



ELSEVIER

Contents lists available at ScienceDirect

C. R. Acad. Sci. Paris, Ser. I

www.sciencedirect.com



Statistics

## Large deviations theorems in nonparametric regression on functional data

*Théorèmes de grandes déviations pour la régression non paramétrique sur des données fonctionnelles*

Mohamed Cherfi

Laboratoire de statistique théorique et appliquée (LSTA), équipe d'Accueil 3124, université Pierre et Marie Curie – Paris 6, tour 15-25, 2ème étage, 4, place Jussieu, 75252 Paris cedex 05, France

## ARTICLE INFO

## Article history:

Received 19 May 2010

Accepted after revision 19 April 2011

Available online 12 May 2011

Presented by Paul Deheuvels

## ABSTRACT

In this Note we prove large deviations principles for the Nadaraya–Watson estimator of the regression of a real-valued variable with a functional covariate. Under suitable conditions, we show pointwise and uniform large deviations theorems.

© 2011 Published by Elsevier Masson SAS on behalf of Académie des sciences.

## R É S U M É

L'objet de cette Note est d'établir un principe de grandes déviations ponctuel et un principe de grandes déviations uniforme pour l'estimateur à noyau de la régression sur des données fonctionnelles.

© 2011 Published by Elsevier Masson SAS on behalf of Académie des sciences.

## 1. Introduction

Let  $\{(Y_i, X_i), i \geq 1\}$  be a sequence of independent and identically distributed random vectors. The random variables  $Y_i$  are real, with  $\mathbb{E}|Y| < \infty$ , and the  $X_i$  are random vectors with values in a semi-metric space  $(\mathcal{X}, d(\cdot, \cdot))$ .

Consider now the functional regression model,

$$Y_i := \mathbb{E}(Y|X_i) + \varepsilon_i = r(X_i) + \varepsilon_i \quad i = 1, \dots, n, \quad (1)$$

where  $r$  is the regression operator mapping  $\mathcal{X}$  onto  $\mathbb{R}$ , and the  $\varepsilon_i$  are real variables such that, for all  $i$ ,  $\mathbb{E}(\varepsilon_i|X_i) = 0$  and  $\mathbb{E}(\varepsilon_i^2|X_i) = \sigma_\varepsilon^2(X_i) < \infty$ . Note that in practice  $\mathcal{X}$  is a normed space which can be of infinite dimension (e.g., Hilbert or Banach space) with norm  $\|\cdot\|$  so that  $d(x, x') = \|x - x'\|$ , which is the case in this paper.

Ferraty and Vieu [4] provided a consistent estimate for the nonlinear regression operator  $r$ , based on the usual finite-dimensional smoothing ideas, that is

$$\hat{T}_n(x) := \frac{\sum_{i=1}^n Y_i K\left(\frac{\|x - X_i\|}{h_n}\right)}{\sum_{i=1}^n K\left(\frac{\|X_i - x\|}{h_n}\right)}, \quad (2)$$

where  $K(\cdot)$  is a real-valued kernel and  $h_n$  the bandwidth, is a sequence of positive real numbers converging to 0 as  $n \rightarrow \infty$ . Note that the bandwidth  $h_n$  depends on  $n$ , but we drop this index for notational convenience. In what follows  $K_h(u)$  stands

E-mail address: mohamed.cherfi@gmail.com.

for  $K(\frac{u}{h})$ . The estimator defined in (2) is a generalization to the functional framework of the classical Nadaraya–Watson regression estimator. The asymptotic properties of this estimate have been studied extensively by several authors, we cite among others [7], for a complete survey see the monograph by Ferraty and Vieu [5].

There exists a wide large deviations literature, for an extensive overview on the area we refer to the monograph of Dembo and Zeitouni [3]. The large deviations behavior of the Nadaraya–Watson estimate of the regression function, have been studied at first by Louani [12], sharp results have been obtained by Joutard [9] in the univariate framework. In the multidimensional case Mokkadem et al. [13] obtained pointwise large and moderate deviations results for the Nadaraya–Watson and recursive kernel estimators of the regression.

Large deviations results have many implications in statistical analysis, we refer to [1] for a good survey on the subject. The main concern in the nonparametric setup is with asymptotic efficiency and the inaccuracy rate. These questions have been investigated in [11] for the density case. As an application of large deviations results, Louani [12] gives the inaccuracy rate in conditional distribution function estimation. Another application can be found in [2], where the authors provide Bahadur efficiency of symmetry tests. We refer to the monograph of Nikitin [14] for an account of results on the asymptotic efficiency.

All the results of large deviations in nonparametric regression only dealt with the case of scalar covariates. In this Note, our aim is to obtain such results in the functional case. Inaccuracy rate in nonparametric regression estimation on functional data will be considered in a future work. We are also interested to obtain a minimax Bahadur-type risk of estimating a nonparametric regression function, in the spirit of Korostelev [10].

In this Note, we are interested in the problem of establishing large deviations principles of the regression operator estimate  $\hat{r}_n(\cdot)$ . The results stated in the Note deal with pointwise and uniform large deviations probabilities of  $\hat{r}_n(\cdot)$  from  $r(\cdot)$ .

## 2. Results

Let  $F_x(h) = P[\|X_i - x\| \leq h]$ , be the cumulative distribution of the real variable  $W_i = \|X_i - x\|$ . As in [7], let  $\varphi(\cdot)$  be the real valued function defined by

$$\varphi(u) = \mathbb{E}\{r(X) - r(x) \mid \|X - x\| = u\}. \quad (3)$$

Before stating our results, we will consider the following conditions:

(C.1) The kernel  $K$  is of class  $C^1$  with continuous derivative on the compact support  $[0, 1]$ .

(C.2) The operator  $r$  verifies the following Lipschitz property:

$$\text{There exists } \beta \text{ such that } \forall(u, v) \in \mathcal{X}^2, \exists C, |r(u) - r(v)| \leq C\|u - v\|^\beta. \quad (4)$$

(C.3) There exist three functions  $\ell(\cdot)$ ,  $\phi(\cdot)$  (supposed increasing and strictly positive and tending to zero as  $h$  goes to zero) and  $\zeta_0(\cdot)$  such that

$$(i) F_x(h) = \ell(x)\phi(h) + o(\phi(h)),$$

$$(ii) \text{ for all } u \in [0, 1], \lim_{h \rightarrow 0} \frac{\phi(uh)}{\phi(h)} =: \lim_{h \rightarrow 0} \zeta_h(u) = \zeta_0(u).$$

(C.4)  $\varphi'(0)$  exists.

**Remark 1.** There exist many examples fulfilling the decomposition mentioned in condition (C.3), see for instance Proposition 1 in [7]. The conditions stated above are classical in nonparametric estimation for functional data, we refer to [7] and the references therein. However this condition is somewhat restrictive but is necessary to prove our results, we refer among others to [8] for less restrictive conditions.

Let now introduce the following functions:

$$I(t) = t\lambda \int_0^1 K'(u) \exp\{-t\lambda K(u)\} \zeta_0(u) du; \quad (5)$$

$$\Gamma_x^+(\lambda) = \inf_{t>0} \{\ell(x)I(t)\}; \quad \Gamma_x^-(\lambda) = \inf_{t>0} \{\ell(x)I(-t)\} \text{ and } \Gamma_x(\lambda) = \max\{\Gamma_x^+(\lambda); \Gamma_x^-(\lambda)\}.$$

Let  $x$  be an element of the functional space  $\mathcal{X}$  and  $\lambda > 0$ . Our first theorem deals with pointwise large deviations probabilities.

**Theorem 2.1.** Assume that the conditions (C.1)–(C.4) are satisfied. If  $n\phi(h) \rightarrow \infty$ , then for any  $\lambda > 0$  and any  $x \in \mathcal{X}$ , we have

$$(a) \quad \lim_{n \rightarrow \infty} \frac{1}{n\phi(h)} \log P(\hat{r}_n(x) - r(x) > \lambda) = \Gamma_x^+(\lambda), \quad (6)$$

$$(b) \quad \lim_{n \rightarrow \infty} \frac{1}{n\phi(h)} \log P(\hat{r}_n(x) - r(x) < -\lambda) = \Gamma_x^-(\lambda), \quad (7)$$

$$(c) \quad \lim_{n \rightarrow \infty} \frac{1}{n\phi(h)} \log P(|\hat{r}_n(x) - r(x)| > \lambda) = \Gamma_\chi(\lambda). \quad (8)$$

To establish uniform large deviations principles for the regression estimator we need the following additional assumptions. Let  $\mathcal{C}$  be some compact subset of  $\mathcal{X}$  and  $B(z_k, \xi)$  a ball centered at  $z_k \in \mathcal{X}$  with radius  $\xi$ , such that for any  $\xi > 0$ ,

$$\mathcal{C} \subset \bigcup_{k=1}^{\tau} B(z_k, \xi), \quad (9a)$$

$$\exists \alpha > 0, \quad \exists C > 0, \quad \tau \xi^\alpha = C. \quad (9b)$$

Before stating the theorem about the uniform version of our result, we introduce the following function:

$$g(\lambda) = \sup_{x \in \mathcal{C}} \Gamma_\chi(\lambda). \quad (10)$$

**Theorem 2.2.** Assume that the conditions (C.1)–(C.4) are satisfied. If  $n\phi(h) \rightarrow \infty$ , then for any compact set  $\mathcal{C} \subset \mathcal{X}$  satisfying conditions (9) and for any  $\lambda > 0$ ,

$$\lim_{n \rightarrow \infty} \frac{1}{n\phi(h)} \log P\left(\sup_{x \in \mathcal{C}} |\hat{r}_n(x) - r(x)| > \lambda\right) = g(\lambda). \quad (11)$$

**Remark 2.** (i) The above conditions on the covering of the compact set  $\mathcal{C}$  by a finite number of balls, the geometric link between the number of balls  $\tau$  and the radius  $\xi$  are necessary to prove uniform convergence in the context of functional nonparametric regression and many functional nonparametric settings, see the discussion in [6].

(ii) More recently, Ferraty et al. [8] proved a very nice result about the uniform almost complete convergence of  $\hat{r}_n(\cdot)$  by use of Kolmogorov's  $\epsilon$ -entropy of  $\mathcal{C}$ . It will be of interest to prove Theorem 2.2, under their conditions.

## References

- [1] R.R. Bahadur, Some limit theorems in statistics, in: Conference Board of the Mathematical Sciences Regional Conference Series, in: Appl. Math., vol. 4, Society for Industrial and Applied Mathematics, Philadelphia, PA, 1971.
- [2] N. Berrahou, D. Louani, Efficiency of some tests when testing symmetry hypothesis, J. Nonparametr. Stat. 18 (7–8) (2006) 465–482 (2007).
- [3] A. Dembo, O. Zeitouni, Large Deviations Techniques and Applications, second ed., Appl. Math. (New York), vol. 38, Springer-Verlag, New York, 1998.
- [4] F. Ferraty, P. Vieu, Nonparametric models for functional data, with application in regression, time-series prediction and curve discrimination, in: The International Conference on Recent Trends and Directions in Nonparametric Statistics, J. Nonparametr. Stat. 16 (1–2) (2004) 111–125.
- [5] F. Ferraty, P. Vieu, Nonparametric functional data analysis, in: Theory and Practice, in: Springer Ser. Statist., Springer, New York, 2006.
- [6] F. Ferraty, P. Vieu, Erratum of: “Nonparametric models for functional data, with application in regression, time-series prediction and curve discrimination” [J. Nonparametr. Stat. 16 (1–2) (2004) 111–125], J. Nonparametr. Stat. 20 (2) (2008) 187–189.
- [7] F. Ferraty, A. Mas, P. Vieu, Nonparametric regression on functional data: inference and practical aspects, Aust. N. Z. J. Stat. 49 (3) (2007) 267–286.
- [8] F. Ferraty, A. Laksaci, A. Tadj, P. Vieu, Rate of uniform consistency for nonparametric estimates with functional variables, J. Statist. Plann. Inference 140 (2) (2010) 335–352.
- [9] C. Joutard, Sharp large deviations in nonparametric estimation, J. Nonparametr. Stat. 18 (3) (2006) 293–306.
- [10] A. Korostelev, A minimaxity criterion in nonparametric regression based on large-deviations probabilities, Ann. Statist. 24 (3) (1996) 1075–1083.
- [11] D. Louani, Large deviations limit theorems for the kernel density estimator, Scand. J. Statist. 25 (1) (1998) 243–253.
- [12] D. Louani, Some large deviations limit theorems in conditional nonparametric statistics, Statistics 33 (2) (1999) 171–196.
- [13] A. Mokkadem, M. Pelletier, B. Thiam, Large and moderate deviations principles for kernel estimators of the multivariate regression, Math. Methods Statist. 17 (2) (2008) 146–172.
- [14] Y. Nikitin, Asymptotic Efficiency of Nonparametric Tests, Cambridge University Press, Cambridge, 1995.