Quadratic error of the conditional quantile for functional data in the local linear estimation with missing at random

W. Bouabsa

Abstract. We consider in this paper the asymptotic mean square error and estimator convergence rates based on the conditional quantile function of the local linear process by taking into account missing random, demonstrate the practicality of our estimate approach and highlight its superiority to standard kernel estimation, A simulation analysis was performed on finite-sized samples.

M.S.C. 2010: Primary: 62G05; 62G20; 62G07; 62G32; 62G30, Secondary: 62E20.
Key words: Nonparametric local linear estimation; mean squared error; conditional quantile function, functional variable, missing at random.

1 Introduction

The conditional quantile estimate is a very good and important subject of statistics. This approximation is used to create predictive intervals, to assess citation curves and provides a predictive method when the regression is not quite suited to other conditions to better predict the effect of the X explanatory variable on the Y answer variable.

Stone [24] appears to be the first to arrive at the conditional quantile calculation in recent decades, obtaining the probability convergence of the estimator based on the empirical estimation of the conditional cumulative distribution. In the i.i.d. case, Samanta [23] defined asymptotic normality and uniform convergence (see for more information Berlinet et al. [1]) and Roussas [22].

The first results having a practical highlighting factor were provided by Cardot et al. [5]. A conditional quantile estimator has been built on the spline method, regarded as a continuous linear form defined on a Hilbert space. Ferraty et al. [14] considered the nonparametric approach to that model.
They evaluated the virtually complete convergence rate for the Kernel estimator in the single i.i.d. case, and in each of these cases, Ezzahrioui and Ould-Said ([9], [10])

Applied Sciences, Vol. 24, 2022, pp. 12-24.

studied the asymptotic normality of the estimator (the strong mixing conditions, or the i.i.d. case).
Recently, Laksaci et al. ([19],[18]) proposed an alternative based on the L^1 estimate formula. Under the mixing hypothesis, Dabo-Niang and Laksaci [7] considered the convergence of the nonparametric quantile regression in the L^p-norm.
The local linear estimation methodology has many advantages in the case of finite-dimensional data over the kernel process, such as bias reduction and adaptation of edge effects (see [12] for more information); in the recent work [4], the local linear estimation was investigated in the functional data case.
Some missing data exists in numerous fields, including surveys, clinical trials and longitudinal studies. Responses may be missing, and methods for processing missing data often depend on the mechanism generating the missing values, see Efromovich [11]. In several practical works, such as pharmaceutical tracing, sample surveys or reliability, it might happen that data are not completely examined, and randomly few answers are provided.
Baíllo and Grané [2] developed an estimator of local linear regression and examined its asymptotic behavior (conditional mean squared error convergence rate), when the explanatory variable takes values in a Hilbert space. Barrientos et al. [3] developed the locally modeled regression estimator, which describes the almost complete convergence (with rate) of the presented estimator. Demongeot et al. [8] proposed estimating the conditional density function on the basis of the local modeling approach, when the explanatory variable is functional.

The remainder of our paper is organized as follows. In Section 2, we present our functional model, give basic notations and describe our assumptions. In Section 3 we state the main theoretical result, and the result of the paper about the mean squared convergence; further, a simulation study is given to prove efficiency in Section 4. Conclusions and perspectives conclude the paper in Section 5.

2 Model estimation

2.1 The conditional quantile estimate for the kernel

Let's consider the $(X_i, Y_i)_i, i \geq 1$ sequence of an (i.i.d) independent and identically random pair according to the (X, Y) pair distribution, with X taking values in a semi-metric space $(\mathcal{F}, d)$ and Y being real-valued random variables.

In the following, x will be a fixed in $\mathcal{F}$, $\mathcal{N}_x$ (resp. $\mathcal{N}_y$) will denote a fixed neighborhood of a fixed point x (resp. of y).
The conditional distribution function $F^x(y)$ of Y for given $X = x$ is defined as follows

$$F^x(y) = \mathbb{P}\left(Y \leq y | X = x\right).$$

The conditional quantile of the order $\alpha \in (0,1)$, is defined by

$$t_\alpha(x) = \inf\{y \in \mathbb{R} : F^x(y) \geq \alpha\}.$$

We note that $F^x(y)$ admits a unique conditional quantile, $\forall x \in \mathcal{F}$.
Let $\alpha \in (0,1)$; then the α-th conditional quantile, denoted by $t_\alpha(x)$, satisfies the

following equation:

$$F^x(t_\alpha(x)) = \alpha. \tag{2.1}$$

Therefore, the new conditional quantile estimator can be specified as

$$\widehat{t_\alpha}(x) = \inf\{y \in \mathbb{R} : \widehat{F}^x(y) \geq \alpha\},$$

which satisfies

$$\widehat{F}^x(\widehat{t_\alpha}(x)) = \alpha. \tag{2.2}$$

Our key goal is to determine the conditional quantile function $\widehat{F}^x(\widehat{t_\alpha}(x))$.
Nonetheless, in the case of incidentally missing response variable, an available incomplete sample of n from (X, Y, δ) is $\{(X_i, Y_i, \delta_i), 1 \leq i \leq n\}$; if X_i is completely observed, $\delta_i = 1$ if Y_i is observed, and $\delta_i = 0$ if not.

Meantime, the Bernoulli δ random variable is satisfied

$$\mathbb{P}(\delta = 1|X, Y) = P(\delta = 1|X) = P(X),$$

where $P(X)$ is a functional operator, known as the conditional probability - given the predictor - of the observing response, often unknown. This method demonstrates that, given X, δ and Y are conditionally independent. Missing at random is a common concept of missing data for statistical analysis, and is fair in many realistic cases (we refer to Ferraty et al. [17], and Ling et al. [20]).

The $F^x(\cdot)$ function, as indicated by Fan [13], can be seen as a nonparametric regression model with variable response $G(h_G^{-1}(\cdot - Y_i))$, where $\int G = 1$ and h_G is a sequence of positive real numbers. That is motivated by the fact that

$$\mathbb{E}[G(h_G^{-1}(y - Y_i))|X_i = x] \to F^x(y) \text{ as } h_G \to 0.$$

We use a technique that extends the linear local ideas to the infinite dimensional model (see Barrientos et al. [3] and Demongeot et al. [8]). We combine this idea with the consideration of the data missing at random.
After that we adopt the fast functional local modeling of the $F^x(y)$ conditional distribution function calculated for the arg min value estimated at $\widehat{a}$, where the pair $(\widehat{a}, \widehat{b})$ is achieved with the rule of optimization

$$\min_{(a,b)\in\mathbb{R}^2} \sum_{i=1}^{n} \left(h_G^{-1} G(h_G^{-1}(y - Y_i)) - a - b\beta(X_i,\ x)\right)^2 \delta_i K(h_K^{-1}\varpi(x, X_i)). \tag{2.3}$$

With the locating functions $\varpi(.;.)$ and $\beta(.;.)$ defined from $\mathcal{F}^2$ into $\mathbb{R}$, such that:

$$\forall \xi \in \mathcal{F}; \beta(\xi;\xi) = 0 \ \ and \ \ d(.;.) = |\varpi(.;.)|,$$

where K is a kernel function, G is a distribution function (df) and $h = h_K := h_{K,n}$ and $h_G = h_{G,n}$ are suites of positive real numbers, going to zero as n goes to infinity.

Obviously, the basic calculations lead to

$$\widehat{F}^x(y) = \frac{\sum_{1\le i,j\le n} Q_{ij}(x)G({h_G}^{-1}(y-Y_j))}{\sum_{1\le i,j\le n} Q_{ij}(x)}, \qquad \forall y \in \mathbb{R}, \tag{2.4}$$

with $Q_{ij}(x) = \beta_i(\beta_i - \beta_j)\delta_i\delta_j K(h_K^{-1}\varpi(x,X_i))K(h_K^{-1}\varpi(x,X_j))$ and $\beta_i = \beta(X_i,x)$. For $n \ge 1$ and for $y \in \mathbb{R}$, this is the final form is estimator.

Before dealing with the asymptotic estimator (2.4), we require more notes and simple hypotheses

(H1) $\phi_x(r) := \phi_x(-r;r) > 0$. For any $r > 0$, a $\chi_x(.)$ function exists, such that

$$\forall t \in (-1,1), \lim_{h_K\to 0} \frac{\phi_x(t\, h_K; h_K)}{\phi_x(h_K)} = \chi_x(t).$$

(H2) For every $l \in \{0,2\}$ and $j = 0,1$, we denote the functions

$$\Upsilon_{l,j}(x,y) = \frac{\partial^l F^{x^j}(y)}{\partial y^l} \;\; \textit{and} \;\; \Upsilon_{l,j}(s) = \mathbb{E}[\Upsilon_{l,j}(X,y) - \Upsilon_{l,j}(x,y)|\beta(x,X) = s], \tag{2.5}$$

and $\Upsilon^1_{l,j}(0)$, $\Upsilon^2_{l,j}(0)$ for the function $\Upsilon_{l,j}(.)$ exist.

(H3) The function $\beta(.,.)$ and $\varpi(.,.)$ are such that

$$\forall z \in F, |\varpi(x,z)| = d(x,z) \;\; \textit{and} \;\; C_1|\varpi(x,z)| \le |\beta(x,z)| \le C_2|\varpi(x,z)|,$$

where $C_1 > 0; C_2 > 0$,

$$\sup_{v\in B(x;r)} |\beta(v;x) - \varpi(x;v)| = o(r),$$

and

$$h_K \int_{B(x;h_K)} \beta(v;x)dP(v) = o\left(\int_{B(x;h_K)} \beta^2(v;x)dP(v)\right),$$

with $B(x;r) = \{z \in \mathcal{F}/|\varpi(z;x)| \le r\}$, where $dP(x)$ is the cumulative distribution of X.

(H4) The kernel function K is positive and differentiable,represented in $(-1;1)$ and satisfies

$$K^2(1) - \int_{-1}^{1} (K^2(v))^{(1)}\chi_s(v)d(v) > 0.$$

(H5) The kernel G is a differentiable function and $G^{(1)}$ is positive bounded Lipschitzian continuous function such that

$$\int |t|^{b_2} G^{(1)}(t)dt < \infty \;\; \textit{and} \;\; \int (G^{(1)})^2(t)dt < \infty \;\; \textit{and} \;\; \int G^{(1)}(t)dt = 1.$$

(H6)

$$\exists 0 < \beta < 1, \; F^x(y) \le 1 - \beta, \; \forall (x,y) \in \mathcal{F} \times \mathbb{R}.$$

(H7) The bandwidths h_K and h_G satisfy
$\lim_{n\to\infty} h_K = 0$, $\lim_{n\to\infty} h_G = 0$ width $\lim_{n\to\infty} n\, h_G^{(j)}\phi_x(h_K) = \infty$, for j=0,1.

(H8) $\forall (y_1, y_2) \in \mathbb{R}^2$, $|G(y_1) - G(y_2)| \leq C|y_1 - y_2|$.

(H9) The $F^x(y)$ conditional distribution function is differentiable, continuous, and it has a uniformly bounded first derivative, denoted by $f^x(y)$ and satisfies $\forall\ (y_1, y_2)\ \in \mathbb{R}\times\mathbb{R}$, $\forall\ (x_1, x_2)\ \in N_{x_1}\times N_{x_2}$, there exist some constants $C, b_1, b_2 > 0$, such that, for $j = 0, 1$, we have

$$|F^{x_1^{(j)}}(y_1) - F^{x_2^{(j)}}(y_2)| \leq C(d(x_1, x_2)^{b_1} + |y_1 - y_2|^{b_2}).$$

(H10) The operator $P(.)$ on N_x is continuous and $P(X) > 0$.

Remark 2.1. Several Comments on the hypotheses follow: Hypothesis (H1) is the explanatory variable concentration property in small balls. The $\chi_x(.)$ function plays a crucial role in any asymptotic analysis, particularly for the term variance. The condition (H2) is used to monitor the regularity of the functional space of our model; this is important to determine the convergence rate bias concept. The assumption (H3) is a technical assumption. As established by Rachdi et al. [21], Barrientos [3], the assumptions (H4) and (H5) on the kernels K, G and $G^{(1)}$ are standard conditions for the quadratic error determination for functional results. The hypotheses (H6), (H7) and (H8), (H9) are technical conditions similar to those assumed in Ferraty et al. [16]. The assumption (H10) characterize the case of the functional estimation in missing data at random; this is also used in Ling et al. [20] as technical condition for the concision of the proofs of the main results.

3 Results

In this section we are going to state our theoretical results. In the first subsection, the proof of our main Theorem 3.1 is demonstrated in terms of Theorems 3.2 and the following Lemmas: Lemma 3.3, Lemma 3.4, Lemma 3.5, Lemma 3.6 and Lemma 3.7. Note that our result extends the complete data from Demongeot et al. [8] to the missing at random case. The completely observed data is obtained by taking $\delta = 1$ in our case of study.

3.1 Main results: Quadratic error convergence

Theorem 3.1. *Under assumptions (H1), (H4) and (H8), we obtain*

$$\begin{aligned}\mathbb{E}\left[\widehat{t_\alpha}(x) - t_\alpha(x))\right]^2 &= B_{F,G}^2(x,y)h_G^4 + B_{F,K}^2(x,y)h_K^4 + \frac{V_{GK}^F(x,y)}{n\phi_x(h_K)}\\ &\quad + o(h_G^4) + o(h_K^4) + o\left(\frac{1}{n\phi_x(h_K)}\right).\end{aligned}$$

where,

$$B_{F,G}(x,y) = \frac{1}{2}\frac{\partial^2 F^x(y)}{\partial y^2}\int t^2 G^{(1)}(t)\mathrm{d}t,$$

$$B_{F,K}(x,y) = \frac{1}{2}\Upsilon^{(2)}_{0,0}(0)\left[\frac{(K(1)-\int_{-1}^{1}(v^2K(v))^{(1)}\chi_x(v)\mathrm{d}v)}{\left(K(1)-\int_{-1}^{1}K^{(1)}(v)\chi_x(v)\mathrm{d}v\right)}\right],$$

and

$$V^F_{GK}(x,y) = F^x(y)(1-F^x(y))\left[\frac{\left(K^2(1)-\int_{-1}^{1}(K^2(v))^{(1)}\chi_x(v)\mathrm{d}v\right)}{\left(K(1)-\int_{-1}^{1}(K(v))^{(1)}\chi_x(v)\mathrm{d}v\right)^2}\right].$$

We note that

$$\widehat{F}^x(y) = \frac{\widehat{F}^x_M(y)}{\widehat{F}^x_R},$$

where,

$$\widehat{F}^x_M(y) = \frac{1}{n(n-1)\mathbb{E}[Q_{12}(x)]}\sum_{1\le i\ne j\le n} Q_{ij}(x)G(h_G^{-1}(y-Y_j)),$$

and

$$\widehat{F}^x_R = \frac{1}{n(n-1)\mathbb{E}[Q_{12}(x)]}\sum_{1\le i\ne j\le n} Q_{ij}(x).$$

Proof. Taking into account Lemma 3.6, at point $\widehat{t_\alpha}(x)$ we write for the function $\widehat{F}^x$ the Taylor expansion of order one at some $t_\alpha{}^{**}(x)$, comprising $\widehat{t_\alpha}(x)$ and $t_\alpha(x)$, and we get

$$(\widehat{t_\alpha}(x)-t_\alpha(x))\widehat{F}^{x^{(1)}}(t^{**}_\alpha(x)) = \widehat{F}^x(\widehat{t_\alpha}(x))-\widehat{F}^x(t_\alpha(x))$$

$$|\widehat{t_\alpha}(x)-t_\alpha(x)| = \frac{1}{\widehat{F}^{x^{(1)}}(t^{**}_\alpha(x))}[|\widehat{F}^x(t_\alpha(x))-F^x(t_\alpha(x))|];$$

then,

$$|\widehat{t_\alpha}(x)-t_\alpha(x)|\widehat{F}^{x^{(1)}}(t^{**}_\alpha(x)) = O_{a.s}\left(\left|\widehat{F}^x(t_\alpha(x))-F^x(t_\alpha(x))\right|\right).$$

The result given by Lemma 3.6 ensures that $t_\alpha{}^*(x)$ tends to $t_\alpha(x)$ a.s., and it provides

$$|\widehat{t_\alpha}(x)-t_\alpha(x)| = O_{a.s}\left(\left|\widehat{F}^x(t_\alpha(x))-F^x(t_\alpha(x))\right|\right).$$

Therefore, we conclude that the Theorem is a simple consequence of Theorem 3.2. □

Theorem 3.2. *Under assumptions (H1)-(H4)-(H7) and (H8)-(H9), we have*

$$\sup_{y\in[t_\alpha(x)-\epsilon,t_\alpha(x)+\epsilon]}\mathbb{E}\left[\widehat{F}^x(y)-F^x(y)\right]^2 = B^2_{F,G}(x,y)h_G^4 + B^2_{F,K}(x,y)h_K^4 + \frac{V^F_{GK}(x,y)}{n\phi_x(h_K)} + o(h_G^4)+o(h_K^4)+o\left(\frac{1}{n\phi_x(h_K)}\right),$$

Proof.

$$\sup_{y\in[t_\alpha(x)-\epsilon,t_\alpha(x)+\epsilon]} \mathbb{E}\left[\widehat{F}^x(y)-F^x(y)\right]^2 = \sup_{y\in[t_\alpha(x)-\epsilon,t_\alpha(x)+\epsilon]} \left(\left[\mathbb{E}\left[\widehat{F}^x(y)\right]-F^x(y)\right]^2 + Var\left[\widehat{F}^x(y)\right]\right).$$

In order to simplify the bias and variance of the second term in the correct equality, we just use the Ferraty et al. analysis (see [15]), to get

$$\sup_{y\in[t_\alpha(x)-\epsilon,t_\alpha(x)+\epsilon]} \mathbb{E}\big[\widehat{F}^x(y)\big]-F^x(y) = \sup_{y\in[t_\alpha(x)-\epsilon,t_\alpha(x)+\epsilon]} \left(\mathbb{E}[\widehat{F}^x_M(y)]-F^x(y)\right)$$
$$+\sup_{y\in[t_\alpha(x)-\epsilon,t_\alpha(x)+\epsilon]} \frac{\mathbb{E}[\widehat{F}^x_M(y)(\widehat{F}^x_R-\mathbb{E}[\widehat{F}^x_R])]}{(\mathbb{E}[\widehat{F}^x_R])^2} + \sup_{y\in[t_\alpha(x)-\epsilon,t_\alpha(x)+\epsilon]} \frac{\mathbb{E}[\widehat{F}^x(y)(\widehat{F}^x_R-\mathbb{E}[\widehat{F}^x_R])^2]}{(\mathbb{E}[\widehat{F}^x_R])^2},$$

and

$$\sup_{y\in[t_\alpha(x)-\epsilon,t_\alpha(x)+\epsilon]} Var\big[\widehat{F}^x(y)\big] = \sup_{y\in[t_\alpha(x)-\epsilon,t_\alpha(x)+\epsilon]} Var\left(\widehat{F}^x_M(y)\right)$$
$$-\sup_{y\in[t_\alpha(x)-\epsilon,t_\alpha(x)+\epsilon]} 4\left(\mathbb{E}[\widehat{F}^x_M(y)]\right)Cov\left(\widehat{F}^x_M(y),\widehat{F}_R(x)\right)$$
$$+\sup_{y\in[t_\alpha(x)-\epsilon,t_\alpha(x)+\epsilon]} 3\big(\mathbb{E}[\widehat{F}^x_M(y)]\big)^2 Var\left(\widehat{F}^x_R\right)+o\left(\frac{1}{n\phi(h_K)}\right).$$

□

The following Lemmas will be useful for proving Theorem 3.2, hence Theorem 3.1 as well.

Lemma 3.3. *Under the hypotheses of Theorem 3.2, we get*

$$\sup_{y\in[t_\alpha(x)-\epsilon,t_\alpha(x)+\epsilon]} \left(\mathbb{E}\left[\widehat{F}^x_N(y)\right]-F^x(y)\right) = B_{F,G}(x,y)h_G^2 + B_{F,K}(x,y)h_K^2 + o(h_G^2)+o(h_K^2).$$

Proof. By the compactness of $S=\{t_\alpha(x)-\epsilon, t_\alpha(x)+\epsilon\}$, the proof is reduced to evaluating the following quantities

$$\mathbb{E}\left[\delta_1 K_1 G_1^k \beta_1^l\right] \text{ for } l=0,1,2, \text{ and } k=0,1.$$

We make use of the variables with respect to X and δ; they are conditionally independent of Y, and G is a distribution function. So, for all $l=0,1,2$, and $k=0,1$, we have

$$\mathbb{E}\left[\delta_1 K_1 G_1^k \beta_1^l\right] = (o(1)+P(X))\mathbb{E}\left[K_1\beta_1^l\right]) = O(h_K^l h_G^k \phi_x(h_K)). \tag{3.1}$$

For the quantities $\mathbb{E}[\widehat{F}^x_M(y)]$, we use integration by parts, and get

$$\mathbb{E}[\widehat{F}^x_M(y)] = \frac{1}{\mathbb{E}[Q_{12}]}\mathbb{E}[Q_{12}\mathbb{E}[G_2|X_2]] \text{ with } \mathbb{E}[G_2|X_2] = \int G_2^{(1)}(t)F^{X_2}(y-h_G t)\mathrm{d}t.$$

So, considering that

$$\mathbb{E}\left[Q_{12}\mathbb{E}[G_2|X_2]\right] = \mathbb{E}\left[\delta_1\beta_1^2(x)K_1(x)\beta_2(x)K_2(x)P(X_2)(\mathbb{E}[G_2|X_2]\right), \tag{3.2}$$

we obtain

$$\mathbb{E}\left[\widehat{F}^x_M(y)\right] = F^x(y) + \frac{h_G^2}{2}\frac{\partial^2 F^x(y)}{\partial y^2}\int t^2 G_2^{(1)}(t)\mathrm{d}t + o(h_G^2)$$
$$+\frac{h_K^2}{2}\Upsilon^{(2)}_{0,0}(0)\frac{\left(K(1)-\int_{-1}^1 (u^2K(v))^{(1)}\chi_x(v)\mathrm{d}v\right)}{\left(K(1)-\int_{-1}^1 K^{(1)}(v)\chi_x(v)\mathrm{d}v\right)} + o(h_K^2).$$

□

Lemma 3.4. *Under the hypotheses of Theorem 3.2, we have*

$$\sup_{y\in[t_\alpha(x)-\epsilon,t_\alpha(x)+\epsilon]} Var\left[\widehat{F}_M^x(y)\right] = \frac{V_{GK}^F(x,y)}{n\phi_x(h_K)} + o\left(\frac{1}{n\phi_x(h_K)}\right).$$

Proof. It is clear that, by the compactness of $S = \{t_\alpha(x)-\epsilon, t_\alpha(x)+\epsilon\}$, we have

(3.3)

$$\begin{aligned} Var[\widehat{F}_M^x(y)] &= \frac{1}{(n(n-1)h_G(\mathbb{E}[Q_{12}]))^2}\Big[n(n-1)\mathbb{E}[Q_{12}^2(G_2)^2] + n(n-1)\mathbb{E}[Q_{12}Q_{21}G_2G_1] \\ &\quad +n(n-1)(n-2)\mathbb{E}[Q_{12}Q_{13}G_2G_3] + n(n-1)(n-2)\mathbb{E}[Q_{12}Q_{23}G_2G_3] \\ &\quad +n(n-1)(n-2)\mathbb{E}[Q_{12}Q_{31}G_2G_1] + n(n-1)(n-2)\mathbb{E}[Q_{12}Q_{32}(G_2)^2] \\ &\quad -n(n-1)(4n-6)\mathbb{E}[Q_{12}G_2]^2\Big]. \end{aligned}$$

Like in the 3.2 and 3.1, by using the same steps as in Lemma 3.3, we get

$$(3.4)\quad \begin{cases} \mathbb{E}[Q_{12}^2G_2^2] = O(h_K^4\phi_x^2(h_K)), \mathbb{E}[Q_{12}Q_{21}G_1G_2] = O(h_K^4\phi_x^2(h_K)), \\ \mathbb{E}[Q_{12}Q_{13}G_2G_3] = (F^x(y))^2\mathbb{E}[\beta_1^4K_1^2]\mathbb{E}^2[K_1] + o(h_K^4\phi_x^3(h_K)), \\ \mathbb{E}[Q_{12}Q_{23}G_2G_3] = (F^x(y))^2\mathbb{E}[\beta_1^2K_1]\mathbb{E}[\beta_1^2K_1^2]\mathbb{E}[K_1] + o(h_K^4\phi_x^3(h_K)), \\ \mathbb{E}[Q_{12}Q_{31}G_2G_1] = (F^x(y))^2\mathbb{E}[\beta_1^2K_1]\mathbb{E}[\beta_1^2K_1^2]\mathbb{E}[K_1] + o(h_K^4\phi_x^3(h_K)), \\ \mathbb{E}[Q_{12}Q_{32}G_2^2] = F^x(y)\mathbb{E}^2[\beta_1^2K_1]\mathbb{E}[K_1^2] + o(h_K^4\phi_x^3(h_K)), \\ \mathbb{E}[Q_{12}G_1] = O(h_K^2\phi_x^2(h_K)). \end{cases}$$

Therefore, from (3.3) and (3.4), it follows that

$$Var[\widehat{F}_M^x(y)] = \frac{F^x(y)(1-F^x(y))}{\mathbb{E}[K_1^2]}(\mathbb{E}[K_1])^2 + o\left(\frac{1}{n\phi_x(h_K)}\right).$$

Finally, we infer

$$Var[\widehat{F}_M^x(y)] = \frac{F^x(y)(1-F^x(y))}{n\phi_x(h_K)}\left[\frac{\left(K^2(1)-\int_{-1}^1(K^2(v))^{(1)}\chi_x(v)\mathrm{d}v\right)}{\left(K(1)-\int_{-1}^1(K(v))^{(1)}\chi_x(v)\mathrm{d}v\right)^2}\right] + o\left(\frac{1}{n\phi_x(h_K)}\right).$$

□

Lemma 3.5. *Under the hypotheses of Theorem 3.1, we get*

$$\sup_{y\in[t_\alpha(x)-\epsilon,t_\alpha(x)+\epsilon]} Cov(\widehat{F}_M^x(y), \widehat{F}_R^x) = O\left(\frac{1}{n\phi_x(h_K)}\right).$$

Proof. By using 3.2 and 3.1, easy calculations lead to:

$$\begin{aligned} Cov\left(\widehat{F}_M^x(y), \widehat{F}_R^x\right) &= \frac{1}{(n(n-1)h_G(\mathbb{E}[Q_{12}]))^2}Cov\Big(\sum_{1\le i\neq j\le n} Q_{ij}G_j^{(1)}, \sum_{1\le i'\neq j'\le n} Q_{i'j'}\Big) \\ &= \frac{1}{(n(n-1)(\mathbb{E}[Q_{12}]))^2}\Big[n(n-1)\mathbb{E}[Q_{12}^2G_1^{(1)}] + n(n-1)\mathbb{E}[Q_{12}Q_{21}G_2^{(1)}] \\ &\quad +n(n-1)(n-2)\mathbb{E}[Q_{12}Q_{13}G_2^{(1)}] + n(n-1)(n-2)\mathbb{E}[Q_{12}Q_{23}G_2^{(1)}] \\ &\quad +n(n-1)(n-2)\mathbb{E}[Q_{12}Q_{31}G_2^{(1)}] + n(n-1)(n-2)\mathbb{E}[Q_{12}Q_{32}G_2^{(1)}] \\ &\quad -n(n-1)(4n-6)(\mathbb{E}[Q_{12}G_2^{(1)}]\mathbb{E}[Q_{12}]\Big]. \end{aligned}$$

By direct manipulation, we obtain

$$\begin{cases} \mathbb{E}[Q_{12}^2 G_2^{(1)}] = \mathbb{E}[Q_{12}Q_{21}G_2^{(1)}] = O(h_K^4 h_G \phi_x^2(h_K)), \\ \mathbb{E}[Q_{12}Q_{13}G_2^{(1)}] = \mathbb{E}[Q_{12}Q_{31}G_2^{(1)}] = O(h_K^4 h_G \phi_x^3(h_K)), \\ \mathbb{E}[Q_{12}Q_{23}G_2^{(1)}] = \mathbb{E}[Q_{12}Q_{32}G_2^{(1)}] = O(h_K^4 h_G \phi_x^3(h_K)), \end{cases}$$

When $\mathbb{E}[Q_{12}] = O(h_K^2 \phi_x^2(h_K))$, we get

$$Cov\left(\widehat{F}_M^x(y), \widehat{F}_R^x\right) = O\left(\frac{1}{n\phi_x(h_K)}\right).$$

□

Lemma 3.6. *Under the conditions of Theorem 3.1, we have*

$$\widehat{t_\alpha}(x) \longrightarrow t_\alpha(x). \quad a.s. \tag{3.5}$$

Proof. By the compactness of $S = \{t_\alpha(x) - \epsilon, t_\alpha(x) + \epsilon\}$, the proof for this Lemma is focused on the decomposition shown in the study of Chaouch and Khardani (see [6]), for $F^x(y)$ of order α has a unique quantile. Then $\forall \epsilon > 0$, we have

$$\delta(\epsilon) = \min\left\{F^x(t_\alpha(x)+\epsilon) - F^x(t_\alpha(x)), F^x(t_\alpha(x)) - F^x(t_\alpha(x)-\epsilon)\right\},$$

and

$$\forall \epsilon > 0 \ \exists\ \sigma(\epsilon) > 0,\ \forall y \in \mathbb{R}, |t_\alpha(x) - y| \geq \epsilon \Rightarrow |F^x(t_\alpha(x)) - F^x(y)| \geq \sigma(\epsilon).$$

Considering the equations (2.1) equation (2.2), we infer

$$\begin{aligned} \exists\ \sigma(\epsilon) > 0,\ \mathbb{P}(|\widehat{t}_\alpha(x) - t_\alpha(x)| > \epsilon) &\leq \mathbb{P}(|\widehat{F}^x(\widehat{t}_\alpha(x)) - \widehat{F}^x(t_\alpha(x))| > \sigma(\epsilon)) \\ &= \mathbb{P}(|F^x(t_\alpha(x)) - \widehat{F}^x(t_\alpha(x))| > \sigma(\epsilon)) \\ &\leq \sup_{y\in S} |\widehat{F}^x(y) - F^x(y)|. \end{aligned}$$

Since $F^x(y)$ is continuously differentiable, we obtain

$$\infty > \sum_n \mathbb{P}\left(\sup_{y\in S} |\widehat{F}^x(y) - F^x(y)| > \delta(\epsilon)\right) \geq \sum_n \mathbb{P}\left(|\widehat{t}_\alpha(x) - t_\alpha(x)| > \epsilon\right),$$

Which finishes the proof. □

Lemma 3.7. *Under hypotheses (H1), (H4) and (H8), we have*

$$\exists\ \delta > 0 \quad such\ that \sum_{n=1}^{\infty} \mathbb{P}[|\widehat{F}^{x^{(1)}}(t_\alpha^{**}(x))| < \delta] < \infty. \tag{3.6}$$

Proof. Compared to the proof of Theorem 3.1 of Dabo-Niang and Laksaci (see [7]) and by combining equation(3.5) with of Ferraty and Vieu [14, Lemma 11.17], we have

$$\begin{aligned} \left|\widehat{F}^{x^{(1)}}(t_\alpha^{**}(x)) - f^x(t_\alpha(x))\right| &= \left|\widehat{F}^{x^{(1)}}(t_\alpha^{**}(x)) - f^x(t_\alpha^{*}(x)) + f^x(t_\alpha^{**}(x)) - f^x(t_\alpha(x))\right| \\ &\leq \left|\widehat{F}^{x^{(1)}}(t_\alpha^{**}(x)) - f^x(t_\alpha^{**}(x))\right| + |f^x(t_\alpha^{**}(x)) - f^x(t_\alpha(x))| \\ &\leq \left|\widehat{F}^{x^{(1)}}(t_\alpha^{**}(x)) - f^x(t_\alpha^{**}(x))\right| + |f^x(t_\alpha^{**}(x)) - f^x(t_\alpha(x))|. \end{aligned}$$

It follows that

$$\widehat{F}^{x^{(1)}}(t^*_\alpha(x)) - F^x(t_\alpha(x)) \longrightarrow 0 \quad \text{a.co.};$$

then

$$\{|\widehat{F}^{x^{(1)}}(t^{**}_\alpha(x))| < \delta\} \subset \{|\widehat{F}^{x^{(1)}}(t^{**}_\alpha(x)) - F^x(t_\alpha(x))| > \delta\}.$$

We further use Ferraty and Vieu [14, Proposition A.4] and just take $\dfrac{F^{x^{(1)}}(t_\alpha(x))}{2} = \delta.$

Hence

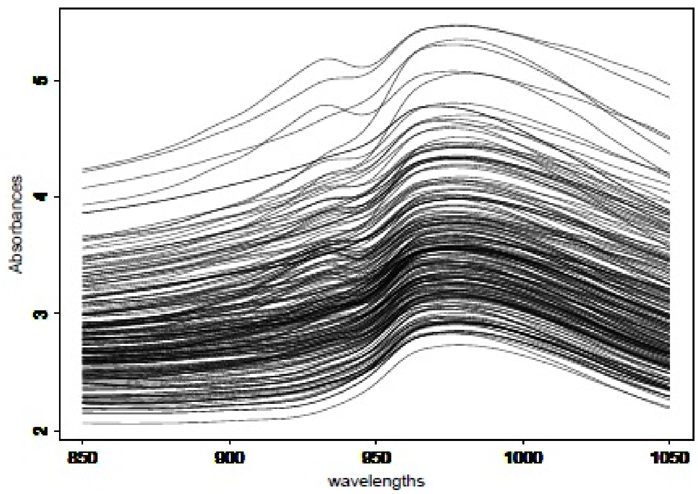
$$\sum_{n=1}^{\infty} \mathbb{P}[|\widehat{F}^{x^{(1)}}(t^{**}_\alpha(x))| < \sigma] < \infty.$$

□

4 Real data application

The conditional quantile function is a very useful tool to define the combination of two random variables. In this section we shall explain the estimation by the local linear approach of this non-parametric model. The main goal is to use virtual real data to demonstrate the applicability in practical case of this system. We shall provide the conditional quantile as a statistical method closely linked to the conditional distribution function estimation. Our aim is to demonstrate the supremacy of the local linear method over the kernel method, and to demonstrate the utility of the conditional quantile in the forecasting context.

For this, we apply the described method to some chemometrical real data. We work on spectrometric data used to classify samples, according to a physicochemical property which is not directly accessible, and therefore requires a specific analysis. The data were obtained using a near-infrared Tecator InfratecFood and Feed Analyzer (with wavelengths ranging between 850 and 1050 nanometers). The measurement is performed through transmitting a sample of finely chopped meat which is then analyzed to determine its fat, by using a chemical process. The spectra correspond to the absorption (-log 10 of the device-measured transmittance) 100 for wavelengths regularly distributed between 850 and 1050 nm. Meat samples were divided into two categories, depending on whether they contain more or less than 20%fat (77 was more spectra corresponding to 20%fat and 138 was less than 20%fat). The problem then is to discriminate the spectra, in order to avoid the costly and time-consuming chemical analysis. For 215 selected pieces of meat, the figure shows absorbance versus wavelength (850-1050). We remind that the main objective of spectrometric analysis is to allow the proportion of some specific chemical content to be discovered (see Ferraty and Vieu [14] for further information on spectrometric data). At this point, one would like to use the X spectrometric curve to predict the proportion of fat in the Y product in meat parts. The data can be found on the Website "http://lib.stat.cmu.edu/datasets/tecator".



The 215 spectrometric curves, $\{X_i(t),\ t \in [850,\ 1050],\ i = 1, \cdots, 215.\}$

However, as described in the real data application (see Ferraty et al. [14]), the prediction problem can be studied by using the conditional quantile approach. Starting from this idea, our objective is to give a comparative study to estimate the conditional quantile of the both methods: the kernel method estimation defined in Ferraty et al. [14] and our model defined in 2.1 and estimated in 2.2.

We divide our samples into two subsets:

$(X_j, Y_j)_{j=1,\ldots 172}$ training sample,
$(X_i, Y_i)_{i=173,\ldots 215}$ test sample.

For the latter, we assume the response values are unknown, and we shall approximate them by 2.1, and estimate them by 2.2.

To verify the effectiveness of this model, in this forecast analysis we compare the quantities $(Y_i)_{i=173,\ldots 215}$ with $\widehat{F}^x(\widehat{t_\alpha}(x))_{i=173,\ldots 215}$,

$$Error(h_k, h_G) = |Y_i - \widehat{F}^x(\widehat{t_\alpha})(X_i)|. \tag{4.1}$$

In brief, we can conclude that the calculation of the conditional quantile used in reading the local linear method is very useful as a forecast model, and we find that this model simple to handle and that our programs swiftly generate the results, providing an error of the local linear conditional quantile estimate of 1.1455, while the classical quantile calculation error used in Ferraty et al. (srr [14]) is 1.53. This result indicates the efficiency of our research.

5 Conclusions

The leading term of the mean square error of the estimator of the conditional quantile by the local linear technique is described in this work. In terms of mean squared error, our estimator for the conditional quantile function with missing at random compares favorably to previous estimators. Our theoretical and practical studies show that the linear local strategy is preferable to the traditional kernel approach. From a theoretical point of view, intriguing possibilities emerge.

In the future, it will be critical to investigate our estimator's asymptotic normality in order to perform statistical tests. The kNN method is a smoothing method that replaces an adaptive estimator. The most essential characteristic of this method is that it enables the creation of a neighborhood that is tailored to the local structure of data. As a result, studying the asymptotic properties of the kNN estimator of the conditional quantile function is substantial. This will be taken into account in our future articles.

Acknowledgements. The authors express their gratitude to the anonymous referees as well as to the Editor in Chief for their insightful remarks and suggestions, which considerably improved this research.

References

[1] A. Berlinet, A. Gannoun and E. Matzner-lober, *Asymptotic properties of convergent estimators of conditional quantiles*, C. R. Acad.Sci. Paris, Serie 1, **326** (1998), 611-614.

[2] A. Baìllo, and Grané, *A Local linear regression for functional predictor and scalar response*, Journal of Multivariate Analysis, **100** (2009), 102-111.

[3] J. Barrientos-Marin, F. Ferraty, and P. Vieu, *Locally modeled regression and functional data*, Journal of Nonparametric Statistics, **22** (5)(2010), 617-632.

[4] E. Boj, P.Delicado, and J. Fortiana, *Distance-based local linear regression for functional predictors*, Computational Statistics and Data Analysis, **54** (2010), 429-437.

[5] H. Cardot, C. Crambes and P. Sarda, *Estimation spline de quantiles conditionnels pour variables explicatives fonctionnelles*, C. R. Math. Paris, **339** (2004), 141-144.

[6] M. Chaouch and S. Khardani, *Randomly censored quantile regression estimation using functional stationary ergodic data*, J. Nonparametric Statistics, **27**, 1 (2015), 65-87.

[7] S. Dabo-Niang, A. Laksaci, *Nonparametric quantile regression estimation for functional dependent data*, Comm. in Statist. Theory and Methods, **41** (2012), 1254-1268.

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

[9] M. Ezzahrioui and E. Ould-Saïd, *Asymptotic results of a nonparametric conditional quantile estimator for functional time series*, Comm in Statist. Theory and Methods, **37** (2008), 2735-2759.

[10] M. Ezzahrioui and E. Ould-Saïd, *Asymptotic normality of the kernel estimators of the conditional quantile in the normed space*, Far East J. Theoretical Statist., **25** (2008), 15-38.

[11] S. Efromovich, *Nonparametric regression with responses missing at random*, Journal of Statistical Planning and Inference, **141** (2011).

[12] J. Fan, I. Gijbels, *Local Polynomial Modelling and its Applications*, Monographs on Statistics and Applied Probability 66, Chapman& Hall, 1996.

[13] J. Fan, *Design-adaptive nonparametric regression*, Journal of the American Statistical Association, **87** (1992), 998-1004.

[14] F. Ferraty, A. Laksaci and P. Vieu, *Estimating some characteristics of the conditional distribution in nonparametric functional models*, J. Statist. Inf. for Stoch. Proc., **9** (2006), 47-76.

[15] F. Ferraty, A. Mas, and P. Vieu, *Non-paramétrique regression on functional data: inference and practical aspects*, Aust. N. Z. J. Stat., **49**, 3 (2007), 207-286.

[16] F. Ferraty, A. Rabhi, and P. Vieu, *Estimation non-paramétrique de la fonction de hasard avec variable explicative fonctionnelle*, Revue de Mathématiques Pures et Appliquées, **53** (2008), 1-18.

[17] F. Ferraty, F. Sued, P. Vieu, *Mean estimation with data missing at random for functional covariables*, Statistics, **47** (2013), 688-706.

[18] A. Laksaci, M. Lemdani and E. Ould-Saïd, *L^1-norm kernel estimator of conditional quantile for functional regressors: consistency and asymptotic normality*, Statist. Probab. Lett., **79** (2009), 1065-1073.

[19] A. Laksaci, M. Lemdani and E. Ould-Saïd, *Asymptotic results for an L^1-norm kernel estimator of the conditional quantile for functional dependent data with application to climatology*, Sankhya, **A 73** (2011), 125-141.

[20] N. Ling, Y. Liu, P. Vieu, *Conditional mode estimation for functional stationary ergodic data with responses missing at random*, Statistics, **50**, 5 (2016), 1-23.

[21] M. Rachdi, A. Laksaci, J. Demongeot, A. Abdali, and F. Madani, *Theoretical and practical aspects of the quadratic error in the local linear estimation of the conditional density for functional data*, Computational Statistics and Data Analysis, **73** (2014), 53-68.

[22] G. Roussas, *Nonparametric functional estimation and related topics*, **7** (1991), 509-529.

[23] M. Samanta, *Non-parametric estimation of conditional quantiles*, Statist. Probab. Lett., **7** (1989), 407-412.

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

Author's address:

Wahiba Bouabsa
Statistics Stochastic Processes, Sidi Bel Abbès University.
BP 89, Sidi Bel Abbès 22000, Algeria.
E-mail: wahiba_bouab@yahoo.fr