Let $(A_i, B_i)_{i\in\mathbb{N}}$ be a sequence of random variables with values in $(\Omega \times \mathbb{R}, \mathcal{A} \otimes \mathcal{B})$, independent but not necessarily identically distributed,where $(\Omega, \mathcal{A})$ is a general measurable space, let $G : \mathbb{R} \times \Omega \to \mathbb{R}^+$ a measurable function such that : $\forall w, w' \in \mathbb{R}$:

$$w \leq w' \Longrightarrow G(w,z) \leq G(w',z); \qquad \forall z \in \Omega.$$

Let c be a not random positive real number and T a real random variable, we define, $\forall n \in \mathbb{N}^*$

$$C_n(T) = \frac{\sum_{i=1}^{n} B_i G(T, A_i)}{\sum_{i=1}^{n} G(T, A_i)}.$$

Lemma 2.5.1 Burba *et al.* ([4])
Let $(D_n)_{n\in\mathbb{N}}$ be a sequence of real random variables and $(u_n)_{n\in\mathbb{N}}$ be a decreasing sequence of positive numbers.
If $l = \lim u_n \neq 0$, and if, for all increasing sequence $\beta_n \in]0,1[$, there exist two sequences of real random variables $(D_n^+(\beta_n))_{n\in\mathbb{N}}$ and $(D_n^-(\beta_n))_{n\in\mathbb{N}}$:

(L1) $\forall n \in \mathbb{N}; D_n^- \leq D_n^+$, *and* $\mathbb{I}_{D_n^- \leq D_n \leq D_n^+} \to 1 \qquad a.co.$

(L2) $$\frac{\sum_{i=1}^{n} G(D_n^-, A_i)}{\sum_{i=1}^{n} G(D_n^+, A_i)} - \beta_n = O(u_n) \qquad a.co.$$

(L3)
$$C_n(D_n^-) - c = O(u_n) \qquad a.co.$$
$$C_n(D_n^+) - c = O(u_n). \qquad a.co.$$

Then :

$$C_n(D_n) - c = O(u_n). \qquad a.co.$$

If l=0 and if (L1), (L2), (L3) hold for any increasing sequence $\beta_n \in]0,1[$ with limit 1, the same result holds.

Lemma 2.5.2 Burba *et al.* ([4])
Let $(D_n)_{n\in\mathbf{N}}$ be a sequence of real random variables and $(v_n)_{n\in\mathbb{N}}$ be a decreasing positive sequence.
If $l' = \lim v_n \neq 0$, and if, for all increasing sequence $\beta_n \in]0,1[$, there exist two sequences of real random variables $(D_n^+(\beta_n))_{n\in\mathbb{N}}$ and $(D_n^-(\beta_n))_{n\in\mathbb{N}}$:

(L'1) $\forall n \in \mathbb{N}; D_n^- \leq D_n^+$, *and* $\mathbb{1}_{D_n^- \leq D_n \leq D_n^+} \to 1$ *a.co.*

(L'2) $\dfrac{\sum_{i=1}^{n} G(D_n^-, A_i)}{\sum_{i=1}^{n} G(D_n^+, A_i)} - \beta_n = o(v_n)$ *a.co.*

(L'3)

$$C_n(D_n^-) - c = o(v_n) \qquad a.co.$$
$$C_n(D_n^+) - c = o(v_n). \qquad a.co.$$

Then :

$$C_n(D_n) - c = o(v_n). \qquad a.co.$$

If l'=0 and if (L'1), (L'2), (L'3) hold for any increasing sequence $(\beta_n) \in]0,1[$ *with limit 1, the same result holds.*

Burba et al. ([4]) use in their consistency proof of the k-NN kernel estimate for independent data a Chernoff-type exponential inequality to check conditions $(L1)$ or $(L'1)$.

Lemma 2.5.3 Burba et al. ([4])
Let $(X_1, X_2, ..., X_n)$ *be independent random variable in* $\{0,1\}$. *We note* $X = \sum_{i=1}^{n} X_i$ *and* $\mu = \mathbb{E}(X)$, *then,* $\forall \delta > 0$ *;*

$$\mathbb{P}[X > (1+\delta)\mu] < [e^{\delta}/(1+\delta)^{1+\delta}]^{\mu}$$
$$\mathbb{P}[X < (1-\delta)\mu] < [e^{-\delta^2/2\mu}].$$

2.6 Appendix :

2.6.1 Proof of section 3.3.1

Proof of Lemma 2.3.1 :
On the one hand, to prove the first result, we apply Lemma 2.5.2 with :

$$\begin{cases} v_n = 1 \\ H_n = D_n \\ \widehat{f^x}(y) = C_n(D_n) \\ f^x(y) = c. \end{cases} \tag{2.19}$$

Choose $\beta_n \in]0,1[$, (D_n^-) and (D_n^+) such that :

$$\begin{cases} \varphi_x(D_n^-) = \sqrt{\beta_n}\varphi_x(h) = \sqrt{\beta_n}\dfrac{k_n}{n} \\ \varphi_x(D_n^+) = \dfrac{1}{\sqrt{\beta_n}}\varphi_x(h) = \dfrac{1}{\sqrt{\beta_n}}\dfrac{k_n}{n}. \end{cases} \tag{2.20}$$

Define

$$\begin{cases} h^- = D_n^- = \varphi_x^{-1}(\sqrt{\beta_n}\dfrac{k_n}{n}) \\ h^+ = D_n^+ = \varphi_x^{-1}(\dfrac{1}{\sqrt{\beta_n}}\dfrac{k_n}{n}). \end{cases} \tag{2.21}$$

Ferraty and Vieu ([13]) proved under the conditions of Theorem 2.3.1 that :

$$\frac{1}{n\varphi_x(h)}\sum_{i=1}^{n} K(h^{-1}d(x,X_i)) \to 1. \qquad a.co.$$

Under the conditions (2.20) and (2.21), we have :

$$\begin{cases} \dfrac{1}{n\varphi_x(D_n^-)}\displaystyle\sum_{i=1}^{n} K\left((D_n^-)^{-1}d(x,X_i)\right) \to 1 & a.co. \\ \dfrac{1}{n\varphi_x(D_n^+)}\displaystyle\sum_{i=1}^{n} K\left((D_n^+)^{-1}d(x,X_i)\right) \to 1. & a.co. \end{cases}$$

Then :

$$\frac{\displaystyle\sum_{i=1}^{n} K\left((D_n^-)^{-1}d(x,X_i)\right)}{\displaystyle\sum_{i=1}^{n} K\left((D_n^+)^{-1}d(x,X_i)\right)} \to \beta_n \qquad a.co.$$

so that $(L'2)$ is checked. Now by using Lemma (6.15) in Ferraty and Vieu ([13]) under the conditions of Theorem 2.3.1 and :

$$D_n \longrightarrow 0\ ,\ \frac{\log n}{n\varphi_x(D_n)} \longrightarrow 0\ (n \to \infty) \tag{2.22}$$

we have :

$$C_n(D_n^-) \to c \qquad a.co.$$
$$C_n(D_n^+) \to c \qquad a.co.$$

so $(L'3)$ is verified. Finally, we check $(L'1)$. The first part is obvious, and the second one that, $\forall \epsilon > 0$:

$$\sum_{n\geq 0} \mathbb{P}\left[|\mathbb{1}_{D_n^- < H_n < D_n^+} - 1| > \epsilon\right] < \infty.$$

We know that ;

$$\mathbb{P}[|\mathbb{1}_{D_n^- < H_n < D_n^+} - 1| > \epsilon] \leq \underbrace{\mathbb{P}[H_n < D_n^-]}_{A_1} + \underbrace{\mathbb{P}[H_n > D_n^+]}_{A_2}.$$

$$A_1 \leq \mathbb{P}\left[\sum_{i=1}^{n} \mathbb{1}_{B(x,D_n^-)} > k_n\right].$$

And by using Lemma 2.5.3 with :

$$\begin{cases} X_i = \mathbb{1}_{B(x,D_n^-)} \\ X = \sum_{i=1}^{n} \mathbb{1}_{B(x,D_n^-)} \\ \mathbb{P}(X_i = 1) = \varphi_x(D_n^-) \\ \mu = \mathbb{E}(X) = \sum_{i=1}^{n} \mathbb{E}\left[\mathbb{1}_{B(x,D_n^-)}\right] = n\varphi_x(D_n^-) \end{cases} \tag{2.23}$$

we get :

$$\begin{aligned} \mathbb{P}[H_n < D_n^-] &< \left[e^{(\frac{1}{\sqrt{\beta}}-1)}/(\frac{1}{\sqrt{\beta}})^{-\frac{1}{\sqrt{\beta}}}\right]^{n\varphi_x(D_n^-)} \\ &< n^{(-\log\sqrt{\beta_n}e^{(1-\sqrt{\beta_n})})\frac{-k_n}{\log n}}. \end{aligned}$$

Under the hypotheses (2.22) and as $\sqrt{\beta}(e^{(1-\sqrt{\beta})}) < 1$ then :

$$A_1 = \mathbb{P}[H_n < D_n^-] < \infty.$$

Turning now to the study of A_2, we obtain :

$$\mathbb{P}[H_n > D_n^+] = \mathbb{P}\left[\sum_{i=1}^{n} \mathbb{1}_{B(x,D_n^+)} < n\sqrt{\beta}\varphi_x(D_n^+)\right]$$

under the modification (2.23) and by applying the Lemma 2.5.3, we obtain :

$$\begin{aligned} \mathbb{P}[H_n > D_n^+] &< e^{\frac{-k_n(1-\sqrt{\beta})^2}{2\sqrt{\beta}}} \\ &< \left(n^{\frac{(1-\sqrt{\beta})^2}{2\sqrt{\beta}}}\right)^{\frac{-k_n}{\log n}}. \end{aligned}$$

Since $\dfrac{(1-\sqrt{\beta})^2}{2\sqrt{\beta}} > 0$ and $\dfrac{n\varphi_x(h)}{\log n} \to \infty$ then :

$$\mathbb{P}[H_n > D_n^+] < \infty.$$

Finally :

$$\mathbb{P}\left[|1\!\!1_{D_n^- < H_n < D_n^+} - 1| > \epsilon\right] < \infty.$$

On the other hand, we prove the second result. For this, we use the precedents steps with :

$$\begin{cases} v_n = 1 \\ H_n = D_n \\ \widehat{F}^x(y) = C_n(D_n) \\ F^x(y) = c. \end{cases} \tag{2.24}$$

Proof of Lemma 2.3.2 :
It is clear that :

$$|1 - \widehat{F}^x(y)| < \frac{1 - F^x(y)}{2} \Longrightarrow |\widehat{F}^x(y) - F^x(y)| > \frac{1 - F^x(y)}{2}$$

turning now, to the term of probability, we obtain :

$$\mathbb{P}\left[|1 - \widehat{F}^x(y)| < \frac{1 - F^x(y)}{2}\right] \leq \mathbb{P}\left[|\widehat{F}^x(y) - F^x(y)| > \frac{1 - F^x(y)}{2}\right]$$

$$\sum_{n\in\mathbb{N}} \mathbb{P}\left[|1 - \widehat{F}^x(y)| < \frac{1 - F^x(y)}{2}\right] \leq \sum_{n\in\mathbb{N}} \mathbb{P}\left[|\widehat{F}^x(y) - F^x(y)| > \frac{1 - F^x(y)}{2}\right].$$

For the second term, by result 2.11, we have :

$$\sum_{n\in\mathbb{N}} \mathbb{P}\left[|\widehat{F}^x(y) - F^x(y)| > \frac{1 - F^x(y)}{2}\right] < \infty$$

then ; for $\delta = \dfrac{1-F^x(y)}{2}$, we obtain :

$$\sum_{n\in\mathbb{N}} \mathbb{P}\left[|\widehat{F}^x(y) - F^x(y)| > \frac{1-F^x(y)}{2}\right] < \infty.$$

Proof of Lemma 2.3.3 :
To prove this lemma, we use Lemma 2.5.1. Choose β_n as an increasing sequence in $]0,1[$ with limit 1.
Furthermore, we choose D_n^- and D_n^+ under (2.20), FERRATY and VIEU ([13]) proved under the conditions of Theorem 2.3.1 that :

$$\widehat{r}_3(x) - \mathbb{E}\left[\widehat{r}_3(x)\right] = O\left(\sqrt{\frac{\log n}{nh_n\varphi_x(h)}}\right)$$

with

$$\widehat{r}_3(x) = \frac{1}{n}\sum_{i=1}^{n} \frac{K\left(h_n^{-1}d(x,X_i)\right) R\left(g_n^{-1}(y-Y_i)\right)}{\mathbb{E}\left[K\left(h_n^{-1}d(x,X_1)\right)\right]}.$$

$$\widehat{r}_3(x) = \frac{1}{n}\sum_{i=1}^{n} \frac{K\left(h_n^{-1}d(x,X_i)\right) \Gamma_i(y)}{\mathbb{E}\left[K\left(h_n^{-1}d(x,X_1)\right)\right]}.$$

$$\Gamma_i(y) = R\left(g_n^{-1}(y-Y_i)\right).$$

Then

$$\widehat{r}_3(x) - \mathbb{E}\left[\widehat{r}_3(x)\right] = \frac{1}{n}\sum_{i=1}^{n} \frac{K\left(h_n^{-1}d(x,X_i)\right)}{\mathbb{E}K\left(h_n^{-1}d(x,X_1)\right)}\Gamma_i(y) - \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left[\frac{K\left(h_n^{-1}d(x,X_i)\right)}{\mathbb{E}K\left(h_n^{-1}d(x,X_1)\right)}\Gamma_i(y)\right]$$

$$= \frac{1}{n\mathbb{E}K\left(h_n^{-1}d(x,X_1)\right)}\sum_{i=1}^{n} K\left(h_n^{-1}d(x,X_i)\right)\Gamma_i(y)$$

$$-\frac{1}{\mathbb{E}K\left(h_n^{-1}d(x,X_1)\right)}\mathbb{E}\left[K\left(h_n^{-1}d(x,X_1)\right)\mathbb{E}(\Gamma_1(y)/X_1)\right]$$

$$= \frac{1}{n\mathbb{E}K\left(h_n^{-1}d(x,X_1)\right)}\sum_{i=1}^{n} K\left(h_n^{-1}d(x,X_i)\right)\Gamma_i(y) - \mathbb{E}(\Gamma_1(y)/X_1).$$

Using the fact that $\mathbb{E}\left[K\left(h_n^{-1}d(x,X_i)\right)\right] = O(\varphi_x(h)$ (see Ferraty and Vieu ([13]) and under the notations (2.20) and (2.21), we have :

$$\begin{cases} \dfrac{1}{n\varphi_x(D_n^-)}\displaystyle\sum_{i=1}^{n} K\left(\frac{d(x,X_i)}{D_n^-}\right)\Gamma_i(y) = \mathbb{E}(\Gamma_1(y)/X_1) + O\left(\sqrt{\frac{\log n}{g_n k_n}}\right) \\ \dfrac{1}{n\varphi_x(D_n^+)}\displaystyle\sum_{i=1}^{n} K\left(\frac{d(x,X_i)}{D_n^+}\right)\Gamma_i(y) = \mathbb{E}(\Gamma_1(y)/X_1) + O\left(\sqrt{\frac{\log n}{g_n k_n}}\right) \end{cases}$$

by this, we obtain :

$$\frac{\displaystyle\sum_{i=1}^{n} K\left(\frac{d(x,X_i)}{D_n^-}\right)}{\displaystyle\sum_{i=1}^{n} K\left(\frac{d(x,X_i)}{D_n^+}\right)} - \beta_n = O\left(\sqrt{\frac{\log n}{g_n k_n}}\right) \ a.co.$$

that $(L2)$ is verified. Now, we apply Lemma (6.15) for Ferraty and Vieu ([13]) under (2.20) and (2.19), we get :

$$C_n(D_n^-) - c = O\left(\varphi_x^{-1}\left(\frac{k}{n}\right)^{\alpha}\right) + O(g_n^{\beta}) + O\left(\sqrt{\frac{\log n}{g_n k_n}}\right) \ a.co.$$

$$C_n(D_n^+) - c = O\left(\varphi_x^{-1}\left(\frac{k}{n}\right)^{\alpha}\right) + O(g_n^{\beta}) + O\left(\sqrt{\frac{\log n}{g_n k_n}}\right) \ a.co.$$

that verifies condition $(L3)$.
Proof of Lemma 2.3.4 :
To verify this Lemma, we pass by the same steps as before, such that; Ferraty and Vieu ([13]) showed that :

$$\widehat{r}_1(x) - 1 = O\left(\sqrt{\frac{\log n}{n\varphi_x(h)}}\right)$$

with

$$\widehat{r}_1(x) = \frac{1}{n}\sum_{i=1}^{n}\frac{K\left(h_n^{-1}d(x,X_i)\right)}{\mathbb{E}K\left(h_n^{-1}d(x,X_1)\right)}.$$

Then

$$\frac{1}{n}\sum_{i=1}^{n} K\left(h_n^{-1} d(x,X_i)\right) - \varphi_x(h) = O\left(\sqrt{\frac{\log n}{n\varphi_x(h)}}\right)$$

under the same choice of $h^- = D_n^-$ and $h^+ = D_n^+$ as above, we have :

$$\begin{cases} \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} K\left(\frac{d(x,X_i)}{D_n^-}\right) = \sqrt{\beta_n}\frac{k_n}{n} + O\left(\sqrt{\frac{\log n}{k_n}}\right) \\ \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} K\left(\frac{d(x,X_i)}{D_n^+}\right) = \frac{1}{\sqrt{\beta_n}}\frac{k_n}{n} + O\left(\sqrt{\frac{\log n}{k_n}}\right) \end{cases}$$

we get :

$$\frac{\displaystyle\sum_{i=1}^{n} K\left(\frac{d(x,X_i)}{D_n^-}\right)}{\displaystyle\sum_{i=1}^{n} K\left(\frac{d(x,X_i)}{D_n^+}\right)} - \beta_n = O\left(\sqrt{\frac{\log n}{k_n}}\right)$$

so that, $(L2)$ is checked. Now wear able to apply Lemma (6.14) in Ferraty and Vieu ([13]) under (3.33), we obtain :

$$C_n(D_n^-) - c = O\left(\varphi_x^{-1}\left(\frac{k_n}{n}\right)^{\alpha}\right) + O(g_n^{\beta}) + O\left(\sqrt{\frac{logn}{k_n}}\right)$$

$$C_n(D_n^+) - c = O\left(\varphi_x^{-1}\left(\frac{k_n}{n}\right)^{\alpha}\right) + O(g_n^{\beta}) + O\left(\sqrt{\frac{logn}{k_n}}\right)$$

and $(L3)$ is verified.

2.6.2 Proof of section 3.3.2

Proof of Lemma 2.3.5 :
We denote :

$$\begin{cases} C_n(H_n) = \widehat{f}^x(y) \\ c = f^x(y). \end{cases} \tag{2.25}$$

Under (2.20) and (2.21), we have :

$$\left(\frac{k_n g_n}{\sigma_f^2(x,y)}\right)^{1/2}[C_n(H_n) - c] = \left(\frac{k_n g_n}{\sigma_f^2(x,y)}\right)^{1/2}\left[C_n(D_n^+) - c\right] + \left(\frac{k_n g_n}{\sigma_f^2(x,y)}\right)^{1/2}\left[C_n(H_n) - C_n(D_n^+)\right]. \tag{2.26}$$

Then, to establish the asymptotic normality of the conditional density function, we need to show the asymptotic normality of the first term in equation (2.26) and the second term converges a.co. to 0.
For this, we remind that, under the same assumptions as Lemma 2.3.5, Quintela-del-Río ([21]) in Theorem 5 proved that

$$\left(\frac{k_n g_n}{\sigma_f^2(x,y)}\right)^{1/2}\left[C_n(D_n^+)-c\right]\xrightarrow{\mathcal{D}}\mathcal{N}(0,1)\quad \text{as } n\to\infty.$$

One the other hand, by hypothesis (H2) and the fact that $\mathbb{1}_{\{D_n^-\leq H_n\leq D_n^+\}}\to 1$ where $\dfrac{k_n}{n}\to 0$ (see Burba et al. ([4])), we have :

$$C_n(D_n^+)\leq C_n(H_n)\leq C_n(D_n^-).$$

Using the fact that :

$$\begin{aligned}\mid C_n(H_n)-C_n(D_n^+)\mid &\leq\mid C_n(D_n^-)-C_n(D_n^+)\mid\\ \leq\mid C_n(D_n^-)-\mathbb{E}\left[C_n(D_n^-)\right]\mid &+\mid C_n(D_n^+)-\mathbb{E}[C_n(D_n^+)]\mid\\ &+\mid \mathbb{E}[C_n(D_n^-)]-\mathbb{E}[C_n(D_n^+)]\mid .\end{aligned}\tag{2.27}$$

For the first term, we can write :

$$\begin{aligned}\mid C_n(D_n^-)-\mathbb{E}\left[C_n(D_n^-)\right]\mid &\leq\mid C_n(D_n^-)-c\mid\\ &+\mid \mathbb{E}\left[C_n(D_n^-)\right]-c\mid\end{aligned}$$

by Lemma (2.3.3), we have :

$$\mid C_n(D_n^-)-c\mid= O\left(\varphi_x^{-1}\left(\frac{k_n}{n}\right)^{\alpha}\right)+O(g_n^{\beta})+O\left(\sqrt{\frac{\log n}{k_n g_n}}\right)$$

and Quintela-del-Río ([21]) proved that :

$$\mid \mathbb{E}\left[C_n(D_n^-)\right]-c\mid= o(g_n^{\beta})+O(\frac{1}{k_n}).\tag{2.28}$$

Finally, under hypothesis (H10), we obtain the almost complete convergence of the first term of (2.27). And to establish the almost complete convergence

of the second term we apply the same steps as precedent.
Finally for the third term, we have :

$$\mid \mathbb{E}[C_n(D_n^-)] - \mathbb{E}[C_n(D_n^+)] \mid \leq \mid \mathbb{E}[C_n(D_n^-)] - c \mid + \mid \mathbb{E}[C_n(D_n^+)] - c \mid$$

the almost complete convergence to 0 of these two terms is verifies in (2.28).
Proof of Lemma 2.3.6 :
To prove this Lemma, we apply the same steps as precedent with :

$$\begin{cases} C_n(H_n) = \widehat{F}^x(y) \\ c = F^x(y). \end{cases} \tag{2.29}$$

Proof of Lemma 2.3.7 :
It is clear that, the result (2.11) of Lemma (2.3.1) permits to conclude that :

$$\widehat{F}^x(y) \to F^x(y) \ \ in \ \ probability.$$

Acknowledgements.
The authors are grateful to a referee for helpful comments and suggestions that improved the paper.

Bibliographie

[1] Attouch, M. ; Benchikh, T. (2012). Asymptotic distribution of robust k-nearest neighbour estimator for functional nonparametric models. *Mathematic Vesnic*, **64**, No. 4, pp. 275-285.

[2] Burba, F. ; Ferraty, F. and Vieu, P.(2009). k-Nearest Neighbour method in functional nonparametric regression. *Journal of Nonparametric Statistics* **21**, 4 , 453–469.

[3] Burba, F. ; Ferraty, F. and Vieu, P. (2008). Convergence de l'estimateur à noyau des k plus proches voisins en régression fonctionnelle non-paramétrique. *Preprint submitted to Elsevier Science*.

[4] Collomb, G. (1980). Estiamation de la regression par la methode des k plus proches avec noyau : quelques propriétés de convergence ponctuelle. *Statistique Non Parametrique Asymptotique*, J-P. Raoult, Ed. vol.821 of *Lecture Notes in Mathematics*. Springer Berlin/Heidelberg, pp. 159–175.

[5] Collomb, G. (1997). Quelques propriétés de la methode du noyau pour l'estimation nonparamétrique de la régression en un point fixe, *C. R. Acad. Sci. Paris*, Série 1 285, 289–293.

[6] Devroye, L.P. *The uniform convergence of nearest neighbour regression function estimators and their application in optimization*, IEEE Trans. Inform. Theory, **24**, 142-151, 1978.

[7] Demongeot, J. ; Laksaci, A. ; Madani, F. and Rachdi, M. (2010). Local linear estimation of the conditional density for functional data. C. R., *Math., Acad. Sci. Paris*, 348, Issues 15-16, 931–934.

[8] Demongeot, J. ; Laksaci, A. ; Madani, F. and Rachdi, M. (2011). *A fast functional locally modeled conditional density and mode for functional time-series*. springer.

[9] DEMONGEOT, J. ; LAKSACI, A. ; MADANI, F. and RACHDI, M. Functional data : local linear estimation of the density and its application. *Statistics*, DOI : 10.1080/02331888.2011.568117, In press.

[10] FERRATY, F. ; LAKSACI A. ; TADJ, A. and VIEU, P.(2011). Kernel regression with functional response. *Electronic Journal of Statistics.* **5**, 159-171.

[11] FERRATY, F. ; LAKSACI A. and VIEU, P.(2006). Estimating Some Characteristics Of The Conditional Distribution in Nonparametric Functional Models. *Statistical Inference for Stochastic Processes.* **9**, 47–76.

[12] FERRATY, F. ; RABHI A. and VIEU, P. (2008). Estimation non parametrique de la fonction de hasard avec variable explicative fonctionnelle. *Roumaine Math*, **53**, 1–18.

[13] FERRATY, F. and VIEU, P. (2006). *Nonparametric Functional Data Analysis.* Springer, New York.

[14] LAKSACI, A. ; MADANI, F. and RACHDI, M. (2012). Kernel conditional density estimation when the regressor is valued in a semi-metric space. *Communications Statistics-theory and methods* Accepted paper.

[15] LAKSACI, A. and MECHAB, B. (2012). Conditional hazard estimate for functional random fiels. *Communications Statistics-theory and methods* Accepted paper.

[16] LAKSACI, A. and MECHAB, B.(2010). Estimation nonparamétrique de la fonction de hasard avec variable explicative fonctionnelle : cas des données spatiales. *Rev. Roumaine Math. Pures Appl.* **55**, 35-51 .

[17] LIAN, H. (2011). Convergence on functional k-nearest neighbor regression estimate with functional responses. *Electronic Journal of Statistics* **5**, 31–40

[18] LI, J. AND TRAN, L.T. (2007). Hazard rate estimation on random fields. *J. Multivariate Anal.* 98, 1337Ũ1355.

[19] MAILLOT, B. and LOUANI, D.(2008). *Propriétés asymptotiques de quelques estimateurs non-paramétriques pour des variables vectorielles et fonctionnelles.* Paris.

[20] MACK, Y. P. (1981). Local properties of k-NN regression estimates. *SIAM J. Algebr Discrete Methods*, **2** , 311-323.

[21] MULLER, S. and DIPPON, J. k-NN Kernel Estimate for Nonparametric Functional Regression in Time Series Analysis. Fachbereich Mathematik, Fakultat Mathematik und Physik (Pfaffenwaldring 57). 014/2011.

[22] Oliveira, P. E.(2005). *Nonparametic density and regression estimate functional data.* Tech. rep., Departamento de Matématica, Univercidade de Coimbra.

[23] Quintela-del-rio (2008). Hasard function given a fonctional variable : nonparametric estimation under strong mixing conditions. *J. Nonparametric stat.*. 413-430.

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

[25] Roussas, G. (1969). Nonparametric estimation of the transition distribution function of a Markov process. *Annals of Mathematical Statistics*, 40 1386–1400.

[26] Roussas, G. (1989). Nonparametric estimation in mixing sequences of random variables. *J. Statist. Plann.*, 18 135–149.

[27] R.M. Royall. (1966). *A class of nonparametric estimates of a smooth regression function*, Ph.D. Diss., Stanford University.

[28] Samanta, M. (1989). Non-parametric estimation of conditional quantiles. *Statist. Proba. Letters*,**7**, 407–412.

[29] Samanta, M. and Thavaneswaran, A. (1990) Non-parametric estimation of conditional model. *Comm. Statist. Theory and Meth.*, **16** , 4515-4524.

[30] Stone, C. J. (1977). Consistent nonparametric regression. *Ann. Statist*, **5**, 595Ű-645.

[31] Waston, G. S. and Leadbetter, M. R. (1964). Hasard analysis. *I. Biometrika*, **51**, 175Ű-184.

Chapitre 3

Estimation of conditional hazard function by the method of k Nearest Neighbor with α-mixing data.

Estimation of conditional hazard function by the method of k Nearest Neighbor with α-mixing data

Mohammed Kadi Attouch , Fatima Zohra Belabed
Lab. de statistique et processus stochastique,
Univ. Djillali Liabès, Sidi Bel Abbès,
BP 89, Sidi Bel Abbès 22000, Algeria

abstract : In this paper, we consider the problem of nonparametric estimation of the conditional hazard function, based on a sample of functional α-mixing data. More precisely, we give a dependent sequence $(X_i, Y_i)_{i=\overline{1,n}}$ valued in $(\mathcal{E} \times \mathbb{R})$ where $\mathcal{E}$ is a semi metric space, and we investigate an estimator of our function by the k Nearest Neighbors (k-NN) method. Our principal aim here is to study the asymptotic properties of our k-NN estimator. The almost complete convergence (its corresponding rates) is easily derived using the properties referred by Ferraty et al. ([12]) for the conditional distribution and conditional density estimates. The asymptotic normality is established for the two mentioned estimates referred by Laksaci and Mechab ([10]) and Quintela-del-Río ([21]). Finally, the applicability of our results is then verified and validated for simulated and some real data.

Keywords : Functional data ; nonparametric regression, k-NN estimator, the conditional hasard function,rate of convergence, random bandwidth, α-mixing data.

Ams : 62G05 · 62G08 · 62G20 · 62G35.

3.1 Introduction

In many fields of applications, as in econometric or seismology, the assumption of independence between the variables observed is not carried out. For that, it is necessary to introduce a structure of probability which makes control the dependence between the variables ; it is the method which present the assumption of the mixture.

The nonparametric estimation of the hazard function and the conditional hazard function has been extensively discussed in the literature. Beginning with Watson and Leadbetter ([25]). Since, several results have been added.

The first result was given by Roussas ([23]). For recent advances and references, Ferraty et al. ([12]) established the almost complete convergence in the (i.i.d) and censured cases. Their results has been extended to dependent case by Quintela-del-Río ([21]). Laksaci and Mechab ([10]) consider too the spacial case, they studied the almost complete convergence. More recently, Ferraty et al. ([13]) give the uniforme version of the almost complete convergence rate in (i.i.d) case.

In this paper, we consider the method of k Nearest Neighbors for the estimate of the conditional hazard function in case α mixing data, which depend of number of neighbors at the point of interest which we want to make a prediction. On the contrary, the traditional estimate depend on the real valued non random bandwidth sequence (h_n) (see Ferraty et al. ([12])).

The organisation of our paper is as follows. In section 2, we introduce the k-NN method estimate and we present the assumptions, after we propose in section 3 the asymptotic properties of the k-NN estimator of the conditional hazard function ; the almost complete convergence (its corresponding rates) and the asymptotic normality. Section 4 is devoted to some discuss on the applicability of our asymptotic results and some technical tools are given in section 5. Finally, the proofs of the auxiliary results are relegated to the appendix in section 6.

3.2 Method and assumptions

Let $(X_i, Y_i)_{i=\overline{1,n}}$ be a dependent sequence identically distributed as (X, Y), where the variable X is of functional nature and Y is a scalar. Formally, we will consider that X is a random variable valued in a semi metric functional space $(\mathcal{E}, d)$

The k-NN estimate of the conditional hazard function of Y given X is defined for any $y \in \mathbb{R}$ and any $x \in \mathcal{E}$ by :

$$\widehat{h}^x(Y = y) = \frac{\widehat{f}^x(y)}{1 - \widehat{F}^x(y)} \tag{3.1}$$

with
$\widehat{F}^x(y)$: is the estimator of the conditional distribution function concluded

from Ferraty et al. ([12]) and Burba et al. ([4]) by :

$$\widehat{F}^x(y) = \frac{\sum_{i=1}^{n} K(H_n^{-1}d(x,X_i))R(g_n^{-1}(y-Y_i))}{\sum_{i=1}^{n} K(H_n^{-1}d(x,X_i))}, if \sum_{i=1}^{n} K(H_n^{-1}d(x,X_i)) \neq 0. \tag{3.2}$$

Otherwise $\widehat{F}^x(y) = 0$.
$\widehat{f}^x(y)$: is the estimator of the conditional density function obtained by differentiating (3.2) with :

$$\widehat{f}^x(y) = \frac{\sum_{i=1}^{n} K(H_n^{-1}d(x,X_i))g_n^{-1}R'(g_n^{-1}(y-Y_i))}{\sum_{i=1}^{n} K(H_n^{-1}d(x,X_i))}, if \sum_{i=1}^{n} K(H_n^{-1}d(x,X_i)) \neq 0 \tag{3.3}$$

otherwise $\widehat{f}^x(y) = 0$, with K is an asymmetrical kernel, H_n is a positive random variable, that is defined as :

$$H_n(x) = \min\{\epsilon \in \mathbb{R}^+ / \sum_{i=1}^{n} 1\!\!1_{B(x,\epsilon)}(X_i) = k\} \tag{3.4}$$

where :

$$B(x,\epsilon) = \{x' \in \mathcal{E}; d(x,x') < \epsilon\}.$$

R is a distribution function and $(g_n)_{n\in\mathbb{N}}$ is a sequence of strictly positive real numbers (depending on n).

To prove the asymptotic properties of the k-NN estimator of the conditional hazard function and to emphasize differences between the k-NN method and the traditional kernel approach, we need some results of Ferraty et al. ([12]) by :

$$\widehat{h}^x_{kernel}(Y=y) = \frac{\widehat{f}^x_{kernel}(y)}{1-\widehat{F}^x_{kernel}(y)} \tag{3.5}$$

with :

$$\widehat{f}^x_{kernel}(y) = \frac{\sum_{i=1}^{n} K(h_n^{-1}d(x,X_i))g_n^{-1}R'(g_n^{-1}(y-Y_i))}{\sum_{i=1}^{n} K(h_n^{-1}d(x,X_i))} \tag{3.6}$$

and

$$\widehat{F}^x_{kernel}(y) = \frac{\sum_{i=1}^{n} K(h_n^{-1}d(x,X_i))R(g_n^{-1}(y-Y_i))}{\sum_{i=1}^{n} K(h_n^{-1}d(x,X_i))}. \tag{3.7}$$

Where K is a kernel function, R is a distribution function, $(h_n)_{n\in\mathbb{N}}$ is a non-random bandwidth and $(g_n)_{n\in\mathbb{N}}$ is a sequence of strictly positive numbers.

Prior to the presentation of our main results, we outline some assumptions. As in any nonparametric functional data problem, the behavior of the estimates is controlled by the concentration properties of the functional variable X around x :

(H1)

$$\varphi_x(\epsilon) \quad = \quad \mathbb{P}\left(X \in B(x,\epsilon)\right) > 0$$

where $B(x,\epsilon)$ being the ball of center x and radius ϵ. Namely ;

$$\forall \epsilon > 0 : \varphi_x(\epsilon) = \mathbb{P}\left[X \in \{x' \in \mathcal{E}, d(x,x') < \epsilon\}\right] > 0.$$

(See Ferraty and View ([11]) for more details).

We distinguish two types of kernel function K :

(H2) K is a kernel of type I, so that : there exist two real constants $C_1, C_2; 0 < C_1 < C_2 < \infty$; such that ;

$$C_1 1\!\!1_{[0,1]} < K < C_2 1\!\!1_{[0,1]}.$$

K is a kernel of type II, so that : the support of K is $[0,1]$ and if its derivative K' exists on $[0,1]$ and satisfies ; for two real constants $-\infty < C_1 < C_2 < 0$;

$$C_1 < K' < C_2.$$

Furthermore, we also suppose that : $\exists C_3 > 0, \exists \epsilon_0$:

$$\forall \epsilon < \epsilon_0, \int_0^\epsilon \varphi_x(u)du > C_3 \epsilon \varphi_x(\epsilon).$$

(H3) Assume that the conditional moment of the response random variable Y is bounded :

$$\forall m \geq 2; \mathbb{E}[\mid Y \mid^m /X = x] = \sigma_m(x) < \infty$$

with $\sigma_m(.)$ continuous on x.

(H4) N_x will denote a fixed neighboorhood of x, S will be a fixed compact subset of $\mathbb{R}$:

We will consider two kinds of nonparametric models. The first one is called the "Lipschitz-type" model that is defined :

$$Lip_{\mathcal{E}\times\mathbb{R}} : \begin{cases} f : \mathcal{E} \times \mathbb{R} \to \mathbb{R}, \forall (x_1, x_2) \in N_x^2, \forall (y_1, y_2) \in S^2, \exists C < \infty; \exists \alpha, \beta > 0 \\ |f(x_1, y_1) - f(x_2, y_2)| \leq C \left(d(x_1, x_2)^\alpha + |y_1 - y_2|^\beta \right). \end{cases}$$

(H5) The second one called the "Continuity type" model that is defined as :

$$C^0_{\mathcal{E}\times\mathbb{R}} = \{ f : \mathcal{E} \times \mathbb{R} \to \mathbb{R}; \forall x' \in N_x, \lim_{d(x,x')\to 0} f(x', y) = f(x, y) \}.$$

(H6) We need the followings additional assumptions on the distribution of two distinct pair (X_i, Y_i) and (X_j, Y_j), we assume that

$$\forall i \neq j : \mathbb{E}\left[Y_i Y_{j/X_i X_j} \right] \leq c < \infty$$

and the joint distribution function :

$$\Psi_{i,j}(h) = \mathbb{P}\left[(X_i, Y_i) \in B(x, \epsilon) \times B(x, \epsilon) \right]$$

satisfy :

$$\exists \epsilon_1 \in (0, 1], 0 < G_x(h) = O[\varphi_x(h)^{1+\epsilon_1}]$$

where :

$$G_x(h) = \max_{i,j \in 1,\dots,n/i \neq j} \Psi_{i,j}(h).$$

(H7) The sequence (X_i, Y_i) is said to be α-mixing if :

$$\alpha(n) = \sup_k \sup_{A \in \mathcal{A}_1^k} \sup_{B \in \mathcal{A}_{k+1}^n} \mid \mathbb{P}(A \cap B) - \mathbb{P}(A).\mathbb{P}(B) \mid \rightarrow_{n\to\infty} 0$$

where : $\mathcal{A}_l^m$ is the σ algebra generated by $(X_i, Y_i) \leq i \leq m$.
Finally with respect to the mixing coefficient, we will impose an arithmetically decaying rate :

$$\alpha(n) = O(n^{-a}), with \qquad a > \frac{1+\epsilon_3}{\epsilon_1\epsilon_2}\; et \qquad g_n\varphi_x(h) = O(n^{-\epsilon_2}); \qquad 0 < \epsilon_2 < \epsilon_3 < 1.$$

(H8) We suppose that $k = k_n$ is a sequence of positive real numbers such that $\frac{k_n}{n} \to 0$ and $\frac{\log n}{k_n} \to 0$.

(H9) Assume that for $S_n(x) = \max[S_{n,1}(x), S_{n,2}(x)]$, the rate of the α-mixing b : there exists $\theta > 2$ such that

$$S_n^{-(b+1)} = o(n^{-\theta}).$$

3.3 Asymptotic properties of k-NN method

3.3.1 Almost complete convergence

We are interested in this section, by the almost complete convergence [1] and rate of convergence [2] of the k-NN estimator of the conditional hazard function $\widehat{h}^x(y)$, where here (X_i, Y_i) is a sequence of α-mixing random variables (see $(H7)$).
To show the almost complete convergence of the k-NN estimator, we remind some results proved by Ferraty et al. ([12]) of $\widehat{h}^x_{kernel}(y)$ define in (3.5) :

Theorem 3.3.1 *Under the "continuity type", suppose (H1)-(H9), then we have :*

$$\widehat{h}^x_{kernel}(y) \longrightarrow h^x_{kernel}(y). \qquad a.co.$$

1. Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of real random variables. We say that $(X_n)_{n\in\mathbb{N}}$ converges almost completely (a.co.) to some r.r.v. X if and only if : $\forall\epsilon > 0, \sum_{n=1}^{\infty} \mathbb{P}[|X_n - X| > \epsilon] < \infty$.

2. Let $(u_n)_{n\in\mathbb{N}}$ be a sequence of positive real number ; We say that $X_n = O_{a.co.}(u_n)$ if and only if : $\exists\epsilon > 0$,so that, $\sum_{n=1}^{\infty} \mathbb{P}[|X_n| > \epsilon u_n] < \infty$. This kind of convergence implies both almost sure convergence and convergence in probability.

Theorem 3.3.2 *Under the "lipschitz type" model and the hypotheses (H1)-(H9), we have :*

$$\widehat{h}^x_{kernel}(y) - h^x_{kernel}(y) = O(h_n^{\alpha}) + O(g_n^{\beta}) + O\left(\sqrt{\frac{\log n}{n g_n \varphi_x(h)}}\right).$$

Now, let us focus on the k-NN method. In the first, we state the almost complete convergence of $\widehat{h}^x(y)$ define in (3.1).

Theorem 3.3.3 *Under the "Continuity type" model and the hypotheses(H1)-(H9), then we have :*

$$\lim_{n\to\infty} \widehat{h}^x(y) = h^x(y). \qquad a.co.$$

Proof. We consider the following decomposition :

$$\widehat{h}^x(y) - h^x(y) = \frac{1}{1-\widehat{F}^x(y)}[\widehat{f}^x(y) - f^x(y)] + h^x(y)\frac{1}{1-\widehat{F}^x(y)}[\widehat{F}^x(y) - F^x(y)]. \tag{3.8}$$

As a direct consequence of the last decomposition and the following Lemmas, we obtain the almost complete convergence of the conditional hazard function. The results of Lemma (3.3.1) are showed in the appendix and the proof of Lemma (3.3.2) is given in Chapter 2. ■

Lemma 3.3.1 *Under the hypotheses of Theorem (3.3.3), we have :*

$$\lim_{n\to\infty} \widehat{f}^x(y) = f^x(y). \qquad a.co. \tag{3.9}$$

And

$$\lim_{n\to\infty} \widehat{F}^x(y) = F^x(y). \qquad a.co. \tag{3.10}$$

Lemma 3.3.2 *Under the hypotheses of Theorem (3.3.3), we have :*

$$\exists \delta > 0, \sum_{n\in\mathbb{N}} \mathbb{P}[1 - \widehat{F}^x(y)) < \delta] < \infty. \tag{3.11}$$

Theorem 3.3.4 *Assume that (H1)-(H9) verifies, and under the "Lipschitz type", then we have :*

$$\widehat{h}^x(y) - h^x(y) = O\left(\varphi_x^{-1}(\frac{k_n}{n})^{\alpha}\right) + O\left(g_n^{\beta}\right) + O\left(\sqrt{\frac{\log n}{k_n g_n}}\right). \qquad a.co.$$

Proof. We consider the decomposition (3.8), and the proof of this Theorem is a consequence of these results which the proof is postponed to the appendix. ■

Lemma 3.3.3 *Under the hypotheses of Theorem (3.3.4), we have :*

$$\widehat{f}^x(y) - f^x(y) = O\left(\varphi_x^{-1}(\frac{k_n}{n})^\alpha\right) + O(g_n^\beta) + O\left(\sqrt{\frac{\log n}{k_n g_n}}\right). \qquad a.co. \quad (3.12)$$

Lemma 3.3.4 *Under the hypotheses of Theorem (3.3.4), we have :*

$$\widehat{F}^x(y) - F^x(y) = O\left(\varphi_x^{-1}(\frac{k_n}{n})^\alpha\right) + O(g_n^\beta) + O\left(\sqrt{\frac{\log n}{k_n}}\right). \qquad a.co. \quad (3.13)$$

3.3.2 Asymptotic normality

This section contain the results of the asymptotic normality of $\widehat{h}^x(y)$. Before giving the theorems, we need some results obtained by Quintela-del-Río ([21]) for the kernel estimator of the conditional hazard function :

Theorem 3.3.5 *Assume that (H1)-(H11) holds, then we have :*

$$\left(\frac{n\varphi_x(h)g_n}{\sigma_h^2(x,y)}\right)^{1/2}\left[\widehat{h}^x_{kernel}(y) - h^x(y)\right] \xrightarrow{\mathcal{D}} \mathcal{N}(0,1) \qquad as\ n \to \infty \qquad (3.14)$$

where

$$\sigma_h^2(x,y) = \frac{1}{1-F^x(y)}h^x(y)\int R'^2(t)dt. \qquad (3.15)$$

This theorem is a result of the two followings theorems :

Theorem 3.3.6 *Assume that (H1)-(H11) holds, then we have :*

$$\left(\frac{(n\varphi_x(h))^{1/2}}{\log n}\right)\left[\frac{\widehat{F}^x_{kernel}(y) - F^x(y)}{\sigma_F(x,y)}\right] \xrightarrow{\mathcal{D}} \mathcal{N}(0,1) \qquad as\ n \to \infty \qquad (3.16)$$

where

$$\sigma_F^2(x,y) = F^x(y)[1-F^x(y)]. \qquad (3.17)$$

Theorem 3.3.7 *Assume that (H1)-(H11) holds, then we have :*

$$\left(\frac{n\varphi_x(h)g_n}{\sigma_f^2(x,y)}\right)^{1/2}\left[\widehat{f}^x_{kernel}(y)-f^x(y)\right]\xrightarrow{\mathcal{D}}\mathcal{N}(0,1)\qquad as\ n\to\infty \tag{3.18}$$

where

$$\sigma_f^2(x,y)=f^x(y)\int R'^2(t)dt. \tag{3.19}$$

To obtain the asymptotic normality of the k-NN estimator of the conditional hazard function defined in (3.1), we have to add the followings assumptions :

(H10) $\lim_{n\to\infty}\left(g_n^2-\varphi_x^{-1}\left(\frac{k}{n}\right)\right)\sqrt{k_n}=0$ and $\frac{1}{k_ng_n}=o(g_n^\beta)$.

(H11) A sequence for integer $q_n\to\infty$, with $q_n=o\left((k_ng_n)^{\frac{1}{2}}\right)$, exists such that $(\frac{n}{g_n})^{1/2}\alpha(q_n)\to 0$.

Theorem 3.3.8 *Assume that (H1)-(H11), holds, then for any $x\in\mathcal{A}$, we have :*

$$\left(\frac{k_ng_n}{\sigma_h^2(x,y)}\right)^{1/2}\left[\widehat{h}^x(y)-h^x(y)\right]\xrightarrow{\mathcal{D}}\mathcal{N}(0,1)\qquad as\ n\to\infty \tag{3.20}$$

where

$$\sigma_h^2(x,y)=\frac{\alpha_2h^x(y)}{\alpha_1^2\left(1-F^x(y)\right)}\quad \left(with : \alpha_j=K^j(1)-\int_0^1(K^j)'(s)\alpha_x(s)\,ds\quad for\ j=1,2\right) \tag{3.21}$$

$$\mathcal{A}=\{x\in\mathcal{E},\ f^x(y)[1-F^x(y)]\neq 0\}$$

$\xrightarrow{\mathcal{D}}$ *means the convergence in distribution.*

Proof. We consider the decomposition (3.8) and the proof of Theorem (3.3.8) is a consequence of the followings lemmas. ■

Lemma 3.3.5 *Under the conditions (H1)-(H11), we have :*

$$\frac{(k_n)^{1/2}}{\sigma_F(x,y)\log n}\left[\widehat{F}^x(y)-F^x(y)\right]\xrightarrow{\mathcal{D}}\mathcal{N}(0,1)\quad as\ n\to\infty$$

where

$$\sigma_F^2(x,y)=F^x(y)\left[1-F^x(y)\right]. \tag{3.22}$$

Lemma 3.3.6 *Under the hypotheses of Theorem (3.3.8), we have :*

$$\left(\frac{k_n g_n}{\sigma_f^2(x,y)}\right)^{1/2}\left[\widehat{f}^x(y)-f^x(y)\right]\xrightarrow{\mathcal{D}}\mathcal{N}(0,1)\quad as\ n\to\infty$$

where

$$\sigma_f^2(x,y)=f^x(y)\int R'^2(t)dt. \tag{3.23}$$

Lemma 3.3.7 *Under the conditions (H1)-(H11), we have :*

$$(1-\widehat{F}^x(y))\to(1-F^x(y))\quad in\ probability.$$

3.4 Applications

3.4.1 A Simulation study

In this section we will show the effectiveness of k-NN method compared to the kernel estimation using simulated data. For this we considered a sample of a diffusion process on interval $[0,1]$, $Z_1(t)=2-\cos(\pi tW)$ and $Z_2(t)=\cos(\pi tW)$, where W is the standard normal distribution and take $X(t)=AZ_1(t)+(1-A)Z_2(t)$, where A is random variable Bernoulli distributed. We carried out the simulation with 200-sample of the curve X which is represented by the following graph :

For the scalar response variable, we took $Y=Ar_1(X)+(1-A)r_2(X)+\varepsilon$ where r_1 (resp. r_2) is the nonlinear regression model $r_1(X)=0.25\times\int_0^1(X'(t))^2dt+\epsilon$, with ϵ is $U([0,0.5])$ (resp. $r_2(X)$ is the null function) and ε is an α-mixing process generated by this model :

$$\varepsilon_i=\frac{1}{\sqrt{2}}(\varepsilon_{i-1}+\eta_i),\ i=1,\ldots,200,$$

with η_i is normally independent identically distributed random variables. We generated standard normal distributed random variable ε_0 independently of η_i.

We choose a quadratic kernel K defined by :

$$K(x)=\frac{3}{2}(1-x^2)\mathbb{1}_{[0,1]}.$$

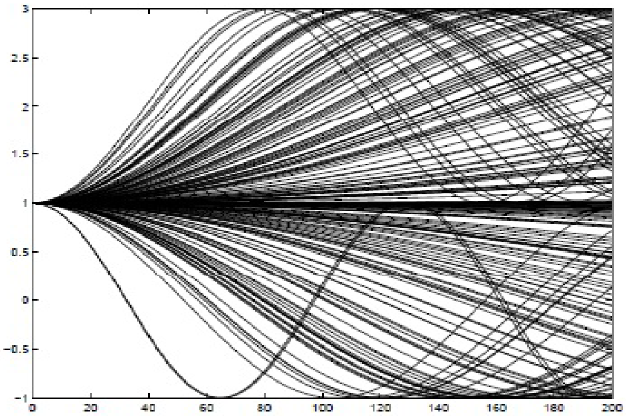


FIGURE 3.1 – The 200 curves X.

In practice, the semi-metric choice is based on the regularity of the curves X. For this we use the semi-metric defined by L_2-distance between the q^{th} derivatives of the curves. In order to evaluate the MSE (Mean Square Error) we proceed by the following algorithm :

- *Step 1.* We split our data into two subsets ; the first sample, of size n=120 corresponds to the learning sample which will be used, as a sample, to compute our conditional hazard function estimators for the 80 remaining curves (considered as the test sample).
  - $(X_j, Y_j)_{j\in J}$ learning sample,
  - $(X_i, Y_i)_{i\in I}$ test sample.
- *Step 2.*
  - We use the learning sample for computing the hazard function estimator $\widehat{h}_j$, for all $j \in J$.
  - We set : $i^* = \arg\min_{j\in J} d(X_i, X_j)$.
  - We put : $\forall i \in I$

$$\widehat{T_i} = \widehat{h}^{X_{i^*}}(Y_i) \text{ for kernel method.}$$
$$\widehat{T_i} = \widehat{h}^{X_{k_{opt}}}(Y_i) \text{ for } k\text{-NN method,}$$

where :

$$X_{i^*} : \text{is the nearest curve to } X_j$$
$$k_{opt} : \arg\min_a(CV(a))$$

with :

$$CV(a) = \frac{1}{n}\left[\sum_{i\in J}\int(\widehat{f}^{-i}_{(a,b)}(X_i, y))^2 dy - 2\sum_{i\in J}\widehat{f}^{-i}_{(a,b)}(X_i, Y_i)\right]$$

and :

$$\widehat{f}^{-k}_{(a,b)}(x,y) = \frac{b^{-1}\sum_{i\in J, i\neq k} K\left(\frac{d(x,X_i)}{a}\right) R\left(\frac{y-Y_i}{b}\right)}{\sum_{i\in J} K\left(\frac{d(x,X_i)}{a}\right)}.$$

– *Step 3.* The error used to evaluate this comparison is the mean of square error (MSE) expressed by

$$\frac{1}{card(I)} \sum_{i\in I} \left| h(Y_i) - \widehat{T}(X_i, Y_i) \right|^2,$$

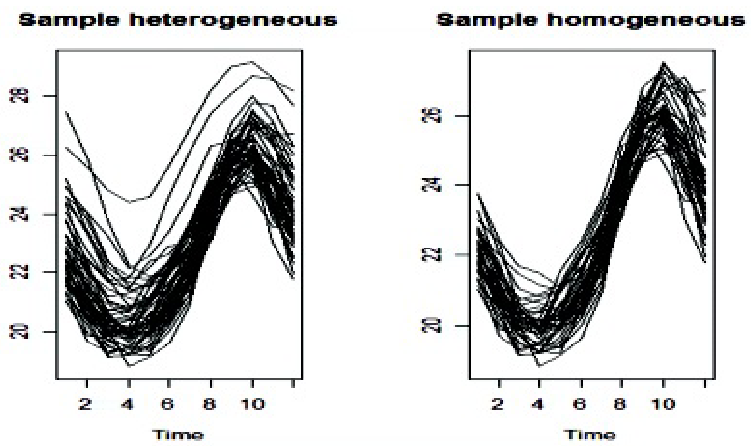
where $\widehat{T}$ designate the estimator used : kernel or k-NN method estimation and h is the true hazard function.

Consequently, the k-NN method gives better results than the kernel method. This is confirmed by the MSE-k-NN$=$ 2.170035 and MSE-Kernel $=$ 13.66790.

3.4.2 Real data application

To highlight the efficiency and robustness of the method of k nearest neighbors with respect to the kernel method in estimating the conditional hazard function, we will test these two methods in the presence or not of heterogeneous data.
To do this, based on the study of Burba et al. (2009) which emphasizes the effect of the nature of the data (homogeneous or heterogeneous) on the quality of the estimate, especially the superiority of the k-nearest neighbors in the presence of very heterogeneous data.
For this purpose, we apply the described algorithm used in the simulation study to El Ni*ño* time series which gives monthly Sea Surface temperatures. More information and other data-sets can be found about the phenomenon called El Ni*ño* in web site http ://www.cpc.ncep.noaa.gov/data/indices (see also :http ://kdd.ics.uci.edu/databases/el_nino/el_nino.data.html). Our study concerns the monthly time series of the Sea Surface Temperature (SST) from June, 1950 up to May, 2004.
A useful way to display such a time series consists in cutting it into 54 pieces or 54 "annual curves" (see Fig. 1). More precisely, let $(X(k), k = 1, \ldots, 648$ be our El Niño time series. We can build, for $i = 1, \ldots, 54$, the following subsequences : $\forall t \in \{1, \ldots, 12\}$, $X_i(t) = X(12 \times (i-1) + t)$,

$X_i = (X_i(1), ..., X_i(12))$ corresponding to the variations of the SST at the i^{th} year.
The sample of size 54 was splitted into learning sample of size 44 (with all data), 30 (without the heterogeneous data, 14 values) and testing sample of size 10. Figue 3.4.2 displays the curves of learning sample for all data and the curves of learning sample without the heterogeneous one.



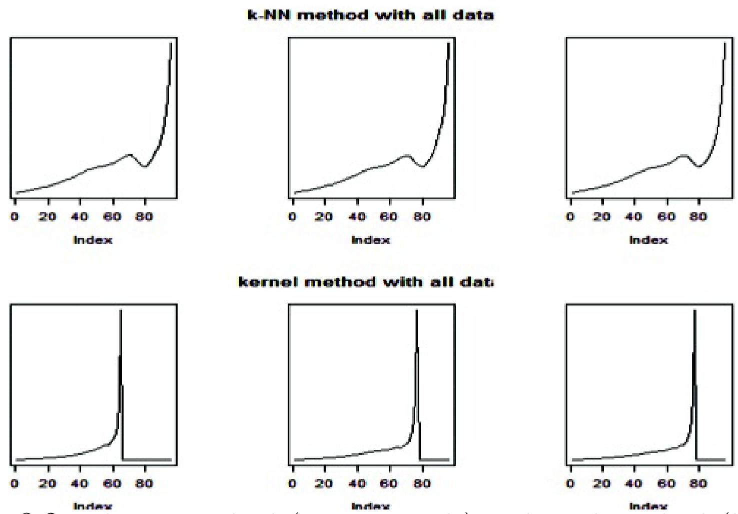


FIGURE 3.2 – k-NN method (upper panels) *vs* kernel method (lower panels) of conditional hazard function far all data.

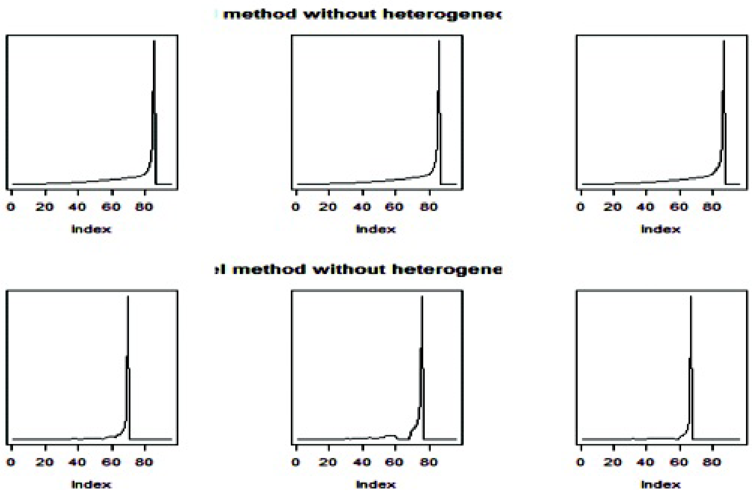


FIGURE 3.3 – k-NN method (upper panels) *vs* kernel method (lower panels) of conditional hazard function without heterogeneous data.

We plot the conditional hazard function estimated for the first 3 values of testing sample, Figure 3.2 depicts that the k-NN method in presence of heterogeneous data give better estimation of the conditional hazard function prediction (regular function) than the kernel method estimation (non-regular function) and when the data are homogeneous the two method give the same result that can be easily seen in Figure 3.3.

3.5 General technical tools

Let $(A_i, B_i)_{i\in\mathbb{N}}$ be a sequence of random variables with values in $(\Omega \times \mathbb{R}, \mathcal{A} \otimes \mathcal{B})$, independent but not necessarily identically distributed,where $(\Omega, \mathcal{A})$ is a general measurable space, let $G : \mathbb{R} \times \Omega \to \mathbb{R}^+$ a measurable function such that : $\forall w, w' \in \mathbb{R}$:

$$w \leq w' \Longrightarrow G(w, z) \leq G(w', z); \qquad \forall z \in \Omega.$$

Let c be a not random positive real number and T a real random variable, we define, $\forall n \in \mathbb{N}^*$

$$C_n(T) = \frac{\sum_{i=1}^{n} B_i G(T, A_i)}{\sum_{i=1}^{n} G(T, A_i)}.$$

Lemma 3.5.1 COLLOMB ([5])-BURBA ET AL. ([4])
Let $(D_n)_{n\in\mathbb{N}}$ be a sequence of real random variables and $(u_n)_{n\in\mathbb{N}}$ be a decreasing sequence of positive numbers.
If $l = \lim u_n \neq 0$, and if, for all increasing sequence $\beta_n \in]0,1[$, there exist two sequences of real random variables $(D_n^+(\beta_n))_{n\in\mathbb{N}}$ and $(D_n^-(\beta_n))_{n\in\mathbb{N}}$:

(L1) $\forall n \in \mathbb{N}; D_n^- \leq D_n^+$, *and* $\mathbb{1}_{D_n^- \leq D_n \leq D_n^+} \to 1 \qquad a.co.$

(L2) $$\frac{\sum_{i=1}^{n} G(D_n^-, A_i)}{\sum_{i=1}^{n} G(D_n^+, A_i)} - \beta_n = O(u_n) \qquad a.co.$$

(L3)
$$C_n(D_n^-) - c = O(u_n) \qquad a.co.$$
$$C_n(D_n^+) - c = O(u_n). \qquad a.co.$$

Then :
$$C_n(D_n) - c = O(u_n). \qquad a.co.$$

If l=0 and if (L1), (L2), (L3) hold for any increasing sequence $\beta_n \in]0,1[$ with limit 1, the same result holds.

Lemma 3.5.2 BURBA ET AL. ([4])
Let $(D_n)_{n\in\mathbf{N}}$ be a sequence of real random variables and $(v_n)_{n\in\mathbb{N}}$ be a decreasing positive sequence.
If $l' = \lim v_n \neq 0$, and if, for all increasing sequence $\beta_n \in]0,1[$, there exist two sequences of real random variables $(D_n^+(\beta_n))_{n\in\mathbb{N}}$ and $(D_n^-(\beta_n))_{n\in\mathbb{N}}$:

(L'1) $\forall n \in \mathbb{N}; D_n^- \leq D_n^+$, *and* $\mathbb{1}_{D_n^- \leq D_n \leq D_n^+} \to 1 \qquad a.co.$

(L'2) $$\frac{\sum_{i=1}^{n} G(D_n^-, A_i)}{\sum_{i=1}^{n} G(D_n^+, A_i)} - \beta_n = o(v_n) \qquad a.co.$$

(L'3)

$$C_n(D_n^-) - c = o(v_n) \qquad a.co.$$

$$C_n(D_n^+) - c = o(v_n). \qquad a.co.$$

Then :

$$C_n(D_n) - c = o(v_n). \qquad a.co.$$

If l'=0 and if (L'1), (L'2), (L'3) hold for any increasing sequence $(\beta_n) \in]0,1[$ *with limit 1, the same result holds.*

Burba et al. ([4]) use in their consistency proof of the k-NN kernel estimate for independent data a Chernoff-type exponential inequality to check conditions $(L1)$ or $(L'1)$.

Lemma 3.5.3 *Let* (X,Y) *be a valued random vector in* $\mathbb{R}^{\shortmid} \times \mathbb{R}$ *such that* $Y \in L_p(P)$ *for some* $p \in [1,+\infty[$. *Let* d *a real number such that* $\|Y+d\|_p > 0$ *and* $\epsilon \in (0, \|Y+d\|_p)$. *Then there exists a random variable* Z *such that :*
- $P_Z = P_Y$ *and* Z *is independent of* X.
- $P(\mid Z - Y \mid > \epsilon) \leq 11 \left(\frac{\|Y+d\|_p}{\epsilon}\right)^{\frac{p}{2p+1}} [\alpha(\sigma(X), \sigma(Y))]^{\frac{p}{2p+1}}$, *where* $\sigma(X)$ *is the* σ *algebra generated by* X.

Lemma 3.5.4 *Let* (X_i) *a arithmetically* α *mixing sequence in the semi metric space* (E,d) *with* $\alpha(n) \leq cn^{-b}$, b *and* $c > 0$. *Define* $\Delta_i(x) = \mathbb{1}_{B(x,h)}(X_i)$.*then we have :*

$$\sum_{i=1}^{n} \sum_{j=1}^{n} \mid Cov(\Delta_i(x), \Delta_j(x)) \mid = O\left(n\varphi_x(h)\right) + O(\chi(x,h)^{1-s} n^{1+s})$$

where $\chi(x,h) = \max G_x(h), \varphi_x(h)^2$ *and* $s = \dfrac{1}{b+1}$.

3.6 Appendix

Proof of Lemma 3.3.1 :
On the one hand, to prove the first result, we apply lemma (3.5.2). The main

difference to the proof of the independent case concerns verification of (L'1). To verify (L'2) and (L'3), we need small modifications :

$$\begin{cases} v_n = 1 \\ H_n = D_n \\ \widehat{f}^x(y) = C_n(D_n) \\ f^x(y) = c. \end{cases} \tag{3.24}$$

Choose $\beta_n \in (0,1)$, (D_n^-) and (D_n^+) such that :

$$\begin{cases} \varphi_x(D_n^+) = \dfrac{1}{\sqrt{\beta_n}}\varphi_x(h) = \dfrac{1}{\sqrt{\beta_n}}\dfrac{k_n}{n} \\ \varphi_x(D_n^-) = \sqrt{\beta_n}\varphi_x(h) = \sqrt{\beta_n}\dfrac{k_n}{n}. \end{cases} \tag{3.25}$$

Then we define

$$\begin{cases} h^+ = D_n^+ = \varphi_x^{-1}(\dfrac{1}{\sqrt{\beta_n}}\dfrac{k_n}{n}) \\ h^- = D_n^- = \varphi_x^{-1}(\sqrt{\beta_n}\dfrac{k_n}{n}). \end{cases} \tag{3.26}$$

FERRATY and VIEU ([11]) proved under the conditions of Theorem (3.3.3) that :

$$\widehat{f}^x_{kernel}(y) \longrightarrow_{a.co} f^x(y) \tag{3.27}$$

and to apply this result, we have to show that the covariance term S_n fulfills following condition : $\exists\theta > 2$ such that :

$$S_n^{-(b+1)} = o(n^{-\theta})$$

with b is the rate of mixing coefficient. If the condition on the rate b of the mixing coefficient and $(H6)$ holds, we have by lemma (11-5) in ([11]) :

$$S_n(x) = O\left(\frac{n}{\varphi_x(h^-)}\right) = O\left(\frac{n^2}{k}\right)$$

and it is true for h^+. Consequently, we can apply the result (3.27), then we obtain :

$$C_n(D_n^-) \to c \qquad\qquad a.co.$$

and

$$C_n(D_n^+) \to c. \qquad\qquad a.co.$$

Thus condition $(L'3)$ is verified.
In parallel Ferraty and Vieu ([11]) proved under the conditions of Theorem (3.3.1) that :

$$\frac{1}{n\varphi_x(h)}\sum_{i=1}^{n} K(h^{-1}d(x,X_i)) \to 1 \qquad a.co.$$

under the conditions (3.25) and (3.26), we have :

$$\begin{cases} \dfrac{1}{n\varphi_x(D_n^-)}\displaystyle\sum_{i=1}^{n} K\left((D_n^-)^{-1}d(x,X_i)\right) \to 1 & a.co. \\ \dfrac{1}{n\varphi_x(D_n^+)}\displaystyle\sum_{i=1}^{n} K\left((D_n^+)^{-1}d(x,X_i)\right) \to 1 & a.co. \end{cases}$$

we get :

$$\frac{\displaystyle\sum_{i=1}^{n} K\left((D_n^-)^{-1}d(x,X_i)\right)}{\displaystyle\sum_{i=1}^{n} K\left((D_n^+)^{-1}d(x,X_i)\right)} \to \beta_n \qquad a.co.$$

so that $(L'2)$ is checked.
Finally, we prove $(L'1)$; the first part is obvious, and the second one that, $\forall\epsilon > 0$:

$$\sum_{n\geq 0} \mathbb{P}[|1\!\!1_{D_n^- < H_n < D_n^+} - 1| > \epsilon] < \infty.$$

Let $\epsilon > 0$ be fixed, we know that ;

$$\mathbb{P}\left[|1\!\!1_{D_n^- < H_n < D_n^+} - 1| > \epsilon\right] \leq \underbrace{\mathbb{P}[H_n < D_n^-]}_{A_1} + \underbrace{\mathbb{P}[H_n > D_n^+]}_{A_2}.$$

We star by A_1, then we have :

$$A_1 \leq \mathbb{P}\left[\sum_{i=1}^{n} 1\!\!1_{B(x,D_n^-)}(X_i) > k_n\right]$$

$$\leq \mathbb{P}\left[\sum_{i=1}^{n}\left(1\!\!1_{B(x,D_n^-)}(X_i) - \varphi_x(D_n^-)\right) > k_n - n\varphi_x(D_n^-)\right] = A_{1,n}. \qquad (3.28)$$

The second step, we centred the random variables, then the plan here is to split the data into a block schema and we applique lemma (3.5.3) ; we devise the set $\{1, 2, 3, ..., n\}$ into blocks of length $2l_n$, set $m_n = [n/2l_n]$ where $[.]$ is the gaussian bracket [3] and $f_n = n - 2l_n m_n < 2l_n$.
Let

$$U_n(j) = \sum_{i=(j-1)l_n+1}^{jl_n} \left(\mathbb{1}_{B(x,D_n^-)}(X_i) - \varphi_x(D_n^-) \right)$$

and we define :

$$B_{n,1} = \sum_{j=1}^{m_n} U_n(2j-1)$$

$$B_{n,2} = \sum_{j=1}^{m_n} U_n(2j)$$

$$R_n = \sum_{i=2l_n m_n}^{n} \left(\mathbb{1}_{B(x,D_n^-)}(X_i) - \varphi_x(D_n^-) \right)$$

thus, we obtain in (3.28)

$$A_{1,n} \le \mathbb{P}\left[B_{n,1} > \frac{k_n - n\varphi_x(D_n^-)}{3} \right] + \mathbb{P}\left[B_{n,2} > \frac{k_n - n\varphi_x(D_n^-)}{3} \right] + \mathbb{P}\left[R_n > \frac{k_n - n\varphi_x(D_n^-)}{3} \right]. \tag{3.29}$$

Now, we apply lemma (3.5.3) in the first term of (3.29) with :
- $d = l_n m_n$ lead to $0 < l_n m_n \le \|U_n(2j-1) + d\|_\infty \le 2l_n + l_n m_n$.
- $\epsilon = \dfrac{k_n - n\varphi_x(D_n^-)}{6m_n} = \dfrac{k_n - n\sqrt{\beta_n}\frac{k_n}{n}}{6m_n} = \dfrac{k_n(1-\sqrt{\beta_n})}{6m_n} \in (0, \|U_n(2j-1) + d\|_\infty)$.
- $Y = U_n(2j-1)$.

And we can construct $\widetilde{U}_n(2j-1)$ such that :
- The random variables $\widetilde{U}_n(2j-1)$ are independents
- $\widetilde{U}_n(2j-1)$ has the same distribution as $U_n(2j-1)$.

Then :

$$\mathbb{P}\left[|\widetilde{U}_n(2j-1) - U_n(2j-1)| > \epsilon \right] \le 11 \left[\frac{\|U_n(2j-1) + d\|_\infty}{\epsilon} \right]^{1/2}$$

$$. \sup |\mathbb{P}(AB) - \mathbb{P}(A)\mathbb{P}(B)|$$

3. $[x] = \max\{y \in \mathbb{Z}/z \le x\}, x \in \mathbb{R}$.

where the supremum is taken over all sets $A, B \in \sigma\left(U_n(1), U_n(3), ..., U_n(2m_n-1)\right)$.
This leads to :

$$\mathbb{P}\left[B_{n,1} > \frac{k_n - n\varphi_x(D_n^-)}{3}\right]$$

$$= \mathbb{P}\left[\sum_{j=1}^{m_n}\left(U_n(2j-1) + \widetilde{U}_n(2j-1) - \widetilde{U}_n(2j-1)\right) > \frac{k_n - n\varphi_x(D_n^-)}{3}\right]$$

$$\leq \mathbb{P}\left[\sum_{j=1}^{m_n}\widetilde{U}_n(2j-1) > \frac{k_n - n\varphi_x(D_n^-)}{6}\right] + \mathbb{P}\left[\sum_{j=1}^{m_n}(U_n(2j-1) - \widetilde{U}_n(2j-1)) > \frac{k_n - n\varphi_x(D_n^-)}{6}\right] \tag{3.30}$$

and by applying lemma (3.5.3) in the second term of (3.30), and some calculation we obtain :

$$\mathbb{P}\left[\sum_{j=1}^{m_n}(U_n(2j-1) - \widetilde{U}_n(2j-1)) > \frac{k_n - n\varphi_x(D_n^-)}{6}\right] \leq$$

$$\sum_{j=1}^{m_n}\mathbb{P}\left[(U_n(2j-1) - \widetilde{U}_n(2j-1)) > \frac{k_n - n\varphi_x(D_n^-)}{6m_n}\right] \leq$$

$$m_n\left(\frac{6m_n l_n(m_n+1)}{k_n(1-\sqrt{\beta_n})}\right)^{\frac{1}{2}}\alpha(l_n) \leq C\frac{n^2}{l_n^{\frac{3}{2}}k_n}\alpha(l_n)$$

we choose the sequence l_n such that :

$$l_n^a = \frac{n^2}{2^a r^a k_n} \tag{3.31}$$

with ; r is a positive constant, $a > \dfrac{2}{\gamma-1}$, and by the condition of the mixing coefficient b, we obtain :

$$C\frac{n^2}{l_n^{\frac{3}{2}}k_n}\alpha(l_n) = C.n^{(2-\gamma)(a-3/2-b)/a}$$

consequently, we trouve the second term of (3.30) is finie, then :

$$\mathbb{P}\left[\sum_{j=1}^{m_n}(U_n(2j-1) - \widetilde{U}_n(2j-1)) > \frac{k_n - n\varphi_x(D_n^-)}{6}\right] < \infty.$$

We passe now in the first term of (3.30) and we apply markov inequality for some $t > 0$, we obtain :

$$\mathbb{P}\left[\sum_{j=1}^{m_n} \widetilde{U}_n(2j-1) > \frac{k_n - n\varphi_x(D_n^-)}{6}\right] \leq e^{-t\frac{k_n - n\varphi_x(D_n^-)}{6}} \mathbb{E}\left[e^{t\sum_{j=1}^{m_n} \widetilde{U}_n(2j-1)}\right]$$

$$\leq e^{-t\frac{k_n - n\varphi_x(D_n^-)}{6}} \prod_{j=1}^{m_n} \mathbb{E}\left[e^{t\widetilde{U}_n(2j-1)}\right]. \tag{3.32}$$

Applying : $e^x \leq 1 + x + x^2$, we obtain for large n :

$$e^{t\widetilde{U}_n(2j-1)} \leq 1 + t\widetilde{U}_n(2j-1) + t^2\widetilde{U}_n(2j-1)^2$$

we passe at the step of esperance, and we know that ; $\mathbb{E}(\widetilde{U}_n(2j-1)) = 0$. After the calculations while we applying $e^x > 1 + x$ we get :

$$\mathbb{E}[e^{t\widetilde{U}_n(2j-1)}] \leq 1 + \mathbb{E}[t^2\widetilde{U}_n(2j-1)^2] \leq e^{t^2\mathbb{E}[\widetilde{U}_n(2j-1)^2]}.$$

Furthermore, because $\widetilde{U}_n(2j-1)$ and $U_n(2j-1)$ have the same distribution function and by some calculations, it follows that :

$$\sum_{j=1}^{m_n} \mathbb{E}[\widetilde{U}_n(2j-1)^2] \leq \sum_{i,j=1}^{n} \mid cov(\mathbb{1}_{B(x,D_n^-)}(X_i), \mathbb{1}_{B(x,D_n^-)}(X_i)) \mid .$$

Now we apply lemma (3.5.4) with $\varphi_x(D_n^-) = \sqrt{\beta_n}\frac{k_n}{n}$ and $k \sim n^\gamma$, then $\varphi_x(D_n^-) = O(n^{\gamma-1})$. Under $(H9)$ we have :

$$\mathbb{E}[e^{t\widetilde{U}_n(2j-1)}] \leq C_1\sqrt{\beta_n}k_n + C_2\chi(D_n^-)^{(1-s)}n^{(1+s)}$$

and under $(H6)$ we obtain :

$$\mathbb{E}[e^{t\widetilde{U}_n(2j-1)}] \leq C_1\sqrt{\beta_n}k_n$$

finally, we obtain from (3.32) with $t = r\frac{logn}{k_n}$ and $r > \frac{6}{1-\sqrt{\beta_n}}$

$$\mathbb{P}\left[\sum_{j=1}^{m_n} \widetilde{U}_n(2j-1) > \frac{k_n - n\varphi_x(D_n^-)}{6}\right] \leq e^{-t\frac{k_n - n\varphi_x(D_n^-)}{6}}$$

$$.e^{C_1 \frac{\sqrt{\beta_n} r^2 (logn)^2}{k_n}} .e^{C_2 \frac{\sqrt{\beta_n} r^2 (logn)^2 \chi(D_n^-)^{(1-s)} n^{(1+s)}}{k_n}}$$

$$\leq e^{-r\frac{1-\sqrt{\beta_n}}{6} logn} = n^{-r\frac{1-\sqrt{\beta_n}}{6}} < \infty$$

because :

$$e^{C_1 \frac{\sqrt{\beta_n} r^2 (logn)^2}{k_n}} \longrightarrow 1$$

and

$$e^{C_2 \frac{\sqrt{\beta_n} r^2 (logn)^2 \chi(D_n^-)^{(1-s)} n^{(1+s)}}{k_n}} \longrightarrow 1$$

and thereafter of (3.30), we conclue that :

$$\mathbb{P}\left[B_{n,1} > \frac{k_n - n\varphi_x(D_n^-)}{3}\right] < \infty.$$

Same way, we can show that :

$$\mathbb{P}\left[B_{n,2} > \frac{k_n - n\varphi_x(D_n^-)}{3}\right] < \infty.$$

Now,it is enough for us to prove :

$$\mathbb{P}\left[R_n > \frac{k_n - n\varphi_x(D_n^-)}{3}\right] < \infty.$$

Then we know that :

$$\begin{cases} |R_n| < 4l_n \\ and \dfrac{k_n - n\varphi_x(D_n^-)}{3} = O(k_n) \end{cases}$$

by the definition of l_n in (3.31) and $a > \frac{2}{\gamma - 1}$ we obtain

$$\mathbb{P}\left[R_n > \frac{k_n - n\varphi_x(D_n^-)}{3}\right] < \infty$$

consequently we have :

$$\sum_{1}^{\infty} A_{1,n} < \infty.$$

While following the same step to prove that

$$\sum_{1}^{\infty} A_{2,n} < \infty$$

then (L_1') is verified.

On the other hand, we prove the second result for Lemma (3.3.1). Indeed, we use the precedents steps with :

$$\begin{cases} v_n = 1 \\ H_n = D_n \\ \widehat{F}^x(y) = C_n(D_n) \\ F^x(y) = c. \end{cases} \tag{3.33}$$

Proof of Lemma 3.3.3 :
To prove this result, we use Lemma (3.5.1), the condition $(L1)$ is proven same manner as $(L'1)$ of Lemma (3.5.2), for this the proof can be omitted here. Choose β_n as an increasing sequence in $(0,1)$ with limit 1. Furthermore, we choose D_n^- and D_n^+ under (3.25). Ferraty and Vieu proved under the conditions of Theorem (3.3.3) that :

$$\widehat{r}_1(x) - 1 = O\left(\sqrt{\frac{\log n}{n\varphi_x(h)}}\right)$$

where :

$$\widehat{r}_1(x) = \frac{1}{n}\sum_{i=1}^{n} \frac{K\left(h_n^{-1}d(x,X_i)\right)}{\mathbb{E}K\left(h_n^{-1}d(x,X_1)\right)}$$

then

$$\frac{1}{n}\sum_{i=1}^{n} K\left(h_n^{-1}d(x,X_i)\right) - \varphi_x(h) = O\left(\sqrt{\frac{\log n}{n\varphi_x(h)}}\right). \tag{3.34}$$

If $(H6)$ holds then :

$$s_n(x) = O\left(\frac{n}{\varphi_x(h^+)}\right) = O\left(\frac{n^2}{k_n}\right)$$

the same is true for the bandwidth h^-.
And if $(H9)$ holds, we have for both bandwidth sequences h^+ and h^- :

$$s_n(x) = O\left(\frac{n^2}{k_n}\right) + O\left(\chi(x,h)^{1-s} n^{1+s}\right).$$

Using (3.34) under (3.25) we get :

$$\begin{cases} \frac{1}{n}\sum_{i=1}^{n} K\left(\frac{d(x,X_i)}{D_n^-}\right) = \sqrt{\beta_n}\frac{k_n}{n} + O\left(\sqrt{\frac{\log n}{k_n}}\right) \\ \frac{1}{n}\sum_{i=1}^{n} K\left(\frac{d(x,X_i)}{D_n^+}\right) = \frac{1}{\sqrt{\beta_n}}\frac{k_n}{n} + O\left(\sqrt{\frac{\log n}{k_n}}\right) \end{cases}$$

and we obtain :

$$\frac{\sum_{i=1}^{n} K\left(\frac{d(x,X_i)}{D_n^-}\right)}{\sum_{i=1}^{n} K\left(\frac{d(x,X_i)}{D_n^+}\right)} - \beta_n = O\left(\sqrt{\frac{\log n}{k}}\right)$$

so that, $(L2)$ is verifies. Now we are able to apply Lemma (6.15) for Ferraty and Vieu ([11]) under (3.25) and (3.26), we get :

$$C_n(D_n^+) - c = O\left(\varphi_x^{-1}\left(\frac{k_n}{n}\right)^{\alpha}\right) + O(g_n^{\beta}) + O\left(\sqrt{\frac{logn}{k_n}}\right)$$

and

$$C_n(D_n^-) - c = O\left(\varphi_x^{-1}\left(\frac{k_n}{n}\right)^{\alpha}\right) + O(g_n^{\beta}) + O\left(\sqrt{\frac{logn}{k_n}}\right)$$

so $(L3)$ is checked.

Proof of Lemma 3.3.4 :
To show this Lemma, we use the same idea of the proof of Lemma (3.3.3). For this Ferraty and Vieu ([11]) proved under the conditions of Theorem (3.3.4) that :

$$\widehat{r}_3(x) - \mathbb{E}\widehat{r}_3(x) = O\left(\sqrt{\frac{\log n}{nh_n\varphi_x(h)}}\right)$$

where

$$\widehat{r}_3(x) = \frac{1}{n}\sum_{i=1}^{n}\frac{K\left(h_n^{-1}d(x,X_i)\right)\Gamma_i(y)}{\mathbb{E}K\left(h_n^{-1}d(x,X_i)\right)}.$$

with

$$\Gamma_i(y) = R\left(g_n^{-1}(y-Y_i)\right)$$

then

$$\widehat{r}_3(x)-\mathbb{E}\widehat{r}_3(x) = \frac{1}{n}\sum_{i=1}^{n}\frac{K\left(h_n^{-1}d(x,X_i)\right)}{\mathbb{E}K\left(h_n^{-1}d(x,X_1)\right)}\Gamma_i(y)-\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left[\frac{K\left(h_n^{-1}d(x,X_i)\right)}{\mathbb{E}K\left(h_n^{-1}d(x,X_1)\right)}\Gamma_i(y)\right]$$

$$= \frac{1}{n\varphi_x(h)}\sum_{i=1}^{n}K\left(h_n^{-1}d(x,X_i)\right)\Gamma_i(y) - \mathbb{E}(\Gamma_1(y)/X_1).$$

Under the notations (3.25) and (3.26), we have :

$$\begin{cases} \dfrac{1}{n\varphi_x(h)}\displaystyle\sum_{i=1}^{n}K\left(\frac{d(x,X_i)}{D_n^-}\right)\Gamma_i(y) = \mathbb{E}(\Gamma_1(y)/X_1) + O\left(\sqrt{\frac{\log n}{g_n k_n}}\right) \\ \dfrac{1}{n\varphi_x(h)}\displaystyle\sum_{i=1}^{n}K\left(\frac{d(x,X_i)}{D_n^+}\right)\Gamma_i(y) = \mathbb{E}(\Gamma_1(y)/X_1) + O\left(\sqrt{\frac{\log n}{g_n k_n}}\right) \end{cases}$$

by this, we obtain :

$$\frac{\displaystyle\sum_{i=1}^{n}K\left(\frac{d(x,X_i)}{D_n^-}\right)}{\displaystyle\sum_{i=1}^{n}K\left(\frac{d(x,X_i)}{D_n^+}\right)} - \beta_n = O\left(\sqrt{\frac{\log n}{g_n k_n}}\right) .a.co.$$

that $(L2)$ is verified. Now, we apply the Theorem (3.3.2) under (3.25), we get :

$$C_n(D_n^-) - c = O\left(\varphi_x^{-1}(\frac{k}{n})^{\alpha}\right) + O(g_n^{\beta}) + O\left(\sqrt{\frac{\log n}{g_n k_n}}\right) a.co.$$

$$C_n(D_n^+) - c = O\left(\varphi_x^{-1}(\frac{k}{n})^{\alpha}\right) + O(g_n^{\beta}) + O\left(\sqrt{\frac{\log n}{g_n k_n}}\right) .a.co.$$

that verifies condition $(L3)$.

Proof of Lemma 3.3.5 :
To prove the asymptotic normality of $\widehat{F}^x(y)$ we pose :

$$\begin{cases} C_n(H_n) = \widehat{F}^x(y) \\ c = F^x(y). \end{cases} \tag{3.35}$$

Under (3.25) and (3.26), we obtain :

$$\frac{(k_n)^{1/2}}{\sigma_F(x,y)\log n}\left[C_n(H_n)-c\right] = \frac{(k_n)^{1/2}}{\sigma_F(x,y)\log n}\left[C_n(D_n^+)-c\right] + \frac{(k_n)^{1/2}}{\sigma_F(x,y)\log n}\left[C_n(H_n)-C_n(D_n^+)\right]. \tag{3.36}$$

It is clear, to verify the asymptotic normality of the conditional distribution function, it is enough to show that :

$$\frac{(k_n)^{1/2}}{\sigma_F(x,y)\log n}\left[C_n(D_n^+)-c\right] \to N(0,1)$$

$$\frac{(k_n)^{1/2}}{\sigma_F(x,y)\log n}\left[C_n(H_n)-C_n(D_n^+)\right] \to 0. \qquad a.co.$$

On the one hand, under some conditions allows us to use Theorem (3.3.6) and under (3.35) we can obtain the asymptotic normality of the first term. On the other hand, for the second term, by hypothesis (H2) and the fact that $1\!\!1_{\{D_n^- \le H_n \le D_n^+\}} \to 1$ where $\frac{k_n}{n} \to 0$ (see Burba et al. ([4])), we have :

$$C_n(D_n^+) \le C_n(H_n) \le C_n(D_n^-).$$

Using the fact that :

$$\mid C_n(H_n) - C_n(D_n^+) \mid \le \underbrace{\mid C_n(D_n^-) - \mathbb{E}\left[C_n(D_n^-)\right] \mid}_{M_1} + \underbrace{\mid C_n(D_n^+) - \mathbb{E}[C_n(D_n^+)] \mid}_{M_2}$$

$$+ \underbrace{\mid \mathbb{E}[C_n(D_n^-)] - \mathbb{E}[C_n(D_n^+)] \mid}_{M3}. \tag{3.37}$$

For the first term, we can write :

$$M_1 \le \mid C_n(D_n^-) - c \mid + \mid \mathbb{E}\left[C_n(D_n^-)\right] - c \mid$$

by Lemma (3.3.4), we have :

$$\mid C_n(D_n^-) - c \mid = O\left(\varphi_x^{-1}\left(\frac{k_n}{n}\right)^{\alpha}\right) + O(g_n^{\beta}) + O\left(\sqrt{\frac{\log n}{k_n}}\right)$$

and QUINTELA-DEL-RÍO ([21]) proved that :

$$\mid \mathbb{E}\left[C_n(D_n^-)\right] - c \mid = o(g_n^{\beta}) + O(\frac{1}{k_n}). \tag{3.38}$$

Using fact under hypothesis (H10), we obtain the almost complete convergence of the first term of M_1.
By same manner, we can establish the almost complete convergence of the second term M_2.
Finally for the third term M_3, we have :

$$M_3 \leq \mid \mathbb{E}[C_n(D_n^-)] - c \mid + \mid \mathbb{E}[C_n(D_n^+)] - c \mid$$

the almost complete convergence to 0 of these two terms is verifies in (3.38).

Proof of Lemma 3.3.6 :
To show this Lemma, we apply the same steps as precedent with :

$$\begin{cases} C_n(H_n) = \widehat{f}^x(y) \\ c = f^x(y). \end{cases} \tag{3.39}$$

Proof of Lemma 3.3.7 :
The result (3.11) of Lemma (3.3.1) allows us to conclude :

$$\widehat{F}^x(y) \rightarrow F^x(y) \;\; in \;\; probability.$$

Bibliographie

[1] Attouch, M., Benchikh, T. Asymptotic distribution of robust k. nearest neighbour estimator of functional nonparametric models. *Agence Nationala de Recherche Universitaire* 4 (2012), 275-285

[2] Attouch, M., Laksaci, A., Ouled said, E. Asymptotic distribution of robust estimator for functional nonparametric models. *Communication in statistics-theory and methods* 8 (2008), 1-19

[3] Burba, F., Ferraty, F., and Vieu, P. k-Nearest Neighbour method in functional nonparametric regression. *Journal of Nonparametric Statistics* 21, 4 (2009), 453-469.

[4] Burba, F., Ferraty, F., and Vieu, P. Convergence de l'estimateur à noyau des k plus proches voisins en régression fonctionnelle non-paramétrique. *Preprint submitted to Elsevier Science* (2008).

[5] Collomb, G. Estiamation de la regression par la methode des k plus proches avec noyau : quelques propriétés de convergence ponctuelle. In *Statistique Non Parametrique Asymptotique*, J-P. Raoult, Ed. vol.821 of *Lecture Notes in Mathematics.* Springer Berlin/Heidelberg, 1980, pp. 159-175.

[6] Collomb, G. Quelques propriétés de la methode du noyau pour l'estimation nonparamétrique de la régression en un point fixe, *C. R. Acad. Sci. Paris*, Série 1 285 (1997), pp. 289-293.

[7] Demongeot, J., Laksaci, A., Madani, F., Rachdi, M. Local linear estimation of the conditional density for functional data. C. R., *Math., Acad. Sci. Paris*, 348, Issues 15-16, pages 931-934, (2010).

[8] Demongeot, J., Laksaci, A., Madani, F., Rachdi, M. Functional data : local linear estimation of the density and its application. *Statistics*, DOI : 10.1080/02331888.2011.568117 (à paraitre en 2012).

[9] Demongeot, J., Laksaci, A., Madani, F., Rachdi, M. *A fast functional locally modeled conditional density and mode for functional time-series.* springer (2011).

[10] Laksaci, A., Mechab, B. Conditional hazard estimate for functional random fields.

[11] Ferraty, F., and Vieu, P. *Nonparametric Functional Data Analysis.* Springer, New York, (2006).

[12] Ferraty, F., Rabhi A., and Vieu, P. Estimation non parametrique de la fonction de hasard avec variable explicative fonctionnelle. *Roumaine Math*, 53 (2008), 1-18.

[13] Ferraty, F., Laksaci A., and Vieu, P. Estimating Some Characteristics Of The Conditional Distribution in Nonparametric Functional Models. *Statistical Inference for Stochastic Processes.*9 (2006), 47-76.

[14] Laksaci, A., Madani, F., Rachdi, M. Kernel conditional density estimation when the regressor is valued in a semi-metric space. *Communications Statistics-theory and methods*, 2012.

[15] Laksaci, A., Mechab, B. Nonparametric estimation of the hazard function given a functional variable : spatial data case. *Rev. Roumaine Math. Pures Appl 53.* 1-18 (2008).

[16] Laksaci, A., Mechab, B. Conditional hazard estimate for functional random fields.*Rev. Roumaine Math. Pures Appl 53.* (2008).

[17] Lian, H. Convergence on functional k-nearest neighbor regression estimate with functional responses. *Electronic Journal of Statistics* 5 (2011), 31-40

[18] Maillot, B., Louani, D. *Propriétés asymptotiques de quelques estimateurs non-paramétriques pour des variables vectorielles et fonctionnelles.* 2008. Paris.

[19] Muller, S., Dippon, J. k-NN Kernel Estimate for Nonparametric Functional Regression in Time Series Analysis. Fachbereich Mathematik, Fakultat Mathematik und Physik (Pfaffenwaldring 57). 014/2011.

[20] Oliveira, P. E. *Nonparametic density and regression estimate functional data.* Tech. rep., Departamento de Matématica, Univercidade de Coimbra, 2005.

[21] Quintela-del-rio, A. (2008). Hasard function given a fonctional variable : nonparametric estimation under strong mixing conditions. *J. Nonparametric stat.*. 413-430.

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

[23] Roussas, G. Nonparametric estimation of the transition distribution function of a Markov process. *Annals of Mathematical Statistics* , 40 1386-1400 (1969).

[24] Samanta, M. Non-parametric estimation of conditional quantiles. *Statist. Proba. Letters*, 7, 407-412 (1989).

[25] WASTON, G. S. and LEADBETTER, M. R. (1964). Hasard analysis. *I. Biometrika*, **51**, 175Ű-184.

Chapitre 4

Conclusion et Perspectives

4.1 Conclusion

Dans cette contribution, nous avons abordé une étude générale de l'estimation non paramétrique de la fonction de hasard conditionnelle par la méthode des k plus proches voisins (k-NN) quand la variable explicative est fonctionnelle. Deux cas ont été étudiés, le premier cas s'agit des variables indépendantes et identiquement distribuées tandis que le deuxième cas s'agit des variables α mélangeantes.

Pour le premier cas, nous avons construit l'estimateur de k plus proches voisins de la fonction de hasard conditionnelle. Ensuite, nous avons étudié ses quelques propriétés asymptotiques. Il est claire que les propriétés asymptotiques de l'estimateur de la fonction de hasard conditionnelle s'obtiennent relativement facilement à partir, de la littérature connu en matière d'estimation, de fonction de répartition et de densité conditionnelles. A ce sujet, nous avons prouvé la convergence presque complète avec taux de convergence de nos estimateurs et nous avons établit la normalité asymptotique.

Pour le deuxième cas, nous avons considéré les mêmes estimateurs de k plus proches voisins définit dans le premier cas et nous avons opté pour une présentation permettant de trouver les mêmes propriétés asymptotiques que le cas indépendants en mettant des conditions supplémentaires sur la dépendance.

De plus, notre méthode et les estimateurs qu'on a donné sont utilisables pour d'autres modèles non paramétrique à des deux fonctions de densité conditionnelle et de la répartition conditionnelle. Ainsi nous pouvons dire que notre contribution ouvre des perspectives sur de nombreuses questions de recherches. Dans la section suivante, nous citons quelques exemples de perspectives proposés.

4.2 Perspectives

Pour conclure les travaux de notre thèse et pour améliorer et étendre nos résultats, nous listons ci dessous quelques futurs développements possibles.

- **Estimation de la fonction de hasard conditionnelle par la méthode des k plus proches voisins avec des variables censurées :**
  Dans la pratique, lors d'applications médicales en particulier, on peut se trouver en presence de variables censurées, c'est pour ça l'estimation de la fonction de hasard conditionnelle est très importante. Dans le futur proches, nous essaierons de développer les résultats de Ferraty et Rabhi et Vieu (2008).
- **Le cas spacial :**
  La généralisation de nos résultats au cadre spacial soulève de nombreuses questions. Les propriétés asymptotiques des estimateurs introduit peuvent être établies pour des modes de convergences.
- **Le choix optimal du paramètre k :**
  L'etude du choix optimal du paramètre k n'a pas été conduite, mais elle constitue une perspective de recherche à court terme.
  D'autres thèmes peuvent être abordé à long terme tels le conditionnement par p variables fonctionnelles par la méthode des k plus proches voisins, le choix de la semi métrique,...

Bibliographie générale

ATTOUCH, M. ; BENCHIKH, T. (2012). Asymptotic distribution of robust k-nearest neighbour estimator for functional nonparametric models. *Mathematic Vesnic*, **64**, No. 4, pp. 275-285.

ATTOUCH, M., LAKSACI, A., OULED SAID, E. (2008). Asymptotic distribution of robust estimator for functional nonparametric models. *Communication in statistics-theory and methods* 8, 1-19

BURBA, F. ; FERRATY, F. and VIEU, P.(2009). k-Nearest Neighbour method in functional nonparametric regression. *Journal of Nonparametric Statistics* **21**, 4 , 453–469.

BURBA, F. ; FERRATY, F. and VIEU, P. (2008). Convergence de l'estimateur à noyau des k plus proches voisins en régression fonctionnelle nonparamétrique. *Preprint submitted to Elsevier Science.*

COLLOMB, G. (1980). Estiamation de la regression par la methode des k plus proches avec noyau : quelques propriétés de convergence ponctuelle. *Statistique Non Parametrique Asymptotique*, J-P. Raoult, Ed. vol.821 of *Lecture Notes in Mathematics.* Springer Berlin/Heidelberg, pp. 159–175.

COLLOMB, G. (1997). Quelques propriétés de la methode du noyau pour l'estimation nonparamétrique de la régression en un point fixe, *C. R. Acad. Sci. Paris*, Série 1 285, 289–293.

DEMONGEOT, J., LAKSACI, A., MADANI, F., RACHDI, M.(2010). Local linear estimation of the conditional density for functional data. C. R., *Math., Acad. Sci. Paris*, 348, Issues 15-16, pages 931-934.

Demongeot, J., Laksaci, A., Madani, F., Rachdi, M.(à paraitre en 2012). Functional data : local linear estimation of the density and its application. *Statistics*, DOI : 10.1080/02331888.2011.568117.

Demongeot, J., Laksaci, A., Madani, F., Rachdi, M.(2011). *A fast functional locally modeled conditional density and mode for functional time-series.* springer.

Devroye, L.P. *The uniform convergence of nearest neighbour regression function estimators and their application in optimization*, IEEE Trans. Inform. Theory, **24**, 142-151, 1978.

Ferraty, F. ; Laksaci A. ; Tadj, A. and Vieu, P.(2011). Kernel regression with functional response. *Electronic Journal of Statistics.* **5**, 159-171.

Ferraty, F. ; Laksaci A. and Vieu, P.(2006). Estimating Some Characteristics Of The Conditional Distribution in Nonparametric Functional Models. *Statistical Inference for Stochastic Processes.* **9**, 47–76.

Ferraty, F. ; Rabhi A. and Vieu, P. (2008). Estimation non parametrique de la fonction de hasard avec variable explicative fonctionnelle. *Roumaine Math*, **53**, 1–18.

Ferraty, F. and Vieu, P. (2006). *Nonparametric Functional Data Analysis.* Springer, New York.

Laksaci, A. ; Madani, F. and Rachdi, M. (2012). Kernel conditional density estimation when the regressor is valued in a semi-metric space. *Communications Statistics-theory and methods* Accepted paper.

Laksaci, A. and Mechab, B. (2012). Conditional hazard estimate for functional random fiels. *Communications Statistics-theory and methods* Accepted paper.

Laksaci, A. and Mechab, B.(2010). Estimation nonparamétrique de la fonction de hasard avec variable explicative fonctionnelle : cas des données spatiales. *Rev. Roumaine Math. Pures Appl.* **55**, 35-51 .

LIAN, H. (2011). Convergence on functional k-nearest neighbor regression estimate with functional responses. *Electronic Journal of Statistics* **5**, 31–40.

LI, J. AND TRAN, L.T. (2007). Hazard rate estimation on random fields. *J. Multivariate Anal.* 98, 1337Ű1355.

MAILLOT, B. and LOUANI, D.(2008). *Propriétés asymptotiques de quelques estimateurs non-paramétriques pour des variables vectorielles et fonctionnelles.* Paris.

MACK, Y. P. (1981). Local properties of k-NN regression estimates. *SIAM J. Algebr Discrete Methods*, **2** , 311-323.

MULLER, S. and DIPPON, J. k-NN Kernel Estimate for Nonparametric Functional Regression in Time Series Analysis. Fachbereich Mathematik, Fakultat Mathematik und Physik (Pfaffenwaldring 57). 014/2011.

OLIVEIRA, P. E.(2005). *Nonparametic density and regression estimate functional data.* Tech. rep., Departamento de Matématica, Univercidade de Coimbra.

QUINTELA-DEL-RIO (2008). Hasard function given a fonctional variable : nonparametric estimation under strong mixing conditions. *J. Nonparametric stat.*. 413-430.

RAMSAY, J. and SILVERMAN, B. (2005) *Functional Data Analysis*, 2nd Ed. Springer Series in Statistics. Springer, New York.

ROUSSAS, G. (1969). Nonparametric estimation of the transition distribution function of a Markov process. *Annals of Mathematical Statistics*, 40 1386–1400.

ROUSSAS, G. (1989). Nonparametric estimation in mixing sequences of random variables. *J. Statist. Plann.*, 18 135–149.

R.M. ROYALL (1966). *A class of nonparametric estimates of a smooth* *regression function*, Ph.D. Diss., Stanford University.

Samanta, M. (1989). Non-parametric estimation of conditional quantiles. *Statist. Proba. Letters*,**7**, 407–412.

Samanta, M. and Thavaneswaran, A. (1990) Non-parametric estimation of conditional model. *Comm. Statist. Theory and Meth.*, **16** , 4515-4524.

Stone, C. J. (1977). Consistent nonparametric regression. *Ann. Statist*, **5**, 595Ű-645.

Waston, G. S. and Leadbetter, M. R. (1964). Hasard analysis. *I. Biometrika*, **51**, 175Ű-184.