Available online at www.sciencedirect.com
SCIENCE DIRECT
C. R. Acad. Sci. Paris, Ser. I 336 (2003) 75–80
COMPTES RENDUS MATHEMATIQUE

Statistique/Probalilités

Estimation non paramétrique de la régression avec variable explicative dans un espace métrique

Kernel regression estimation when the regressor takes values in metric space

Sophie Dabo-Niang [a], Noureddine Rhomari [b]

[a] *Laboratoire de statistique, CREST-INSEE, 3, avenue Pierre Larousse, 92245 Malakoff cedex, France*
[b] *Université Mohamed I, Faculté des sciences, 60 000 Oujda, Maroc*

Reçu le 6 janvier 2002 ; accepté après révision le 13 novembre 2002

Présenté par Paul Deheuvels

Résumé

Nous étudions l'estimateur à noyau de la régression quand la variable explicative prend ses valeurs dans un espace semi-métrique. Nous établissons sa consistance en moyenne d'ordre p et presque sûre et nous donnons des bornes supérieures de ces erreurs d'estimation sous des conditions générales. Nous appliquons ces résultats à la discrimination de variables d'un espace semi-métrique et les illustrons par l'exemple du processus de Wiener comme variable explicative. ***Pour citer cet article : S. Dabo-Niang, N. Rhomari, C. R. Acad. Sci. Paris, Ser. I 336 (2003).***

Abstract

We study a nonparametric regression estimator when the explanatory variable takes its values in a semi-metric space. We establish some asymptotic results and give upper bounds of the p-mean and the almost sure estimation errors under general conditions. We end by an application to the discrimination in a semi-metric space and illustrate the results by the example of Wiener process as an explanatory variable. ***To cite this article: S. Dabo-Niang, N. Rhomari, C. R. Acad. Sci. Paris, Ser. I 336 (2003).***

Abridged English version

Let (X, Y) be a random vector from $\mathcal{X} \times \mathbb{R}$ with $E(|Y|) < \infty$, where $(\mathcal{X}, d)$ is a separable semi-metric space equipped with the semi-metric d. The distribution of (X, Y) is often unknown, so is the regression function

Adresses e-mail : niang@ensae.fr (S. Dabo-Niang), rhomari@ensae.fr, rhomari@sciences.univ-oujda.ac.ma (N. Rhomari).

1631-073X/03/$ – see front matter
doi:10.1016/S1631-073X(02)00012-2

$m(x) = E(Y|X = x)$. The aim of this work is to build a consistent estimator m_n of m, based on the sample $(X_1, Y_1), \ldots, (X_n, Y_n)$ of (X, Y). This framework includes the classical regression estimation when $\mathcal{X} = \mathbb{R}^l$, $l \geqslant 1$, which was widely studied by many authors, but also the case of the spaces $\mathcal{X}$ of infinite dimensions. This last case has recently met a growing interest in the modelization of certain phenomena (see Ramsay and Silverman [12] and Bosq [1]). $\mathcal{X}$ is a general semi-metric space and may be any subset of usual function spaces, or subset of probabilities on some given measurable set, etc.

The estimate we consider here is of kernel-type defined, for $x \in \mathcal{X}$, by

$$m_n(x) = \sum_{i=1}^{n} \frac{K(d(X_i, x)/h_n)}{\sum_{j=1}^{n} K(d(X_j, x)/h_n)} Y_i, \quad \text{if } \sum_{j=1}^{n} K\left(\frac{d(X_j, x)}{h_n}\right) \neq 0,$$

else $m_n(x) = 0$; $(h_n)_n$ is a sequence of positive numbers and K a positive measurable function on $\mathbb{R}$.

To our knowledge, there is little literature dealing with nonparametric regression estimation when the explanatory variable is of infinite dimension. Kulkarni and Posner [11] studied the nearest neighbor estimation in a separable metric space $\mathcal{X}$. Ferraty and Vieu [9] was on the kernel method when X is in a semi-normed vector space and has a distribution with a finite fractal dimension. Ferraty et al. [8] extended this last work to dependent observations and obtained the a.s. uniform convergence on compact set.

In this Note we prove the p-mean and (pointwise and integrated) almost sure consistency of m_n under general conditions; the regression operator m may be discontinuous. We also precise the upper bound of each estimation error. We then apply these results to nonparametric discrimination of variables in a semi-metric space (e.g., the curves classification). As an example we present a special case when X is the standard Wiener process and $\mathcal{X} = C[0, 1]$ equipped with the sup-norm. This example illustrates the extension and the contribution of this work; we can apply it to nonparametric regression estimation with an explanatory variable whose distribution has an infinite fractal dimension, and to general curves classification. Let μ denote the distribution of X and B_h^x the closed ball of radius h and center $x \in \mathcal{X}$; see the French version for the other notations and assumptions. Then the main results are,

Theorem 0.1. (i) *If $E(|Y|^p) < \infty$, and if* (3), (4) *and* (6) *are fulfilled then, for all x in $\mathcal{X}$ such that m verifies* (5), *we have $E(|m_n(x) - m(x)|^p) \underset{n\to\infty}{\longrightarrow} 0$.*

(ii) *If in addition $E(|Y - m(X)|^p | X \in B_{rh_n}^x) = \mathrm{O}(1)$ and* (7) *is satisfied, with $p \geqslant 2$, then*

$$E\left(\left|m_n(x) - m(x)\right|^p\right) = \mathrm{O}\left(h_n^{p\tau} + \left(\frac{1}{n\mu(B_{rh_n}^x)}\right)^{p/2}\right).$$

Theorem 0.2. *If Y is bounded,* (3) *is fulfilled, $h_n \to 0$ and $\lim_{n\to\infty} n\mu(B_{rh_n}^x)/\log n = \infty$, then, for all x in $\mathcal{X}$ such that m satisfies* (5), *we have $|m_n(x) - m(x)| \to 0$, a.s., when $n \to \infty$.*

If in addition (5) *is fulfilled μ-almost all x, then $\lim_{n\to\infty} E(|m_n(X) - m(X)||X_1, Y_1, \ldots, X_n, Y_n) = 0$, a.s.*

Corollary 0.3. *If $|Y| \leqslant M$,* (3) *and* (7) *hold, $h_n \to 0$ and $\lim_{n\to\infty} n\mu(B_{rh_n}^x)/\log n = \infty$, then*

$$\left|m_n(x) - m(x)\right| = \mathrm{O}\left(h_n^{\tau} + \left(\frac{\log n}{n\mu(B_{rh_n}^x)}\right)^{1/2}\right), \quad \textit{a.s.}$$

For the discrimination see Theorem 5.1 bellow. In Theorem 0.2 and Corollary 0.3, we can relax the boundedness of Y by assuming the Cramer condition, and in this case the $\log n$ is replaced by $(\log n)^3$. But if we just assume a polynomial moment of Y this $\log n$ becomes a power of n. We remark that if $\mathcal{X} = \mathbb{R}^l$, then [6] $h^l = \mathrm{O}(\mu(B_h^x))$ for μ-almost all x, and all μ; therefore we find the well known optimal rate of convergence. However, in the general case the rate can be slow. Let $\mathcal{X} = C[0, 1]$ equipped with the sup-norm, and X the standard Wiener process. Then for $h_n = (c\log n)^{-1/2}$, $c < 8/\pi^2$, we obtain $E(|m_n(x) - m(x)|^p) = \mathrm{O}((\log n)^{-\tau p/2})$, $p \geqslant 2$, for $x \in C[0,1]$ s.t. $x(0) = 0$, x absolutely continuous and $\int_0^1 x'^2(t)\,\mathrm{d}t < \infty$. Note that the standard Wiener process in $C[0,1]$ does not fulfill the conditions of [9] and [8]. Our conditions and results unify both cases of finite and infinite dimensional regressors.

1. Introduction

Soit (X,Y) un vecteur aléatoire (v.a.) de $\mathcal{X} \times \mathbb{R}$ avec $E(|Y|) < \infty$, $(\mathcal{X}, d)$ un espace semi-métrique séparable muni de la semi-métrique d. Le but de ce travail est de construire un estimateur, $m_n(x)$, consistant de la fonction de régression $m(x) = E(Y|X = x)$, $x \in \mathcal{X}$, basé sur un n-échantillon de (X,Y). Ce cadre inclut le cas où $\mathcal{X} = \mathbb{R}^l$, $l \geqslant 1$, qui a été largement étudié dans la littérature, mais aussi des espaces $\mathcal{X}$ de dimension infinie, éventuellement des espaces fonctionnels. Ce dernier cas a récemment connu un intérêt croissant dans la modélisation de certain phénomènes (voir Ramsay et Silverman [12] et Bosq [1]). $\mathcal{X}$ peut aussi être l'ensemble des probabilités sur un ensemble donné, ...

Etant donné $(X_1,Y_1), \ldots, (X_n,Y_n)$ des observations i.i.d. de (X,Y), un estimateur populaire de m est

$$m_n(x) = \sum_{i=1}^n W_{ni}(x) Y_i, \quad x \in \mathcal{X}, \tag{1}$$

où $(W_{n1}(x), \ldots, W_{nn}(x))$ est un vecteur de poids dont chacune des composantes $W_{ni}(x)$ est une fonction mesurable de $x, X_1, \ldots, X_n$. On s'intéresse, ici, à la méthode du noyau donnée par les poids

$$W_{ni}(x) = \frac{K(d(X_i,x)/h_n)}{\sum_{j=1}^n K(d(X_j,x)/h_n)}, \quad \text{si } \sum_{j=1}^n K\left(\frac{d(X_j,x)}{h_n}\right) \neq 0, \tag{2}$$

sinon $W_{ni}(x) = 0$; $(h_n)_n$ est une suite de réels positifs et K une fonction mesurable positive sur $\mathbb{R}$.

La littérature sur ce sujet est relativement restreinte lorsque $\mathcal{X}$ est de dimension infinie. Kulkarni et Posner [11] ont considéré l'estimateur des voisins les plus proches dans un espace métrique séparable $\mathcal{X}$. La méthode du noyau a été étudiée par Ferraty et Vieu [9] quand la loi de X admet une dimension fractale finie dans un espace vectoriel semi-normé $\mathcal{X}$. Récemment, sous cette même condition, Ferraty et al. [8] ont traité le cas dépendant et ont établi des convergences p.s. uniformes sur des compacts. Pour l'estimation à noyau de la densité en dimension infinie voir [3].

Dans cette Note nous montrons la consistance en moyenne d'ordre p et presque sûre (ponctuelle et intégrée) de l'estimateur (1), (2) sous des conditions générales (m peut être discontinue). Nous précisons aussi des bornes supérieures de ces erreurs d'estimation. Puis nous appliquons ces résultats à la discrimination de variables d'un espace semi-métrique (e.g. classification de courbes non paramétriques) et étudions comme exemple le cas où X est le processus de Wiener dans $C[0,1]$. Cette dérnière application illustre bien l'extension et l'apport de ce travail ; la mesure de Wiener, sur $C[0,1]$, ne vérifie pas les conditions de [9] et [8]. Les bornes que nous obtenons sont optimales en dimension finie. Ce travail étend, notamment, les résultats de [6] aux variables explicatives d'un espace semi-métrique, et généralise [9] ; lorsque la loi de X a une dimension fractale finie on retrouve les résultats de [9]. Nos conditions et résultats unifient les deux cadres des variables explicatives de dimensions finie et infinie.

Dans le Paragraphe 2 nous donnons les hypothèses, puis nous étudions la convergence en moyenne d'ordre p de l'estimateur au Paragraphe 3. Le Paragraphe 4 est consacré à l'étude de la convergence ponctuelle et intégrée presque sûre. Quant au dernier paragraphe, il est consacré à l'application à la discrimination et l'exemple du processus de Wiener. Puis on termine par quelques commentaires.

2. Hypothèses et commentaires

Nous supposons qu'il existe des nombres positives r, a et b tels que

$$a\mathbb{1}_{\{|u| \leqslant r\}} \leqslant K(u) \leqslant b\mathbb{1}_{\{|u| \leqslant r\}}. \tag{3}$$

Notons μ la loi de X et B_h^x la boule fermée de rayon h et de centre x. Soit x dans $\mathcal{X}$ et $p \geqslant 1$, on suppose

$$h_n \to 0 \quad \text{et} \quad \lim_{n\to\infty} n\mu\big(B_{rh_n}^x\big) = \infty, \tag{4}$$

$$\lim_{h\to 0} \frac{1}{\mu(B_h^x)} \int_{B_h^x} \big|m(w) - m(x)\big|^p \,\mathrm{d}\mu(w) = 0, \tag{5}$$

$$E\big[\big|Y - m(X)\big|^p | X \in B_{rh_n}^x\big] = \mathrm{o}\big(\big[n\mu\big(B_{rh_n}^x\big)\big]^{p/2}\big). \tag{6}$$

Remarque 1. (1) (5) est satisfaite pour tout point de continuité de m et tout μ, mais n'implique pas la continuité de m. La fonction $m(x) = \mathbb{1}_{\{x \text{ rationnel dans } [0,1]\}}$, satisfait (5), mais n'est continue en aucun point, (exemple de [10]) ; μ étant la mesure de Lebesgue sur $[0, 1]$.

(2) Si m vérifie (5) pour $p = 1$ et est bornée dans un voisinage de x alors (5) est vraie pour tout p.

(3) Si Y est bornée la condition (6) est trivialement satisfaite. Elle l'est aussi, par exemple, quand (4) est vérifiée et si $E[|Y - m(X)|^p|X \in B_{rh_n}^x] = \mathrm{O}(1)$ ou $\lim_{n\to\infty} n(\mu(B_{rh_n}^x))^{1+2/p} = \infty$.

(4) Si $\mathcal{X}$ est un espace de Hilbert et μ une mesure Gaussienne d'opérateur de covariance $Rx = \sum_i c_i\langle x, e_i\rangle e_i$; (e_i) étant un système orthonormal dans $\mathcal{X}$. Si $c_{i+1} \leqslant c_i i^{-\alpha}$ pour $\alpha > 5/2$, alors (5) et (6) sont vérifiées pour μ-presque tout x dans $\mathcal{X}$ dès que $E|Y|^{p'} < \infty$ pour un $p' > p \geqslant 1$ (voir [14]).

Notons que (5) et (6) sont satisfaites pour μ-presque tout x et tout μ quand $\mathcal{X}$ est de dimension finie [6].

3. Convergence en moyenne d'ordre p

Théorème 3.1. *Si $E(|Y|^p) < \infty$, et si* (3), (4) *et* (6) *sont satisfaites alors, pour tout x dans $\mathcal{X}$ tel que m vérifie* (5), *on a* $E(|m_n(x) - m(x)|^p) \underset{n\to\infty}{\longrightarrow} 0$.

Si μ est discrète, ce résultat a lieu en tout point du support de μ dès que $E(|Y|^p) < \infty$ et $h_n \to 0$.

Pour préciser une borne de l'erreur d'estimation, on a besoin de renforcer l'hypothèse (5) à fin d'évaluer le biais. On suppose que m est « lipschitzienne en moyenne d'ordre p », de paramètres $0 < \tau = \tau_x \leqslant 1$ et $c_x > 0$, dans un voisinage de x :

$$\frac{1}{\mu(B_h^x)} \int_{B_h^x} \big|m(w) - m(x)\big|^p \,\mathrm{d}\mu(w) \leqslant c_x h^{p\tau}, \quad \text{quand } h \to 0. \tag{7}$$

Il est clair que (7) est satisfaite si m est lipschitzienne dans un voisinage de x ; mais elle est plus faible que la condition de Lipschitz et n'implique, en général, même pas la continuité de m en x (voir la Remarque 1). Une condition similaire a été utilisée par [10] en dimension finie.

Corollaire 3.2. *Si $E(|Y|^p) < \infty$, pour un $p \geqslant 2$, $E(|Y - m(X)|^p|X \in B_{rh_n}^x) = \mathrm{O}(1)$ et si* (3), (4), (7) *sont satisfaites, alors on a*

$$E\big(\big|m_n(x) - m(x)\big|^p\big) = \mathrm{O}\left(h_n^{p\tau} + \left(\frac{1}{n\mu(B_{rh_n}^x)}\right)^{p/2}\right).$$

Sans l'hypothèse d'existence des densités, ce corollaire en dimension finie semble nouveau pour Y non bornée.

Remarque 2. (1) S'il existe $a(x) > 0$ tel que $\limsup_{h\to 0} \mu(B_h^x)/h^{a(x)} < \infty$ alors pour $h_n = cn^{-1/(2\tau + a(x))}$ on a $E(|m_n(x) - m(x)|^p) = \mathrm{O}(n^{-p\tau/(2\tau + a(x))})$.

(2) Si $\mathcal{X} = \mathbb{R}^l$, la meilleure vitesse obtenue dans le Corollaire 3.2 est optimale, car $h^l = \mathrm{O}(\mu(B_h^x))$ pour μ-presque tout x, et tout μ (voir [6]). La vitesse est similaire à celle en (1) avec l au lieu de $a(x)$.

4. Convergence presque sûre

Théorème 4.1. *Si Y est bornée,* (3) *est vérifiée, $h_n \to 0$ et $\lim_{n\to\infty} n\mu(B^x_{rh_n})/\log n = \infty$, alors pour tout x dans $\mathcal{X}$ tel que m vérifie* (5) *on a*

$$\left|m_n(x) - m(x)\right| \to 0, \quad \text{p.s., quand } n \to \infty. \tag{8}$$

Si de plus (5) *est satisfaite pour μ-presque tout x, alors*

$$\lim_{n\to\infty} E\left(\left|m_n(X) - m(X)\right| | X_1, Y_1, \ldots, X_n, Y_n\right) = 0, \quad \text{p.s.} \tag{9}$$

Corollaire 4.2. *Si $|Y| \leqslant M$,* (3) *et* (7) *sont vérifiées, $h_n \to 0$ et $\lim_{n\to\infty} n\mu(B^x_{rh_n})/\log n = \infty$, alors*

$$\left|m_n(x) - m(x)\right| = \mathrm{O}\left(h_n^\tau + \left(\frac{\log n}{n\mu(B^x_{rh_n})}\right)^{1/2}\right), \quad \text{p.s.}$$

En fait on obtient, pour n assez grand, $|m_n(x) - m(x)| \leqslant c_x(b/a)(rh_n)^\tau + 6M(b/a)(\log n/(n\mu(B^x_{rh_n})))^{1/2}$, p.s.

Remarque 3. Cette dernière borne est simple et facile à estimer. Nous suggérons de choisir la fenêtre h qui minimise cette borne. D'autre méthodes, comme la validation croisée, peuvent aussi être envisagées.

Remarque 4. (1) [9] a obtenu (8) sous $\lim_{h\to 0} \mu(B^x_h)/h^{a(x)} = c(x) > 0$ et m continue en x. Cette convergence, aussi bien ponctuelle qu'uniforme sur des compacts, a été étendue par [8] aux processus α-mélangeants.

(2) Si $\limsup_{h\to 0} \frac{\mu(B^x_h)}{h^{a(x)}} < \infty$ et $h = c(\log n/n)^{1/(2\tau + a(x))}$, alors $|m_n(x) - m(x)| = \mathrm{O}((\log n/n)^{\tau/(2\tau + a(x))})$, p.s. [9] obtient $\mathrm{O}((\log n/n)^{\gamma(x)/(2\gamma(x) + a(x))})$, quand m est lipschitzienne et s'il existe $a(x), b(x), c(x) > 0$ tels que $\mu(B^x_h) = h^{a(x)}c(x) + \mathrm{O}(h^{a(x)+b(x)})$; $\gamma(x) = \min\{b(x), \tau\}$. [8] a montré que ces résultats ont lieu uniformément sur des compacts sous le mélange fort.

5. Application à la discrimination

En discrimination, le problème est de classer un objet ayant une variable observée X appartenant à $\mathcal{X}$, et une classe inconnue Y dans $\{1, \ldots, M\}$, i.e. décider sur la valeur de Y à partir de X et d'un échantillon de cette paire de variables. X peut être une probabilité, une courbe (e.g. courbe de l'électrocardiogramme d'un malade, et Y un risque). Une régle de décision empirique, g_n, est une fonction mesurable de X et de l'échantillon dans $\{1, \ldots, M\}$; sa probabilité d'erreur est $L_n = P\{g_n(X) \neq Y | X_1, Y_1, \ldots, X_n, Y_n\}$. Soit $L^* = \inf_{g:\mathcal{X}\to\{1,\ldots,M\}} P\{g(X) \neq Y\}$, la probabilité d'erreur de Bayes ; alors on sait que $L_n \geqslant L^*$. Soit g_n définie par, [6], $\sum_{i=1}^n W_{ni}(x)\mathbb{1}_{\{Y_i = g_n(x)\}} = \max_{1\leqslant j\leqslant M} \sum_{i=1}^n W_{ni}(x)\mathbb{1}_{\{Y_i = j\}}$. Un estimateur similaire a été utilisé par [7] pour discriminer des courbes. La consistance du classificateur g_n est assurée par

Théorème 5.1. *Si* (3), (4) *et* (5) *sont vérifiées μ-presque tout x alors $\lim_{n\to\infty} L_n = L^*$ en probabilité.*

Si de plus, $\lim_{n\to\infty} n\mu(B^x_{rh_n})/\log n = \infty$, μ-presque tout x, alors $\lim_{n\to\infty} L_n = L^*$, presque sûrement.

Example. Soit $\mathcal{X} = C[0, 1]$, l'espace des fonctions continues sur $[0, 1]$, muni de la norme uniforme, et X le processus de Wiener standard. Soit $S = \{x \in C[0, 1]\colon\ x(0) = 0,\ x$ absolument continue et $\int_0^1 x'^2(t)\,\mathrm{d}t < \infty\}$. Soit $h = (c\log n)^{-1/2}$, $c < 8/\pi^2$, alors, pour $x \in S$, [2], il existe $C(x) > 0$, tel que $C(x)n^{-c\pi^2/8} \leqslant \mu(B^x_h) \leqslant n^{-c\pi^2/8}$. Donc $E(|m_n(x) - m(x)|^p) = \mathrm{O}((\log n)^{-\tau p/2})$, $p \geqslant 2$.

Notons que $\forall x \in S$ et $\forall a > 0$ on a $\lim_{h\to 0} \mu(B^x_h)/h^a = 0$; la mesure de Wiener sur $C[0, 1]$ a une dimension fractale infinie et ne vérifie donc pas la condition de [9] et [8] (cf. Remarque 4).

Remarque 5. (1) Soit $\mathcal{X}'$ un espace semi-métrique séparable, $\Phi:\mathbb{R}\to\mathbb{R}$ et $\Psi:\mathcal{X}\to\mathcal{X}'$ des fonctions mesurables. Il est clair que les résultats obtenus peuvent être utilisés pour estimer $m_\Phi^\Psi(z)=E(\Phi(Y)|\Psi(X)=z)$. L'introduction de ces fonctions permet de couvrir et d'estimer différents paramètres (aussi fonctionnels) associés à la distribution de (X,Y).

(2) Si Y n'est pas bornée mais admet un moment exponentiel $E(\mathrm{e}^{\theta\|Y\|})<\infty$, pour un $\theta>0$, alors (8) a lieu dès que $h_n\to 0$, $\lim_{n\to\infty} n\mu(B^x_{rh_n})/\log^3 n=\infty$ et m vérifie (5). On peut se contenter juste de l'existence d'un moment polynomial, mais le $\log n$ dans la condition devient une puissance de n.

(3) Nous pensons que les conditions $h_n\to 0$ et $\lim_{n\to\infty} n\mu(B^x_{rh_n})=\infty$ (resp. $\lim_{n\to\infty} n\mu(B^x_{rh_n})/\log n=\infty$) sont nécessaires dans les Théorèmes 3.1 (resp. 4.1).

(4) Si $K(u)=\mathbb{1}_{\{|u|\leqslant r\}}$ et si (4) est satisfaite alors la condition (5) est nécessaire.

(5) Les preuves que nous avons utilisées nous ont permis de ne pas supposer une « structure particulière » de la loi de X, comme dans [9] et [8] ; ce qui a permis d'inclure par exemple le processus de Wiener.

(6) $(\mathcal{X},d)$ peut être par exemple tout sous-ensemble d'espaces fonctionnels usuels (e.g. ensemble des densités de probabiltés, tout sous-espace de fonctions qui n'a pas une structure d'espace vectoriel, ...), ou :

(a) $(\mathcal{P},d_{\mathrm{Pro}})$, ensemble de probabilités sur un certain espace mesurable et d_{Pro} la distance de Prohorov.
(b) $(\mathcal{P}_{\mathbb{R}},d_K)$, ensemble de probabilités sur $\mathbb{R}$ et $d_K(P,Q)=\sup_x|P(-\infty,x]-Q(-\infty,x]|$.
(c) $(\mathrm{D}[0,1],d_{\mathrm{Sko}})$ ensemble de fonctions cadlag sur $[0,1]$ et d_{Sko} la distance de Skorohod.

(7) Les mêmes résultats restent valables lorsque Y est à valeurs dans un espace de Banach séparable, [5].

(8) On peut supprimer l'hypothèse d'indépendance et la remplacer par le β-mélange, [13].

Les détails et preuves sont dans [4], et sont, avec de légères modifications, des adaptations de celles de [6].

Remerciements

Nous remercions un des référés de nous avoir mis au courant des travaux [7] et [8].

Références

[1] D. Bosq, Linear Processes in Function Spaces: Theory and Applications, in: Lecture Notes in Statist., Vol. 149, Springer, 2000.
[2] E. Csaki, A relation between Chung's and Strassen laws of the iterated logarithm, Z. Wahrscheinlichkeitstheorie Verw. Gebiete 19 (1980) 287–301.
[3] S. Dabo-Niang, Estimation de la densité dans un espace de dimension infinie : Application aux diffusions, C. R. Acad. Sci. Paris Sér. I Math. 334 (2002) 213–216.
[4] S. Dabo-Niang, N. Rhomari, Nonparametric regression estimation when the regressor takes values in a metric space, Technical Report 2002-9, LSTA Univ. Paris 6, 2001, http://www.ccr.jussieu.fr/lsta/R2002_9.pdf.
[5] S. Dabo-Niang, N. Rhomari, Kernel regression estimation in a Banach space, Preprint, 2001.
[6] L. Devroye, On the absolute everywhere convergence of nonparametric regression function estimates, Ann. Statist. 9 (1981) 1310–1319.
[7] F. Ferraty, Estimation nonparamétrique et discrimination de courbes, Actes SFC 2001, 17–21 Décembre 2001, pp. 128–132.
[8] F. Ferraty, A. Goia, P. Vieu, Functional nonparametric model for time series: a fractal approach for dimension reduction, Test, 2002, à paraître.
[9] F. Ferraty, P. Vieu, Dimension fractal et estimation de la régression dans des espaces vectoriels semi-normés, C. R. Acad. Sci. Paris Sér. I Math. 330 (2000) 139–142.
[10] A. Krzyżak, The rates of onvergence of kernel regression estimates and clasification rules, IEEE Trans. Inform. Theory IT-32 (1986) 668–679.
[11] S.R. Kulkarni, S.E. Posner, Rates of onvergence of nearest neighbor estimation under arbitrary sampling, IEEE Trans. Inform. Theory 41 (1995) 1028–1039.
[12] J.O Ramsay, B.W. Silverman, Functional Data Analysis, Springer, New York, 1997.
[13] N. Rhomari, Kernel regression estimation in Banach space under dependence, Preprint, 2001.
[14] J. Tišer, Differentiation theorem for Gaussian measures on Hilbert spaces, Trans. Amer. Math. Soc. 308 (1988) 655–666.