The consistency and asymptotic normality of the kernel type expectile regression estimator for functional data

Mustapha Mohammedi, Salim Bouzebda, Ali Laksaci

Version of Record: https://www.sciencedirect.com/science/article/pii/S0047259X20302542
Manuscript_0d327d3f8d4785df1e4969c88273d063

The Consistency and Asymptotic Normality of the Kernel type Expectile Regression Estimator for Functional Data

Mustapha Mohammedi [1a], Salim Bouzebda[b,*], Ali Laksaci [2c]

[a] *Université Djillali Liabès, BP 89, 22000, Sidi Bel Abbès, Algérie*
L.S.P.S., Sidi Bel Abbès, Algérie
[b] *Université de Technologie de Compiègne,*
L.M.A.C., Compiègne, France
[c] *Department of Mathematics, College of Science*
King Khalid University, Abha, 61413, Saudi Arabia

Abstract

The aim of this paper is to nonparametrically estimate the expectile regression in the case of a functional predictor and a scalar response. More precisely, we construct a kernel-type estimator of the expectile regression function. The main contribution of this study is the establishment of the asymptotic properties of the expectile regression estimator. Precisely, we establish the almost complete convergence with rate. Furthermore, we obtain the asymptotic normality of the proposed estimator under some mild conditions. We provide how to apply our results to construct the confidence intervals. The case of functional predictor is of particular interest and challenge, both from theoretical as well as practical point of view. We discuss the potential impacts of functional expectile regression in NFDA with a particular focus on the supervised classification, prediction and financial risk analysis problems. Finally, the finite-sample performances of the model and the estimation method are illustrated using the analysis of simulated data and real data coming from the financial risk analysis.

Keywords: Nonparametric estimation, Kernel type function estimator, Risk measure, Asymmetric least squares regression, Expectiles, Functional data, Almost consistency, Asymptotic normality, Probability convergence, Strong mixing process.
2010 MSC: 62G08, 62E20, 62F12, 62G10, 62G20, 62M10.

1. Introduction

Risk measures such as quantiles and expected shortfall have been studied and developed by several authors from a practical and a theoretical point of view, both risk measures have their own merits and defects. In this paper, we investigate an alternative way based on the least asymmetrically weighted squares estimation, borrowed from the econometrics literature, that is one of the basic tools in statistical applications. This method often involves [44] concept of expectiles, a least squares analogue of the traditional quantiles. The proposed risk measure allows to overcome some drawbacks of the classical measures. It is expressed by

$$\boldsymbol{\theta}_p(Y) = \arg\min_{t\in\mathbb{R}} \mathbb{E}\left[\rho_p(Y,t)\right], \quad \text{for } 0 < p < 1, \tag{1}$$

such that

$$\rho_p(Y,t) = \left|p - \mathbb{1}_{\{(Y-t)\leq 0\}}\right| (Y-t)^2 = p(Y-t)^2\mathbb{1}_{\{(Y-t)>0\}} + (1-p)(Y-t)^2\mathbb{1}_{\{(Y-t)\leq 0\}}, \tag{2}$$

with $\rho_p(Y,t)$ is the scoring function, the level p is called the asymmetry parameter and $\mathbb{1}_A$ stands for the indicator function of the event A. Such risk model was introduced by [3, 44], in the context of linear regression models. It is worth noticing that (2) generalizes the conditional *expec*tation of Y given $X = x$, which coincides with $\boldsymbol{\theta}_p(Y)$ when specifically $p = 1/2$. On the other hand, (2) is similar to the conditional p-quan*tile* of Y given $X = x$, which can be obtained by replacing $(Y-t)^2$ by $|Y-t|$ in (2). This motivates the name the conditional p-expectile. Hence the expectile regression is constructed by combining the ideas of least

[1] *Email address* : mustapha.mohammedi@gmail.com
[2] *Email address* : alilak@yahoo.fr
*Corresponding author. Email address: salim.bouzebda@utc.fr
All authors contributed equally to this work.

square regression to those of the conditional quantile allowing to accumulate the advantages of both models. In particular, its high sensitivity to the extreme value permits for more prudent and reactive risk management, and can be considered as an important advantage, since if we are measuring potential losses, we want that our measure to be sensitive to the extreme tail losses. The main aim of this work is to analyze the asymptotic behavior of this model using the functional nonparametric approach.

Recall that, in a multivariate parametric context, the expectile model has received growing interest. We cite for instance [9] for dapper discussion on the advantages of this model over the classical risk measures such as the quantiles and the expected shortfall. They indicate that the expectiles are perfectly reasonable alternatives to the value at risk (VaR) and the expected shortfall (ES) risk measures. The expectile based risk measure is coherent for $0.5 \leq p < 1$. Moreover, according to [10, 12], the expectile is the only coherent as well as elicitable risk measure. For more details about the statistical significance of the property of elicitability and coherence, we refer the interested reader to the papers of [8, 19, 25, 55]. It turns out that, similar to the conditional mean or the percentile, the expectile regression of order p (or the conditional p-expectile) can be uniquely be defined by

$$\boldsymbol{\theta}(p;x) = \arg\min_{t\in\mathbb{R}} \mathbb{E}\left[\rho_p(Y,t) \mid X = x)\right], \quad \text{for } 0 < p < 1. \tag{3}$$

Although they present differences in their construction, both quantiles and expectiles share similar properties. The main reason for this, as shown in [31], is the fact that expectiles are precisely quantiles but for a transformation of the original distribution. [2] established an important feature is that quantiles and expectiles of the same distribution coincide under the hypothesis of weighted symmetry and pointed out that inference on expectiles is much easier than inference on quantiles. Notice that the quantiles are not always satisfactory and can be criticized for being somewhat difficult to compute as the corresponding loss function is not continuously differentiable. The key advantage of the expectile over the quantile is its efficiency and computing expedience, although it has not a direct interpretation as the quantile in terms of the relative frequency, see [15]. Another substantial difference is that the conditional expectile is based on fundamentally different information and relies on more complete information to measure the risk in the sense that it depends on the form of the entire distribution, and it is easy to see that the conditional expectiles are characterized by the tail expectations in the same way that the conditional quantiles are characterized by the conditional distribution function as shown by the following equation, for more details refer to [1, 44]:

$$\frac{p}{1-p} = \left(\int_{(-\infty,\boldsymbol{\theta}(p;x))} (y - \boldsymbol{\theta}(p;x))^2 dF_Y^x(y)\right)\left(\int_{(\boldsymbol{\theta}(p;x),+\infty)} (y - \boldsymbol{\theta}(p;x))^2 dF_Y^x(y)\right)^{-1}, \tag{4}$$

where $F_Y^x(y)$ represents the conditional distribution function of Y given $\{X = x\}$ and $\boldsymbol{\theta}(p;x)$ is the conditional expectile of order p.

The conditional expectile has been widely studied in the applied area such as econometrics, finance and actuarial science, see, for instance [37] as pioneer work, [18]. Despite this importance, the expectile regression is unexplored compared to both competitive regressions (conditional mean and quantile). [11] have generalized the conditional expectation to the conditional expectile by means of the minimization of an asymmetric quadratic loss function and expose their main properties. [33] have defined the regression quantile estimators via an asymmetric absolute loss, the same authors in [34] have studied a new approach of regression quantiles based upon the estimates of the coefficients of the regression equation. Making use of the expectiles, [16] have proposed the estimations of the VaR and the ES or the CVaR. More details and results on regression are given by many references see [17, 28, 36, 43]. While, these cited works consider the finite dimensional case, in the present work we use this model to analyze the effect of a functional covariate on a scalar response variable. Therefore, we treat in this paper the functional case. Note that analogously to the quantile regression, the conditional expectile allows to a get an exhaustive information on the impact of explanatory variable on the response one by exploring its conditional distribution. On the other hand, the expectile regression can be employed for other standard regression model such as the curves discrimination or the prediction problems. Such questions of functional data analysis are particularly interesting, for many applied areas. Some basic materials in this modern branch of statistics can be found in the monographs of [24, 29, 30, 35, 48] or in the special issues [5, 6, 26]. It should be noted that the principal motivation of the functional statistics is the recent technological development of the measuring instruments allowing the data recording over thinner discretization grids. In this context, the economics or financial data, which are the principal applied areas of the expectile model, constitute a natural source of functional data. Accordingly, the inference statistic by the functional expectile is motivated by the diversity of its application areas including the risk analysis, the discriminant analysis and the forecasting problem.

Notice that the nonparametric estimation in the conditional models when the regressor takes its values in an infinite dimensional space, is an interesting subject and further details can be found in [24], with more details. [1] study the nonparametric weighted symmetry tests, while the almost complete convergence of the conditional quantile kernel estimators is proved by [24] in the independent and identically distributed observations (i.i.d.). The dependent case is considered by many authors, see for example [23, 38, 39]. The papers of [14, 20] investigated the asymptotic normality of this estimator in both cases, i.i.d. and strong mixing variables. Moreover, the kernel method based on the locally linear fit is adapted to the nonparametric estimation of regression

expectiles and percentiles by using the ALS approach, and under the assumptions of stationarity and ρ-mixing, the asymptotic normality for the estimators of conditional expectiles is thoroughly discussed in the paper of [54].

In this contribution we establish the asymptotic properties of the kernel-type estimator of the expectile regression in the case of a functional predictor and a scalar response. To our best knowledge, this problem was open up to present, and it gives the main motivation to our paper. We focus on the almost complete consistency by giving its convergence rate, and we will establish the asymptotic normality of the estimator. The obtained results are stated under standard assumptions in nonparametric functional statistics. Nevertheless, a lot of attention has been paid to others nonparametric functional regression models such as the conditional expectation and the conditional quantiles. Among the wide literature concerning the nonparametric treatment of these functional models we only refer to [4, 24, 38, 39] and the references therein. Finally, we refer the reader to the survey paper [41] for a state of the art as well as some important future tracks on this topic of the nonparametric functional data analysis.

The layout of the article is as follows. In Section 2, we first present and describe the nonparametric model of the expectile regression. In Section 3, we establish our main results of the almost complete convergence of the kernel estimators and the asymptotic normality under some mild conditions in the i.i.d case. In Section 3.2, we extend our results to the dependent setting by showing the consistency in probability for the α-mixing sequences. In Section 3.3, we discuss an application of our results for the construction of the confidence interval. In Section 4, we discuss some connections of the conditional expectile in the supervised classification, prediction and financial risk analysis problems. The implementation of the expectile regression model in practice is given in Section 4.3 and Section 4.4, where a comparison with the quantile regression by illustration on real data is provided. Some concluding remarks are given in Section 5. To avoid interrupting the flow of the presentation, all mathematical developments are relegated to the Section 6. For a better understanding of our proofs, we recall some relevant results in the Appendix.

2. Model and estimators

2.1. The Conditional expectile

Let (X, Y) be a pair of random variable valued in $\mathcal{F} \times \mathbb{R}$, where $(\mathcal{F}, d)$ is a semi-metric space equipped with a semi-metric $d(\cdot,\cdot)$, a semi-metric sometimes called pseudo-metric) $d(\cdot,\cdot)$ is a metric which allows $d(x_1, x_2) = 0$ for some $x_1 \neq x_2$), x is a fixed point in $\mathcal{F}$, defining a topology to measure the proximity between two elements of $\mathcal{F}$ and which is disconnected of the definition of X in order to avoid measurability problems. For $0 < p < 1$, we define the conditional expectile $\boldsymbol{\theta}(p; x)$ of order p of the increasing function $G(\cdot; x)$, defined below, by the unique solution with respect to (w.r.t.) t of the following optimization problem

$$\min_{t\in\mathbb{R}} \left\{ \mathbb{E}\left[p(Y-t)^2 \mathbb{1}_{\{(Y-t)>0\}} \mid X = x\right] + \mathbb{E}\left[(1-p)(Y-t)^2 \mathbb{1}_{\{(Y-t)\leq 0\}} \mid X = x\right] \right\}. \tag{5}$$

By a simple manipulation, we show that $\boldsymbol{\theta}(p; x)$ is the unique solution w.r.t. t of

$$\zeta(p; x) \quad = \quad G(t; x) := \frac{G^1(t; x)}{G^2(t; x)}, \tag{6}$$

where

$$\begin{cases} \zeta(p; x) & = & \dfrac{p}{1-p}, \\ G^1(t; x) & = & -\mathbb{E}\left[(Y-t)\mathbb{1}_{\{(Y-t)\leq 0\}} \mid X = x\right], \\ G^2(t; x) & = & \mathbb{E}\left[(Y-t)\mathbb{1}_{\{(Y-t)>0\}} \mid X = x\right]. \end{cases}$$

We define the conditional expectile $\boldsymbol{\theta}(p; x)$ of order p of the function $G(\cdot; x)$ as follows

$$\boldsymbol{\theta}(p; x) = \inf\{t \in \mathbb{R} : G(t; x) \geq \zeta(p; x)\}. \tag{7}$$

Remark 1. Remark that, for all $t \in \mathbb{R}$, the function $G(\cdot; x)$ is an increasing function.

Remark 2. In the real economical and financial problems, the conditional expectile measure can be considered and interpreted as the threshold that provides a ratio between profit and loss by the quantity $\zeta(p; x)$.

2.2. Nonparametric estimation of the conditional expectile

Let $\{(X_i, Y_i) : i \geq 1\}$ be a sequence of stationary random variables with the same distribution as the couple (X, Y). Let $K(\cdot)$ be a kernel function and $h = h_n$ is a sequence of positive real numbers tending to zero as n tends to infinity and is called the smoothing parameter. The kernel estimator of the function $G(\cdot; x)$ is given by

$$\widehat{G}_{n,h_n}(t;x) = \frac{-\sum_{i=1}^{n} W_{ni}(x)(Y_i - t)\mathbb{1}_{\{(Y_i - t)\leq 0\}}}{\sum_{i=1}^{n} W_{ni}(x)(Y_i - t)\mathbb{1}_{\{(Y_i - t)>0\}}}, \quad \text{for } t \in \mathbb{R}, \tag{8}$$

where

$$W_{ni}(x) = \frac{K_i(x)}{\sum_{i=1}^{n} K_i(x)}, \; K_i(x) = K\left(h^{-1} d(x, X_i)\right).$$

Using the functional kernel local weighted variables, we can write

$$\widehat{G}_{n,h_n}(t;x) = \frac{\widehat{G}_{n,h_n;-}(t;x)}{\widehat{G}_{n,h_n;+}(t;x)}, \tag{9}$$

where

$$\begin{cases} \widehat{G}_{n,h_n;-}(t;x) &= -\frac{1}{n\mathbb{E}\left[K_1(x)\right]} \sum_{i=1}^{n} K_i(x)(Y_i - t)\mathbb{1}_{\{(Y_i - t)\leq 0\}}, \\ \widehat{G}_{n,h_n;+}(t;x) &= \frac{1}{n\mathbb{E}\left[K_1(x)\right]} \sum_{i=1}^{n} K_i(x)(Y_i - t)\mathbb{1}_{\{(Y_i - t)>0\}}. \end{cases}$$

Consequently, we estimate the conditional expectile $\boldsymbol{\theta}(p; x)$ by $\widehat{\boldsymbol{\theta}}_{n,h_n}(p; x)$ that is

$$\widehat{\boldsymbol{\theta}}_{n,h_n}(p;x) = \arg\min_{t\in\mathbb{R}} \widehat{\Psi_p}(t;x), \tag{10}$$

where $\widehat{\Psi_p}(t; x)$ is the estimator of $\Psi_p(t; x)$ defined by

$$\widehat{\Psi_p}(t;x) = \sum_{i=1}^{n} W_{ni}(x)\rho_p(Y_i, t), \text{ for } t \in \mathbb{R},$$

which yields to

$$\widehat{\boldsymbol{\theta}}_{n,h_n}(p;x) = \inf\{t \in \mathbb{R} : \widehat{G}_{n,h_n}(t;x) \geq \zeta(p;x)\}. \tag{11}$$

2.3. Hypotheses and notation

In our analysis, the following conditions are needed to study the asymptotic properties of the estimator $\widehat{\boldsymbol{\theta}}_{n,h_n}(p; x)$. We will denote by C_i, for $i \in \{1, 2, \ldots\}$, some strictly positive constants.

(A1) The function $G(\cdot; x)$ is differentiable in $\mathbb{R}$ and satisfies the following Lipschitz's condition: $\exists a > 0$, $\forall t \in [\boldsymbol{\theta}(p; x) - a, \boldsymbol{\theta}(p; x) + a]$, $\forall x_1, x_2 \in \mathcal{F}$,

$$\left|G^i(t;x_1) - G^i(t;x_2)\right| \leq C_i d^{k_i}(x_1, x_2), \text{ for } k_i > 0, i \in \{1, 2\}.$$

Furthermore, we assume that $\forall x \in \mathcal{F}$, $G^{1'}(\boldsymbol{\theta}(p; x); x)$ and $G^{2'}(\boldsymbol{\theta}(p; x); x)$ are not identical to zero, with $G^{i'}(t; x)$ denoting the derivative of the function $G^i(t; x)$ with respect to t;

(A2) For each $m \geq 2$, $\varphi_m(Y^-) = \mathbb{E}\left[\left|Y^-\right|^m \mid X\right] \leq C_3 < \infty$, $\quad a.s.$; with $Y^- = (Y - t)\,\mathbb{1}_{\{(Y-t)\leq 0\}}$;

(A3) For all $x \in \mathcal{F}$,

$$\mathbb{P}\left(X \in B(x, h_n)\right) = \phi_x(h_n) > 0,$$

where $\phi_x(h_n)$ the small ball probability function, characterizes the classical concentration property (see [24]) and can be written approximately as the product of two independent functions. $B(x, h_n)$ denotes the ball of the center x and radius $h_n > 0$ with

$$\lim_{n\longrightarrow\infty} h_n = 0 \text{ and } \lim_{n\longrightarrow\infty} \phi_x(h_n) = 0;$$

(A4) Assume that

$$\lim_{n\longrightarrow\infty}\frac{n\phi_x(h_n)}{\log n}=\infty;$$

(A5) $K(\cdot)$ is a measurable function with a support $[0,1]$ and there exist two constants, C_4 and C_5 such that $0<C_4<C_5<\infty$, the kernel satisfies

$$0<C_4\leq K(\cdot)\leq C_5<\infty;$$

(B1) The function

$$R(t;x)=\mathbb{E}\left((Y-t)^2\mid X=x\right),$$

is differentiable with respect to t in $\mathbb{R}$ and satisfies the following Lipschitz's condition: $\exists c>0$, $\forall t\in[\boldsymbol{\theta}(p;x)-c,\boldsymbol{\theta}(p;x)+c]$, $\forall x_1,x_2\in\mathcal{F}$,

$$\left|R(t;x_1)-R(t;x_2)\right|\leq C_6 d^r(x_1,x_2),\ \text{for } r>0;$$

(B2) For all $\delta>0$, and $C_7>0$,

$$\mathbb{E}\left[\left|Y-t\right|^{2+\delta}\mid X\right]\leq C_7<\infty,\quad a.s.$$

(B3) The function $K(\cdot)$ satisfies **(A5)** and is differentiable on $(0,1)$ with derivative $K'(\cdot)$ such that there exist two constants, C_8 and C_9 fulfilling $0<C_8<C_9<\infty$, and

$$C_8\mathbb{1}_{(0,1)}(\cdot)\leq K'(\cdot)\leq C_9\mathbb{1}_{(0,1)}(\cdot).$$

Moreover, we assume that $n\phi_x(h_n)\longrightarrow\infty$ and for $\eta>\min(k_1,k_2)$, $nh^{2\eta}\phi_x(h_n)\longrightarrow 0$ as $n\longrightarrow\infty$;

(B4) The concentration property given in **(A3)** holds and there exists a function $\beta_x(\cdot)$ such that

$$\forall s\in[0,1],\ \lim_{t\longrightarrow 0}\frac{\phi_x(st)}{\phi_x(t)}=\beta_x(s).$$

Some comments

- Recall that the preceding conditions are not very restrictive. Our main result is stated under standard conditions of the almost complete consistency in the nonparametric functional statistics. In particular, the structure of the considered assumptions are the same as in the monograph of [24]. They cover the three structural axes of this study (data, model and estimator). Specifically, **(A1)** and **(A2)** allow to explore the dimensionality of the data and the model, respectively.
- The hypothesis **(A1)** is a regularity condition which characterizes the functional space of our model and is needed to evaluate the bias in the asymptotic results of this paper. The hypotheses **(A2)**, **(A4)** and **(A5)** are technical conditions and are, also, similar to those considered in [21].
- The bridge between local weighting and the notion of small ball probabilities is given by the hypothesis **(A3)**, which is a simple normalization of the concentration property of the probability measure on the small balls.
- The hypotheses **(B1)**, **(B2)** and **(B3)** are technical conditions on the functions $R(\cdot,\cdot)$ and $K(\cdot)$, and the hypothesis **(B4)** permits to give the variance term explicitly in the asymptotic normality, and is satisfied for a large family of random functional variables. Indeed, in many examples, the small ball probability function $\phi_x(h_n)$ can be written approximately as the product of two independent functions in terms of x and h, as in the following examples, which can be found in Proposition 1 of [22]:
  1. If $\phi_x(h)\sim Ch^{\upsilon}$ for some $\upsilon>0$ then $\beta_x(s)=s^{\upsilon}$;
  2. If $\phi_x(h)\sim C_1h^{\upsilon}\exp(-C_2h^{-p})$ for some $\upsilon>0$ and $p>0$ then $\beta_x(s)$ is the Dirac's function;
  3. If $\phi_x(h)\sim C\left|\ln(h)\right|^{-1}$ then $\beta_x(s)=\mathbb{1}_{]0,1]}(s)$ the indicator function in $]0,1]$.

3. Main results

In this section, we present our main results concerning the asymptotic properties of the estimator $\widehat{\boldsymbol{\theta}}_{n,h_n}(p;x)$ in both situations: independent and dependent cases. Let us recall the following definition. We say that the sequence $(\boldsymbol{\theta}_n)_n$ converges a.co. to zero, if and only if

$$\forall\tau>0,\ \sum_{n\geq 1}\mathbb{P}(|\boldsymbol{\theta}_n|>\tau)<\infty.$$

Furthermore, we say that $\boldsymbol{\theta}_n=O_{a.co.}(\boldsymbol{\delta}_n)$, if there exists $\tau_0>0$, such that

$$\sum_{n\geq 1}\mathbb{P}(|\boldsymbol{\theta}_n|>\tau_0\boldsymbol{\delta}_n)<\infty.$$

3.1. Independent case

In this paragraph, we suppose that the $\{(X_i, Y_i) : i \geq 1\}$ are i.i.d. The first main result is given by the following proposition and theorem. We establish the almost complete convergence (a.co.) of $\widehat{\boldsymbol{\theta}}_{n,h_n}(p;x)$ to $\boldsymbol{\theta}(p;x)$ in the i.i.d. case.

Proposition 1. *Under the hypotheses **(A1)-(A5)**, we have, for all $t \in \mathbb{R}$, as $n \longrightarrow \infty$,*

$$\left|\widehat{G}_{n,h_n}(t;x) - G(t;x)\right| = O\left(h_n^{k_1}\right) + O\left(h_n^{k_2}\right) + O_{a.co.}\left(\sqrt{\frac{\log n}{n\phi_x(h_n)}}\right). \tag{12}$$

The proof of Proposition 1 is postponed until Section 6.

Theorem 1. *Under the hypotheses of Proposition 1, then, for all $\epsilon > 0$, we obtain*

$$\sum_{n\geq 1} \mathbb{P}\left(\left|\widehat{\boldsymbol{\theta}}_{n,h_n}(p;x) - \boldsymbol{\theta}(p;x)\right| > \epsilon\right) < \infty. \tag{13}$$

The proof of Theorem 1 is postponed until Section 6.

Now, we establish the asymptotic normality of $\widehat{\boldsymbol{\theta}}_{n,h_n}(p;x)$. Recall that the notation $Z \overset{\mathcal{D}}{=} N(\mu,\sigma^2)$ is used whenever the random variable Z follows a normal law with expectation μ and variance σ^2, $\overset{\mathcal{D}}{\longrightarrow}$ denotes the convergence in distribution and $\overset{\mathbb{P}}{\longrightarrow}$ denotes the convergence in probability. Our second main result is presented in the following theorem.

Theorem 2. *Under the hypotheses **(A1)-(A5)-(B1)-(B4)**, we have, as $n \longrightarrow \infty$,*

$$Z_n = \left(\frac{n\phi_x(h_n)}{\sigma_p^2(x)}\right)^{\frac{1}{2}} \left(\widehat{\boldsymbol{\theta}}_{n,h_n}(p;x) - \boldsymbol{\theta}(p;x)\right) \overset{\mathcal{D}}{\longrightarrow} N(0,1), \tag{14}$$

where

$$\sigma_p^2(x) = \frac{\alpha_2(x)\gamma_p\left(\boldsymbol{\theta}(p;x);x\right)}{\alpha_1^2(x)\Lambda_p^2\left(\boldsymbol{\theta}(p;x);x\right)}, \tag{15}$$

$$\alpha_j(x) = K^j(1) - \int_0^1 \left(K^j\right)'(s)\beta_x(s)ds \text{ for } j = 1,2, \tag{16}$$

and

$$\gamma_p\left(\boldsymbol{\theta}(p;x);x\right) = \left(\frac{p}{1-p}\right)^2 R_+^x\left(\boldsymbol{\theta}(p;x)\right) + R_-^x\left(\boldsymbol{\theta}(p;x)\right),$$

where

$$R_+\left(\boldsymbol{\theta}(p;x);x\right) = \mathbb{E}\left((Y_1 - \boldsymbol{\theta}(p;x))^2\, \mathbb{1}_{\{Y_1 > \boldsymbol{\theta}(p;x)\}} \mid X = x\right), R_-\left(\boldsymbol{\theta}(p;x);x\right) = \mathbb{E}\left((Y_1 - \boldsymbol{\theta}(p;x))^2\, \mathbb{1}_{\{Y_1 \leq \boldsymbol{\theta}(p;x)\}} \mid X = x\right),$$

and

$$\Lambda_p\left(\boldsymbol{\theta}(p;x);x\right) = G^{1'}\left(\boldsymbol{\theta}(p;x);x\right) - \left(\frac{p}{1-p}\right) G^{2'}\left(\boldsymbol{\theta}(p;x);x\right).$$

The proof of Theorem 2 is postponed until Section 6.

Remark 3.

- It is clear that the convergence rate of the Proposition 1 is usual in nonparametric functional data analysis. In the sense that the nonparametric path of the model is explored in the bias term. While the functional nature of the model is explored in the variance term through the small ball probability function $\phi_x(h_n)$.
- Recall that the result of Theorem 1 can be viewed as generalisation of Theorem 6.1 in [24] on the regression operator, by the fact that the expectile regression of order $p = 0.5$ reduces to the classical regression $\mathbb{E}[Y \mid X = x]$.
- Once again the asymptotic normality in Theorem 2 generalizes the asymptotic normality of the regression operator obtained in Theorem 2 of [22]. It should be noted that the asymptotic normality is of particular interest of some various asymptotic studies. In particular, it can be used as preliminary result to prove the mean quadratic consistency or to prove the uniform integrability of the estimator, one can refer for more details and discussion to [13].

3.2. Dependent case

Often statistical observations are not independent but are not far from being independent. If not taken into account, dependence can have disastrous effects on statistical inference. The notion of mixing quantifies how close to independence a sequence of random variables is, and it can help us to extend classical results for independent sequences to weakly dependent or mixing sequences. Let us specify the dependence that we will consider in this section.

Definition 1. We call the stationary $\{(X_t, Y_t)\}$ strong mixing if

$$\alpha(j) = \sup_{A\in\widehat{F}_{-\infty}^{0}, B\in\widehat{F}_{j}^{-\infty}} |\mathbb{P}(A\cap B) - \mathbb{P}(A)\mathbb{P}(B)| = 0 \ \ as \ \ j\to\infty,$$

where $\widehat{F}_s^t$ denotes σ-algebra generated by $\{(X_t, Y_t), s \le i \le t\}$. We use the term geometrically strong mixing if: $\alpha(j) \le aj^{-\beta}$ for some $a > 0$ and $\beta > 1$, and exponentially strong mixing if: $\alpha(k) \le b\gamma^k$ for some $b > 0$ and $0 < \gamma < 1$.

The interested reader may refer to [49, 52]. [24] provided some properties of this kind of functional process. To obtain the consistency in probability of $\widehat{\boldsymbol{\theta}}_{n,h_n}(p; x)$, in addition to the assumptions **(A1)-(A3)**, **(A5)**, we need to the following conditions in our analysis.

(C1) The sequence $\{(X_i, Y_i) : i \ge 1\}$ satisfies: $\exists a > 2, \exists c > 0 : \ \forall n \in \mathbb{N}, \ \alpha(n) \le cn^{-a}$ and

$$\begin{cases} \mathbb{E}\left[Y_iY_j \mid X_i, X_j\right] \le C < \infty, & \forall i \ne j, \\ \mathbb{P}\left((X_i, X_j) \in B(x, r)\times B(x, r)\right) = O\left(\phi_x^{1+\epsilon_1}(r)\right) > 0, \quad \text{for some} \quad \epsilon_1 \in (2a^{-1}, 1]. \end{cases}$$

(C2) Assume that

$$\exists \epsilon_2 \in ((a\epsilon_1 - 1)^{-1}, 1), \quad \phi_x(h_n) = O(n^{-\epsilon_2}).$$

The main result of this section to be proved here may now be stated precisely as follows.

Theorem 3. *Under the hypotheses **(A1)-(A3)**, **(A5)** and **(C1)-(C2)**, we have, for all $\epsilon > 0$, as $n \longrightarrow \infty$,*

$$\mathbb{P}\left(\left|\widehat{\boldsymbol{\theta}}_{n,h_n}(p; x) - \boldsymbol{\theta}(p; x)\right| > \epsilon\right) \longrightarrow 0.$$

The proof of this results based on the same ideas of Theorem 1. The proof of Theorem 3 is postponed until Section 6.

Remark 4. Let's point out that Theorem 3 states the asymptotic property of the estimator $\widehat{\boldsymbol{\theta}}_{n,h_n}(p; x)$ in functional time series case using the strong mixing condition. The same result can be obtained using the geometric strong mixing condition for which the coefficient $\alpha(n)$ fulfils the following specific condition

$$\exists C > 0, \exists t \in (0, 1), \quad \alpha(n) \le Ct^n.$$

3.3. Application to confidence intervals

The purpose of a confidence interval is to supplement the functional estimate at a point with information about the uncertainty in this estimate. It is a direct application of the central limit theorem (CLT). In order to provide a confidence interval for the conditional expectile measure, we need first to propose a consistent estimator of the variance given in equation (15). A natural consistent estimator of $\sigma_p^2(x)$ is then

$$\widehat{\sigma}_p^2(x) = \frac{\widehat{\alpha_2}(x)\widehat{\gamma}_p\left(\widehat{\boldsymbol{\theta}}_{n,h_n}(p; x); x\right)}{\widehat{\alpha_1}^2(x)\widehat{\Lambda}_p^2\left(\widehat{\boldsymbol{\theta}}_{n,h_n}(p; x); x\right)}, \tag{17}$$

where $(\widehat{\alpha_i})_{i=1,2}$ are the empirical estimators of the coefficients $(\alpha_i)_{i=1,2}$ given in equation (15). Similarly to [22], $(\widehat{\alpha_i})_{i=1,2}$ are expressed by

$$\widehat{\alpha_1}(x) = \frac{1}{n\phi_x(h_n)}\sum_{i=1}^{n} K_i(x), \quad \widehat{\alpha_2}(x) = \frac{1}{n\phi_x(h_n)}\sum_{i=1}^{n} K_i^2(x).$$

We have now to give estimates of $\gamma_p(\cdot,\cdot)$ and $\Lambda_p(\cdot,\cdot)$ given in (17). Since $G^1(\cdot; x)$, $G^2(\cdot; x)$ and $\boldsymbol{\theta}(p; x)$ are estimated by $\widehat{G}_{n,h_n;-}(\cdot; x)$, $\widehat{G}_{n,h_n;+}(\cdot; x)$ and $\widehat{\boldsymbol{\theta}}_{n,h_n}(p; x)$ respectively, then $\Lambda_p(\cdot,\cdot)$ can be consistently estimated by

$$\widehat{\Lambda}_p\left(\widehat{\boldsymbol{\theta}}_{n,h_n}(p; x); x\right) = \widehat{G}'_{n,h_n;-}(\widehat{\boldsymbol{\theta}}_{n,h_n}(p; x); x) - \left(\frac{p}{1-p}\right)\widehat{G}'_{n,h_n;+}(\widehat{\boldsymbol{\theta}}_{n,h_n}(p; x); x). \tag{18}$$

For $\gamma_p(\cdot,\cdot)$, we use the Nadaraya-Waston type estimator to estimate $R_-(\cdot;x)$ and $R_+(\cdot;x)$ by $\widetilde{R}(\cdot;x)_-$ and $\widetilde{R}_+(\cdot;x)$ respectively, then we obtain a consistent estimator of $\gamma_p(\cdot,\cdot)$:

$$\widehat{\gamma}_p\left(\widehat{\boldsymbol{\theta}}_{n,h_n}(p;x);x\right)=\left(\frac{p}{1-p}\right)^2\widetilde{R}_+\left(\widehat{\boldsymbol{\theta}}_{n,h_n}(p;x);x\right)+\widetilde{R}_-\left(\widehat{\boldsymbol{\theta}}_{n,h_n}(p;x);x\right), \tag{19}$$

with the estimators of $R_-(\cdot;x)$ and $R_+(\cdot;x)$ are given, respectively, by

$$\widetilde{R}_-\left(\widehat{\boldsymbol{\theta}}_{n,h_n}(p;x);x\right) = \frac{\sum_{i=1}^{n}\left(Y_i-\widehat{\boldsymbol{\theta}}_{n,h_n}(p;x)\right)^2K_i(x)\mathbb{1}_{\{(Y_i-\widehat{\boldsymbol{\theta}}_{n,h_n}(p;x))\leq 0\}}}{\sum_{i=1}^{n}K_i(x)}, \tag{20}$$

$$\widetilde{R}_+\left(\widehat{\boldsymbol{\theta}}_{n,h_n}(p;x);x\right) = \frac{\sum_{i=1}^{n}\left(Y_i-\widehat{\boldsymbol{\theta}}_{n,h_n}(p;x)\right)^2K_i(x)\mathbb{1}_{\{(Y_i-\widehat{\boldsymbol{\theta}}_{n,h_n}(p;x))> 0\}}}{\sum_{i=1}^{n}K_i(x)}. \tag{21}$$

Consequently, the equations (18), (19), (20) and (21) allow to construct a consistent estimator of $\sigma_p^2(x)$. On the other hand, an application of Theorem 2 gives

$$\mathbb{P}\left[-U_{1-\frac{\alpha}{2}}\leq\left(\frac{n\phi_x(h_n)}{\sigma_p^2(x)}\right)^{\frac{1}{2}}\left(\widehat{\boldsymbol{\theta}}_{n,h_n}(p;x)-\boldsymbol{\theta}(p;x)\right)\leq U_{1-\frac{\alpha}{2}}\right]=1-\alpha, \tag{22}$$

where $(1-\alpha)$ is the confidence level and $U_{1-\frac{\alpha}{2}}$ the quantile of order $(1-\frac{\alpha}{2})$ of the standard normal law. Thereby, with a confidence level $(1-\alpha)$, the confidence interval is provided by

$$\widehat{\boldsymbol{\theta}}_{n,h_n}(p;x)\pm U_{1-\frac{\alpha}{2}}\left(\frac{\sigma_p^2(x)}{n\phi_x(h_n)}\right)^{\frac{1}{2}}.$$

Of course in practice, we replace $\sigma_p^2(x)$ by its estimator $\widehat{\sigma}_p^2(x)$, defined by

$$\widehat{\sigma}_p^2(x)=n\phi_x(h_n)\frac{\sum_{i=1}^{n}K_i^2(x)\widehat{\gamma}_p\left(\widehat{\boldsymbol{\theta}}_{n,h_n}(p;x);x\right)}{\left(\sum_{i=1}^{n}K_i(x)\right)^2\widehat{\Lambda}_p^2\left(\widehat{\boldsymbol{\theta}}_{n,h_n}(p;x),x\right)},$$

and we obtain the following confidence interval

$$\widehat{\boldsymbol{\theta}}_{n,h_n}(p;x)\pm U_{1-\frac{\alpha}{2}}\left(\frac{\sum_{i=1}^{n}K_i^2(x)\widehat{\gamma}_p\left(\widehat{\boldsymbol{\theta}}_{n,h_n}(p;x);x\right)}{\left(\sum_{i=1}^{n}K_i(x)\right)^2\widehat{\Lambda}_p^2\left(\widehat{\boldsymbol{\theta}}_{n,h_n}(p;x),x\right)}\right)^{\frac{1}{2}}. \tag{23}$$

4. On the potential impact of the functional expectile regression in NFDA

It is well documented that the expectile model (conditional or unconditional) has many theoretical and practical merits. From the theoretical point of view, the expectile regression can be interpreted as the quantilizeation of the conditional expectation. In the sense that the expectiles play the same role as the quantiles by using the tail expectations rather than the tail probabilities. From the empirical point of view, the use of the expectation criterion is more informative than the probability distribution (characterized by the frequency of the data), because the expectation is based on both (values of the data and its frequencies). Similarly to the conditional median the conditional expectation coincides with the expectile regression of order $p = 0.5$. In this context, the hypotheses and asymptotic results are identifiable to the results of [24]. On the other hand, the expectile regression functions are used to construct an alternative estimators for both known risk measures such as Conditional Value at Risk (CVaR) or the Conditional Expected Shortfall (CES), refer to [51]. Thus we can say that the present contribution is of great importance in nonparametric functional statistics. It can be used as an alternative model for several statistical issues such as the discriminant analysis, the prediction problems or the risk analysis.

4.1. Supervised classification and prediction : expectile regression versus classical regression

The supervised curves classification is one of the fundamental problems in the functional data analysis. Recall that the classical regression is a basic tool to solve this issue. As suggested by [24], the functional version of Nadarya-Watson estimator provides an efficient rule of the supervised classification problem. Precisely, from a n-sample $(X_i, Y_i)_i$ where $Y_i = \mathbb{1}_{C_i}$, C_i being a given class, we estimate $\mathbb{P}(C \mid X) = \mathbb{E}[Y \mid X]$ to decide on the affectation of X to the class C. Specifically, the new curve $\mathcal{X}$ is affected to class C if $\mathbb{P}(C \mid X = \mathcal{X}) > \mathbb{P}(\bar{C} \mid X = \mathcal{X})$, where $\bar{C}$ is the complementary set of C. Based on the same reasoning, the expectile regression provides an alternative classifier. Indeed, for a binary random variable $Y = \mathbb{1}_C$, the expectile regression $\boldsymbol{\theta}(p; x)$ of order p is explicitly defined by

$$\boldsymbol{\theta}(p; x) = \frac{p\mathbb{P}(Y = 1 \mid X = x)}{p\mathbb{P}(Y = 1 \mid X = x) + (1 - p)\mathbb{P}(Y = 0 \mid X = x)}. \tag{24}$$

Therefore, our new classification rule affects the new curve $\mathcal{X}$ to the class C when

$$\widehat{\boldsymbol{\theta}}_{n,h_n}(p; \mathcal{X}) \geq p.$$

Such procedure constitutes a generalization of the regression approach given by [24]. In particular, the mean regression can be obtained by taking $p = 0.5$. In addition, the main feature of our new classification procedure is the possibility to take into account the cost of the misclassification. More precisely, assume the cost of the misclassification of C is ϱ_1 against the misclassification of $\bar{C}$ is costed by ϱ_2, the expectile classifier associated to $p = \frac{\varrho_1}{\varrho_1+\varrho_2}$ is more appropriate than the classical regression.
On the other hand, as a generalisation of the regression operator, the expectile regression function plays also an important role in the prediction issue. Recall that the nonparametric functional prediction problem consists to determine a real characteristic Y for a given curve $\mathcal{X}$. For this purpose, we consider the functionals random variables $(X_i(t))_{t\in\mathbb{R}}$ $i = 1, \ldots, n$. For each curve $(X_i(t))_{t\in\mathbb{R}}$, we know its required real characteristic Y_i. Thus, the random variable $\widehat{\boldsymbol{\theta}}_{n,h_n}(p; x_{new})$, defined by (24), is the best approximation of Y, given x_{new} with respect to the loss function $\rho_p(\cdot, \cdot)$ given in (2).

4.2. Financial risk analysis : conditional expectile versus conditional quantile

The second attractive domain of the expectile function is the financial risk management. Such issue is crucial for bankers as well as the investors. At this stage the CVaR or the CES, refer to [51], are is the most common models for financial risk analysis. However, both risk measures tools have some drawbacks. For instance, the CVaR is not coherent. On the other hand, the CES measure is not elicitabile models. In this context, the expectile regression is a good alternative which allows to overcome the limitations of the last cited risk measures. In particular, recall that the expectile regression is the unique risk model that is coherent and elicitabile. Recall that the coherence regroup four properties such as invariant by translation, homogeneity, monotonicity and subadditivity. Each property is instrumental for the investment processing. For instance, the VaR model fails to satisfy the subadditivity. This property guaranties that the risk of a financial portfolio with two combined positions is smaller than the sum of the total risk associated with either position. Thus the VaR model cannot identify the risk measure of diversified financial portfolio. On the other hand, the elicitabile property is also primordial of risk analysis. In particular, the elicitability means that the risk model is obtained with respect to some appropriate criterion. Therefore, the elicitabile property allows to validate the estimation procedure of the risk model by using the backtesting procedure. So, the non elicitability of the expected shortfall exhibits its practical limitations. Because there is non credible rule to check the validity of this risk measure. Thus we can say there is an important gain by using the expectile regression compared to the conditional quantile. Moreover, the standards models CVaR or CES can be alternatively computed by using the expectile regression. In particular, the right-tail $\mathrm{CES}_x(\alpha)$, is defined by

$$\mathrm{CES}_x(\alpha) = \Phi(\alpha, \boldsymbol{\theta}(p; x)) = \left(\frac{1 - p}{(1 - 2p)(1 - \alpha)}\right)\boldsymbol{\theta}(.5; x) - \left(\left(\frac{\alpha(1 - 2p) + p}{(1 - 2p)(1 - \alpha)}\right)\boldsymbol{\theta}(p; x)\right), \tag{25}$$

where $p \in (0, 1)$ such that

$$\boldsymbol{\theta}(p; x) = \mathrm{VaR}_\alpha(x), \tag{26}$$

where $\mathrm{VaR}_\alpha(x)$ is the conditional quantile of Y given $X = x$ of order $\alpha \in (0, 1)$. Noting that the one-to-one correspondence between the expectiles and the quantiles of a related distribution has been proved by [31]. For a recent comparison between quantile and expectile regressions and references, see [53].

Finally let's point out that the expectile regression can be improved as a risk measure by introducing its own expected shortfall. Similarly to the CES, it is defined as the conditional expectation of the financial position given that it exceeds the expectile. The statistical estimation of this model and the identification of its real impact in risk management are fundamental future tracks of the present study. In the same insight, the originality of this study is also motivated by the numerous open questions which can be studied in the future.

4.3. Simulation study

The main aim of this section is to show the easy implementation of the constructed estimator in practice. Of course, the practical use of this estimator is closely related to the flexibility of the selection method of the different parameters involved in the estimator. As all kernel smoothing in nonparametric functional data analysis, the influence of kernel $K(\cdot)$ is less important compared to the others parameters and the choice of the metric $d(\cdot,\cdot)$ depends on the shape of the curves $(X_i)_i$ and its regularity assumption (order of drivability, continuity, discontinuity, ...). Thus the challenge in this kind of estimator is the choice of the bandwidth parameter h_n. Unfortunately, despite the pivotal role of this parameter in practice, its choice stays insufficiently explored from theoretical point of view. It is well known that only two popular rules are studied in functional statistics such as the mean quadratic cross-validation rule by [47] and the Bayesian approach by [50]. However, from the practical point of view, [24] use three cross-validation rules based on the minimization of the mean square prediction error. They evaluate three predicator models that are the conditional expectation, the conditional median and the conditional mode. In this simulation paragraph, we explore the main feature of the expectile regression, that is its multi-connection with classical regression and conditional quantile, to examine the feasibility of the selector procedures, proposed by [24], on the expectile regression estimator. More precisely, we conduct a simulation study to examine the efficiency of $\widehat{\boldsymbol{\theta}}_{n,h_n}(p;x)$ under the two following cross-validation rules

$$\begin{cases} \arg\min\limits_{h\in H_n} \sum\limits_{i=1}^{n} \left(Y_i - \widehat{\boldsymbol{\theta}}_h(0.5;X_i)\right)^2 & \text{RegCV rule,} \\ \arg\min\limits_{h\in H_n} \sum\limits_{i=1}^{n} \left(Y_i - \widehat{\boldsymbol{q}}_h(0.5;X_i)\right)^2 & \text{MedCV rule,} \end{cases}$$

where $\widehat{\boldsymbol{q}}_h(0.5;\cdot)$ is the conditional median estimator and H_n given subset of positive real numbers. For this aim, we generate a functional input-variables using the following formula

$$X(t) = at^3 + bt^2 + ct + e, \quad a, c \backsim N(0, 0.2), \quad b, e \backsim N(0, 1). \tag{27}$$

To take an idea on the suitable metric $d(\cdot,\cdot)$ on this functional variables, we show the shape of the curves in the following Figure 1. In view of the discontinuity of the curves, we can say that the PCA-metric (see [24]) is adequate for this functional explanatory variables. Concerning the output variables Y, we generate $(Y_i)_i$ by the following regression function

$$Y = \sqrt{\int_{-3}^{3} \log(X^2(t)+1)} + \epsilon, \text{ where } \epsilon \text{ is white noise independent of } X.$$

Now, in order to check the behavior of $\widehat{\boldsymbol{\theta}}_{n,h_n}(p;x)$ in both cases (independent and dependent case), we simulate with two white noises. In the first situation, we suppose that $\epsilon \backsim N(0, 0.5)$, and in the second case, we generate it from the ARCH(1) model with parameters (0.01.0.9). For the computational task, we use the quadratic kernel $K(\cdot)$ on $(0, 1)$ and the PCA-metric associated to the third eigenvalue and we optimize the rules *MedCV* and (*RegCV*) using the local k-nearest neighborhood approach for which

$$H_n = \left\{ a \geq 0 : \sum_{i=1}^{n} \mathbb{1}_{B(x,a)}(X_i) = k \right\},$$

where $k \in \{5, 15, 25, \ldots, 0.5n\}$. Finally, the efficiency of the expectile regression is evaluated by using the following backtesting criterion

$$Ex(p) \quad = \quad \frac{1}{n} \sum_{i=1}^{n} \rho_p \left(Y_i, \boldsymbol{\theta}(p;X_i)\right), \tag{28}$$

where we recall that $\rho_p(\cdot,\cdot)$ is given in (2). The comparison results are reported in the Table 1 where we summarize the values of $Ex(p)$ for various sample size $n = 50, 150, 250$ and for three values of $p = 0.1, 0.05, 0.01$.

The obtained results confirm the compatibility of the rules *MedCV* and *RegCV* as bandwidth selector procedures on the expectile regression estimation with slight advantage to the *RegCV*-rule. Moreover, even if, without surprising, the simulation results are affected by the correlation of the data the $Ex(p)$- values shows that $\widehat{\boldsymbol{\theta}}_{n,h_n}(p;x)$ has good performance in the dependence restriction. Finally, we can see that, the computational results incorporate the asymptotic theory about the consistency of the estimator $\widehat{\boldsymbol{\theta}}_{n,h_n}(p;x)$. In sense that the efficiency of $\widehat{\boldsymbol{\theta}}_{n,h_n}(p;x)$ increases with respect to the sample size n and it has satisfactory performance for moderate sample size when $n = 50$.

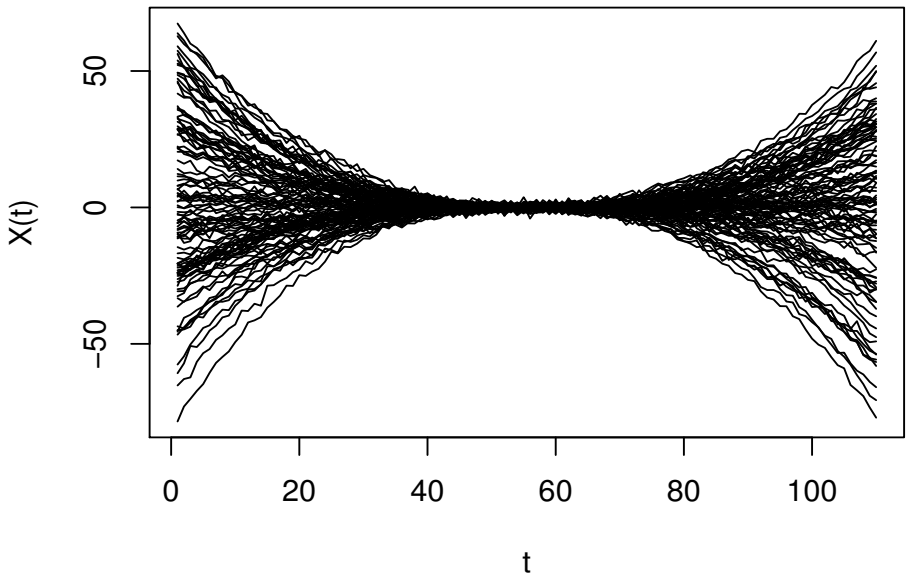

Fig. 1: The curves $X_i(t)$, $i \in \{1, \ldots, 100\}$, $t_j \in [-3, 3]$, $j \in \{1, \ldots 110\}$ as is defined in (27) .

4.4. Data example

In addition to the above discussions on the real impact of the expectile regression compared to its competitive models (quantile and regression), we will examine the easy implementation of this model in practice. In particular, we focus on the risk analysis area and we compare between the quantile and the expectile regressions. Although both models have the same role in the risk management, the expectile regression is more easier to compute and to interpret. Indeed, firstly the smoothing property of the score function associated with the expectile regression allows to increase the precision of its estimation and reduces its computational time cost. Secondly, in the financial risk area, the expectile function can be viewed as a threshold that measures a ratio between profit and loss which makes its interpretation as risk measure more easier than the quantile regression. So, in order to quantify this gain in practice, we conduct an empirical analysis based on the daily returns of $r(t)$ of 20 international financial stock indexes. The data is available in the website "https://fred.stlouisfed.org/categories/32255/downloaddata". It provides data during the period 01/ January/ 2010 to 31/December/2018. From this data, we consider 250 separate months independently chosen from different indexes. Of course, the analyzed functional data keeps all principal features of financial data that are volatility, excess kurtosis, skewness and non-normality (see, the Fig. 2).

Formally, in this real data study, we suppose that the curve is $X(\cdot)$ is the monthly curve of the time series $Z(t) = -100 \log\left(\frac{r(t)}{r(t-1)}\right)$ and the real response variable is $Y = Z(t_l + 1)$, where t_l is the last day of the month. As all statistical modeling, the efficiency of the two approaches is closely linked to the choice of the different parameters involved in the definition of the estimators. However, to conduct a fair comparison between both risk measures, we use the same rule on the choice of the fundamental parameters in the estimators. Precisely, concerning the conditional quantile we use the routine R-code "*funopare.quantile.lcv* " provided by [24] which is available in the website "http://www.lsp.ups-tlse.fr/staph/npfda ". More precisely, we point out that we have used this routine with a quadratic kernel on $(0, 1)$ and the PCA-metric. While the optimal bandwidth was chosen by splitting the data into subsets (Learning and testing samples $(200 + 50)$ and using the 200-observations of the learning to select the best smoothing parameter by the *RegCV*-rule For sake of brevity, we have just considered the most popular method, that is, the cross-validated selected bandwidth. This may be extended to any other bandwidth selector such the bandwidth based on Bayesian ideas [50]. As previously mentioned, the expectile regression estimator is computed by using the same arguments. Finally, the efficiency of the

Table 1: The $Ex(\cdot)$ values (see, equation (28) for different cross-validation rules, white-noise scenarios, sample sizes and orders of quantile).

Cross-validation rule	White-noise type	n	$p = 0.1$	$p = 0.05$	$p = 0.01$
MedCV	IID	50	1.58	1.37	1.26
		150	1.22	1.14	1.02
		250	1.10	1.05	0.93
MedCV	ARCH	50	1.72	1.64	1.51
		150	1.56	1.33	1.12
		250	1.41	1.24	1.08
RegCV	IID	50	1.15	1.06	0.89
		150	0.82	0.74	0.67
		250	0.70	0.61	0.54
RegCV	ARCH	50	1.66	1.43	1.27
		150	1.30	1.16	1.03
		250	1.02	0.96	0.88

two approaches is evaluated by using the following backtesting criterions

$$E(\boldsymbol{\theta}) = \frac{1}{50}\sum_{i=201}^{250}\rho_p\left(Y_i - \boldsymbol{\theta}(p;X_i)\right), \tag{29}$$

$$Q(\boldsymbol{\theta}) = \frac{1}{50}\sum_{i=201}^{250}L_p\left(Y_i - \boldsymbol{\theta}(p;X_i)\right), \tag{30}$$

where

$$L_p(t) = (2p-1)(Y-t) + |Y-t|,$$

and $\boldsymbol{\theta}(p;X_i)$ means quantile or expectile regression. Such errors are evaluated for a fixed value of $p = 0.05$ and for both estimators conditional expectile or conditional quantile. The proposed criterions are performed over 100 replications, of each case, we interchange the observations between the learning and the testing parts. The obtained errors are displayed in the following box-plot figures.

It appears clearly that the expectile regression has slight superiority in terms of precision compared to the quantile regression. Specifically, they are equivalent for the $Q(\cdot)$, in Fig. 3, backtesting measure whereas the $E(\cdot)$, in Fig. 4, measure gives a significative superiority of the expectile regression. Undoubtedly, this gain is due to the ease to determine an accurate estimator of the expectile regression compared to the last absolute error regression. However, we mention that both methods fit with satisfactory performance the volatility of the process.

5. Conclusion and perspectives

A natural perspective of this work is an extension to the multidimensional framework where $\mathbf{Y} \in \mathbb{R}^d$. On the basis of the multidimensional expectile in the paper of [43], let us introduce the conditional multidimensional expectile: Let $\|\cdot\|$ be a norm on $\mathbb{R}^d$. We denote by $(\mathbf{Y}_1)_+$ the vector $(\mathbf{Y}_1)_+ = ((Y_1)_+, \ldots, (Y_d)_+)^\top$ and by $(\mathbf{Y})_-$ the vector $(\mathbf{Y})_- = (\mathbf{Y}_1)_-, \ldots, (Y_d)_-)^\top$. We define the following scoring function

$$s_\alpha(\mathbf{Y},\mathbf{y}) = \alpha \parallel (\mathbf{Y}-\mathbf{y})_+ \parallel^2 + (1-\alpha) \parallel (\mathbf{y}-\mathbf{Y})_+ \parallel^2,$$

for all $\mathbf{y} \in \mathbb{R}^d$. We call a *multivariate conditional expectile* any minimizer

$$\mathbf{y}^* \in \arg\min_{\mathbf{y}\in\mathbb{R}^d} \mathbb{E}[s_\alpha(\mathbf{Y},\mathbf{y}) \mid X = x].$$

It will be of interest to consider this extension in a future investigation. In addition to this issue of multivariate response, another challenging task would be to study the problem of the bandwidth selection in the expectile regression. Such question is primordial for a better use of this estimator in practice. Recall that, in our computational study the optimal smoothing parameter h_n was chosen by some cross-validation methods over a discrete set of bandwidth values defined by the k-Nearest Neighbors (kNN). Even if these selector procedures give satisfactory results by employing the kNN-approach, they lack a mathematical support to highlight the conditions of their optimality and to ensure their asymptotic outcomes. At this stage, the uniform consistency in number of

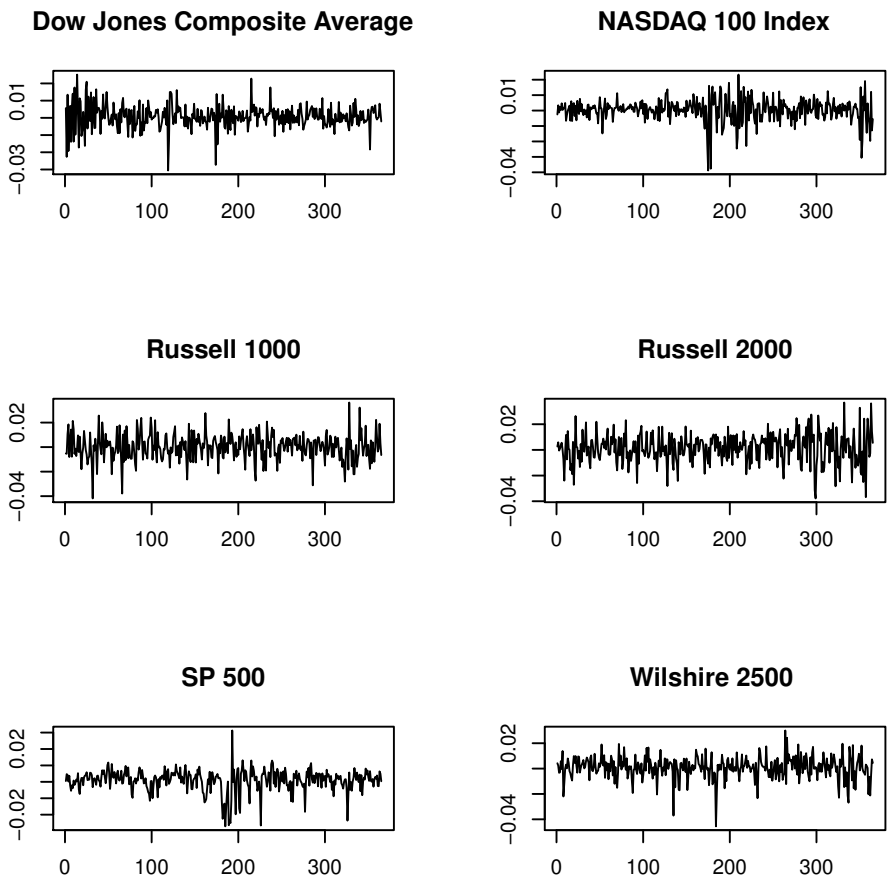

Fig. 2: The daily values of $Z(t) = -100 \log\left(\frac{r(t)}{r(t-1)}\right)$ for six stock index including Dow Jones Composite Average, NASDAQ 100 Index, Russell (1000 and 2000), SP 500 and Wilshire 2500.

neighborhoods, recently developed in functional statistics by [32, 40] may be a best way to overcome this issue. Moreover, it souled by noted that one of the advantages of the expectile regression is its possibility to behave with various statistical strategies (parametric, semiparametric or nonparametric modeling) in the same insight as in the regression and the conditional quantile. Thus, we can say that the linear or semiparametric version of the functional expectile is also a natural prospect of the present contribution. A particular attention will be paid to the semiparametric approach which is actually the trend in functional data analysis (see [7, 45] for more discussions and motivations in this topic). Among the numerous open questions that can be concluded from the present contribution, we mention the construction of pointwise confidence intervals by combining the bootstrap ideas used by [46] on the conditional expectation to those used by [27] on the quantile regression. Finally, let's point out that it would be interested to extend of our results to other functional time series cases (ergodic, long memory, associated process) which requires nontrivial mathematics, that goes well beyond the scope of the present paper.

6. Demonstration of asymptotic results

This section is devoted to the proofs of our results. The aforementioned notation is also used in what follows.

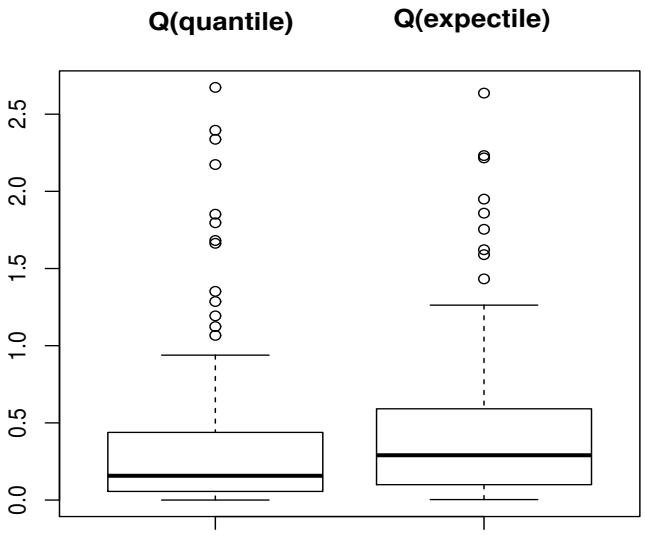

Fig. 3: Comparison between the quantile and the expectile using box-plot of the Q-backtesting measure defined in (30).

Proof of Theorem 1. For $t \in \mathbb{R}$ and by the fact that $\widehat{G}_{n,h_n}$ is an increasing function, for any $\epsilon > 0$ small enough, we have

$$\begin{aligned}\left\{\left|\widehat{\boldsymbol{\theta}}_{n,h_n}(p;x) - \boldsymbol{\theta}(p;x)\right| > \epsilon\right\} &\Rightarrow \left\{\widehat{\boldsymbol{\theta}}_{n,h_n}(p;x) < \boldsymbol{\theta}(p;x) - \epsilon\right\} \cup \left\{\widehat{\boldsymbol{\theta}}_{n,h_n}(p;x) > \boldsymbol{\theta}(p;x) + \epsilon\right\} \\ &\Rightarrow \left\{\widehat{G}_{n,h_n}(\widehat{\boldsymbol{\theta}}_{n,h_n}(p;x);x) \le \widehat{G}_{n,h_n}(\boldsymbol{\theta}(p;x) - \epsilon;x)\right\} \cup \left\{\widehat{G}_{n,h_n}(\widehat{\boldsymbol{\theta}}_{n,h_n}(p;x);x) \ge \widehat{G}_{n,h_n}(\boldsymbol{\theta}(p;x) + \epsilon;x)\right\}.\end{aligned}$$

Since $G(\boldsymbol{\theta}(p;x);x) = \widehat{G}_{n,h_n}(\widehat{\boldsymbol{\theta}}_{n,h_n}(p;x);x)$, we obtain

$$\Rightarrow \left\{\begin{array}{lcl} G(\boldsymbol{\theta}(p;x);x) - G(\boldsymbol{\theta}(p;x) - \epsilon;x) & \le & \widehat{G}_{n,h_n}(\boldsymbol{\theta} - \epsilon) - G(\boldsymbol{\theta}(p;x) - \epsilon;x) \\ G(\boldsymbol{\theta}(p;x);x) - G(\boldsymbol{\theta}(p;x) + \epsilon;x) & \ge & \widehat{G}_{n,h_n}(\boldsymbol{\theta} + \epsilon) - G(\boldsymbol{\theta}(p;x) + \epsilon;x). \end{array}\right.$$

Hence, we infer that

$$\begin{aligned}\left\{\left|\widehat{\boldsymbol{\theta}}_{n,h_n} - \boldsymbol{\theta}\right| > \epsilon\right\} \Rightarrow &\left\{\widehat{G}_{n,h_n}(\boldsymbol{\theta} - \epsilon) - G(\boldsymbol{\theta}(p;x) - \epsilon;x) \ge G(\boldsymbol{\theta}(p;x);x) - G(\boldsymbol{\theta}(p;x) - \epsilon;x)\right\} \\ &\cup \left\{\widehat{G}_{n,h_n}(\boldsymbol{\theta} + \epsilon) - G(\boldsymbol{\theta}(p;x) + \epsilon;x) \le G(\boldsymbol{\theta}(p;x);x) - G(\boldsymbol{\theta}(p;x) + \epsilon;x)\right\}.\end{aligned}$$

This implies that we have

$$\begin{aligned}\left\{\left|\widehat{\boldsymbol{\theta}}_{n,h_n} - \boldsymbol{\theta}\right| > \epsilon\right\} \Rightarrow &\left\{\left|\widehat{G}_{n,h_n}(\boldsymbol{\theta} - \epsilon) - G(\boldsymbol{\theta}(p;x) - \epsilon;x)\right| \ge \left|G(\boldsymbol{\theta}(p;x);x) - G(\boldsymbol{\theta}(p;x) - \epsilon;x)\right|\right\} \\ &\cup \left\{\left|\widehat{G}_{n,h_n}(\boldsymbol{\theta} + \epsilon) - G(\boldsymbol{\theta}(p;x) + \epsilon;x)\right| \ge \left|G(\boldsymbol{\theta}(p;x) + \epsilon;x) - G^x(\boldsymbol{\theta})\right|\right\}.\end{aligned}$$

From the last equation, we obtain that

$$\begin{aligned}\sum_{n\ge1} \mathbb{P}\left(\left|\widehat{\boldsymbol{\theta}}_{n,h_n}(p;x) - \boldsymbol{\theta}(p;x)\right| > \epsilon\right) \le &\sum_{n\ge1} \mathbb{P}\left(\left|\widehat{G}_{n,h_n}(\boldsymbol{\theta}(p;x) - \epsilon;x) - G(\boldsymbol{\theta}(p;x) - \epsilon;x)\right| \ge \left|G(\boldsymbol{\theta}(p;x);x) - G(\boldsymbol{\theta}(p;x) - \epsilon;x)\right|\right) \\ &+ \sum_{n\ge1} \mathbb{P}\left(\left|\widehat{G}_{n,h_n}(\boldsymbol{\theta}(p;x) + \epsilon;x) - G(\boldsymbol{\theta}(p;x) + \epsilon;x)\right| \ge \left|G(\boldsymbol{\theta}(p;x) + \epsilon;x) - G(\boldsymbol{\theta}(p;x);x)\right|\right).\end{aligned}$$

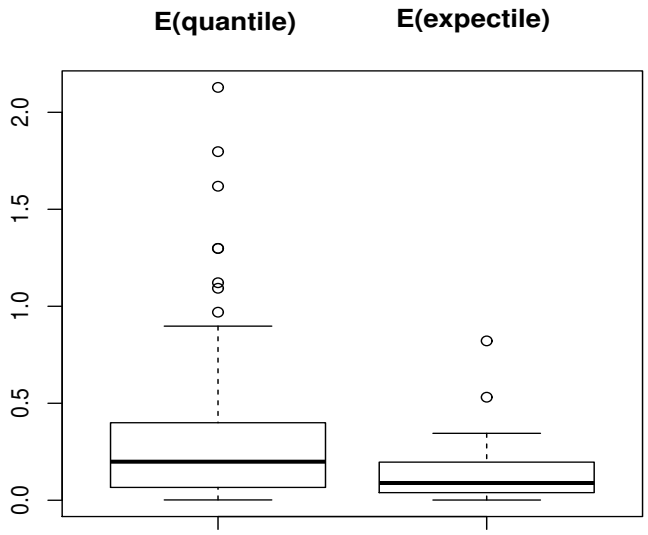

Fig. 4: Comparison between the quantile regression and the expectile regression using box-plot of the E-backtesting measure defined in (30).

Hence, by using the Proposition 1, for all $t \in \mathbb{R}$ and $\tau > 0$ we have,

$$\sum_{n\geq 1} \mathbb{P}\left(\left|\widehat{G}_{n,h_n}(t;x) - G(t;x)\right| > \tau\right) < \infty. \tag{31}$$

Then, the proof of Theorem 1 is a direct consequence of the Proposition 1. □

Proof of Proposition 1. We recall the definitions given in (6) and (9). We first decompose $\widehat{G}_{n,h_n}(t;x) - G(t;x)$ into the sum of two components, as follows:

$$\widehat{G}_{n,h_n}(t;x) - G(t;x) = \frac{\widehat{G}_{n,h_n,-}(t)}{\widehat{G}_{n,h_n,+}(t)} - \frac{G^1(t;x)}{G^2(t;x)} = \frac{1}{\widehat{G}_{n,h_n,+}(t)}\left[\widehat{G}_{n,h_n,-}(t) - G^1(t;x)\right] + \frac{G(t;x)}{\widehat{G}_{n,h_n,+}(t)}\left[G^2(t;x) - \widehat{G}_{n,h_n,+}(t)\right]. \tag{32}$$

Lemmas 1–5 below are given in order to prove Proposition 1. To obtain the consistency results, one usually writes the difference between the estimator and the quantity to be estimated as the sum of a stochastic term and the so-called bias. The mathematical program is to determine the size of the stochastic term and to control the bias separately in the following Lemmas.

Lemma 1. *Assume that the hypotheses **(A1)**(for $i = 1$), **(A2)**, **(A3)**, **(A4)** and **(A5)** are satisfied. We have, as $n \to \infty$,*

$$\left|\widehat{G}_{n,h_n;-}(t;x) - \mathbb{E}\left[\widehat{G}_{n,h_n;-}(t;x)\right]\right| = O_{a.co.}\left(\sqrt{\frac{\log n}{n\phi_x(h_n)}}\right). \tag{33}$$

Lemma 2. *Assume that the hypotheses **(A1)**(for $i = 1$), **(A3)** and **(A5)** are fulfilled. We have as $n \to \infty$,*

$$\left|\mathbb{E}\left[\widehat{G}_{n,h_n;-}(t;x)\right] - G^1(t;x)\right| = O\left(h_n^{k_1}\right). \tag{34}$$

Lemma 3. *Assume that the hypotheses **(A1)**(for $i = 1$), **(A2)**, **(A3)**, **(A4)** and **(A5)** are satisfied. We have, as $n \to \infty$,*

$$\left|\widehat{G}_{n,h_n;+}(t;x) - \mathbb{E}\left[\widehat{G}_{n,h_n;+}(t;x)\right]\right| = O_{a.co.}\left(\sqrt{\frac{\log n}{n\phi_x(h_n)}}\right). \tag{35}$$

Lemma 4. *Assume that the hypotheses **(A1)**(for $i = 2$), **(A3)** and **(A5)** are satisfied. We have, as $n \to \infty$,*

$$\left|\mathbb{E}\left[\widehat{G}_{n,h_n;+}(t;x)\right] - G^2(t;x)\right| = O\left(h_n^{k_2}\right). \tag{36}$$

Lemma 5. *Assume that the hypotheses **(A1)**(for $i = 1$), **(A2)**, **(A3)**, **(A4)** and **(A5)** are satisfied. We have, as $n \to \infty$,*

$$\sum_{n\geq 1} \mathbb{P}\left(\left|\widehat{G}_{n,h_n;+}(t;x)\right| \leq \epsilon\right) < \infty. \tag{37}$$

In order to demonstrate the consistency of $\widehat{\boldsymbol{\theta}}_{n,h_n}(p;x)$, it suffices to show that the estimator $\widehat{G}_{n,h_n;+}(t;x)$ converge almost completely, as $n \to \infty$. The proofs being quite lengthy, we limit ourselves to the main arguments. Hence the proof of the preceding Lemmas 3 and 5 will be omitted.

Proof of Lemma 1. We will use the notation $Y_i^- = (Y_i - t)\mathbb{1}_{\{(Y_i - t)\leq 0\}}$. We have

$$\begin{aligned}
\left|\widehat{G}_{n,h_n;-}(t;x) - \mathbb{E}\left[\widehat{G}_{n,h_n;-}(t;x)\right]\right| &= \left|-\frac{1}{n\mathbb{E}\left[K_1(x)\right]}\sum_{i=1}^{n} K_i(x)Y_i^- - \mathbb{E}\left[-\frac{1}{n\mathbb{E}\left[K_1(x)\right]}\sum_{i=1}^{n} K_i(x)Y_i^-\right]\right| \\
&= \frac{1}{n}\left\{\left|\frac{1}{\mathbb{E}\left[K_1(x)\right]}\sum_{i=1}^{n}\left(K_i(x)Y_i^- - \mathbb{E}\left[K_i(x)Y_i^-\right]\right)\right|\right\} = \frac{1}{n}\left(\left|\sum_{i=1}^{n}\Delta_i\right|\right),
\end{aligned}$$

where we have used the notation $\Delta_i = \frac{1}{\mathbb{E}\left[K_1(x)\right]}\left(K_i(x)Y_i^- - \mathbb{E}\left[K_i(x)Y_i^-\right]\right)$. In order to control the stochastic term, we will make use of the Bernstein's exponential inequality for which the main ingredient is to evaluate asymptotically the m-th order moment of Δ_i:

$$\begin{aligned}
\mathbb{E}\left(\left|\frac{1}{\mathbb{E}\left[K_1(x)\right]}K_1Y_1^-\right|^m\right) &= \frac{1}{\left|\mathbb{E}\left[K_1(x)\right]\right|^m}\mathbb{E}\left[\left|K_1\right|^m\left|Y_1^-\right|^m\right] = \frac{1}{\left|\mathbb{E}\left[K_1(x)\right]\right|^m}\mathbb{E}\left[\mathbb{E}\left[\left|Y_1^-\right|^m \mid X = x\right]K_1^m\right] \\
&= \frac{1}{\left|\mathbb{E}\left[K_1(x)\right]\right|^m}\mathbb{E}\left[\varphi_m(Y_1^-)K_1^m\right] = \mathbb{E}\left[\varphi_m(Y_1^-)(\widehat{K}_1)^m\right],
\end{aligned}$$

where

$$\widehat{K}_i(\cdot) = \frac{K_i(\cdot)}{\mathbb{E}\left[K_i(\cdot)\right]} \text{ and recall that } \varphi_m(Y) = \mathbb{E}\left[\left|Y\right|^m \mid X = x\right].$$

By using the assumptions **(A2)**, **(A3)** and **(A5)**, we get

$$\mathbb{E}\left|\left(\widehat{K}_1Y_1^-\right)\right|^m = \mathbb{E}\left|\left(\frac{1}{\mathbb{E}\left[K_1(x)\right]}K_1Y_1^-\right)\right|^m \leq C_3(\phi_x(h_n))^{-m+1}.$$

This readily implies that

$$\mathbb{E}\left|\left(\widehat{K}_1Y_1^-\right)\right|^m = O\left((\phi_x(h_n))^{-m+1}\right). \tag{38}$$

By the use of the Newton's binomial formula, we obtain that

$$\begin{aligned}
\mathbb{E}\left|\left(Y_1^-\widehat{K}_1 - \mathbb{E}\left[Y_1^-\widehat{K}_1\right]\right)\right|^m &= \mathbb{E}\left|\sum_{l=0}^{m} C_l^m\left(Y_1^-\widehat{K}_1\right)^l\left(\mathbb{E}\left[Y_1^-\widehat{K}_1\right]\right)^{m-l}(-1)^{m-l}\right| \leq \sum_{l=0}^{m} C_l^m\mathbb{E}\left|Y_1^-\widehat{K}_1\right|^l\left|\mathbb{E}\left[Y_1^-\widehat{K}_1\right]\right|^{m-l} \\
&\leq \sum_{l=0}^{m} C_l^m\mathbb{E}\left|Y_1^-\widehat{K}_1\right|^l\left|\mathbb{E}\left[\mathbb{E}\left[Y_1^-\widehat{K}_1\right] \mid X_1\right]\right|^{m-l} \leq C\sum_{l=0}^{m} C_l^m\left(\phi_x(h_n)\right)^{-l+1}\left|G_1^{X_1}(t)\right|^{m-l} \\
&\leq C\max_{0\leq l\leq m}\left(\phi_x(h_n)\right)^{-l+1} = C\left(\phi_x(h_n)\right)^{-m+1}.
\end{aligned}$$

We therefore obtain that

$$\mathbb{E}\left|\left(Y_1^-\widehat{K}_1 - \mathbb{E}\left(Y_1^-\widehat{K}_1\right)\right)\right|^m = O\left((\phi_x(h_n))^{-m+1}\right). \tag{39}$$

Thus, to achieve this proof, it suffices to use the classical Bernstein's inequality (see Corollary A8 in [24]) with

$$a_n^2 = (\phi_x(h_n))^{-1},$$

and by using **(A4)** we obtain, for all $\tau > 0$, that

$$\mathbb{P}\left(\left|\widehat{G}_{n,h_n;-}(t;x) - \mathbb{E}\left[\widehat{G}_{n,h_n;-}(t;x)\right]\right| > \tau\sqrt{\frac{\log n}{n\phi_x(h_n)}}\right) = \mathbb{P}\left(\left|\sum_{i=1}^{n}\Delta_i\right| > n\tau\sqrt{\frac{\log n}{n\phi_x(h_n)}}\right) \leq 2\exp\left(-\frac{1}{2}\vartheta_n\right),$$

where

$$\vartheta_n = \frac{\tau^2\log n}{1+\tau\sqrt{\frac{\log n}{n\phi_x(h_n)}}},$$

and

$$\sum_{n\geq 1}\mathbb{P}\left(\left|\widehat{G}_{n,h_n;-}(t;x) - \mathbb{E}\left[\widehat{G}_{n,h_n;-}(t;x)\right]\right| > \tau\sqrt{\frac{\log n}{n\phi_x(h_n)}}\right) \leq 2\sum_{n\geq 1}\exp\left(-C\tau^2\log n\right).$$

Thereby, for a strictly positive constant C, we choose $\tau > \frac{1}{\sqrt{C}}$ to obtain that

$$\sum_{n\geq 1}\exp\left(-C\tau^2\log n\right) < \infty.$$

This, implies that

$$\left|\widehat{G}_{n,h_n;-}(t;x) - \mathbb{E}\left[\widehat{G}_{n,h_n;-}(t;x)\right]\right| = O_{a.co.}\left(\sqrt{\frac{\log n}{n\phi_x(h_n)}}\right). \tag{40}$$

Hence the proof is complete. □

Proof of Lemma 2. Notice that we have

$$\begin{aligned}
\mathbb{E}\left[\widehat{G}_{n,h_n;-}(t;x)\right] - G^1(t;x) &= \mathbb{E}\left[-\frac{1}{n\mathbb{E}\left[K_1(x)\right]}\sum_{i=1}^{n}K_i(x)Y_i^-\right] - \left(-\mathbb{E}\left[Y^- \mid X = x\right]\right) = \mathbb{E}\left[Y^- \mid X = x\right] - \frac{1}{\mathbb{E}\left[K_1(x)\right]}\mathbb{E}\left[K_1(x)Y_1^-\right] \\
&= \frac{1}{\mathbb{E}\left[K_1(x)\right]}\left\{\mathbb{E}\left[K_1\mathbb{E}\left[Y^- \mid X = x\right]\right] - \mathbb{E}\left[\mathbb{E}\left[K_1Y_1^- \mid X_1\right]\right]\right\} \\
&= \frac{1}{\mathbb{E}\left[K_1(x)\right]}\left\{\mathbb{E}\left[K_1(x)\mathbb{E}\left[Y^- \mid X = x\right]\right] - \mathbb{E}\left[Y_1^- \mid X_1\right]\right\} \\
&= \frac{1}{\mathbb{E}\left[K_1(x)\right]}\left\{\mathbb{E}\left[K_1(x)\left(G_1^{X_1}(t) - G^1(t;x)\right)\right]\right\},
\end{aligned}$$

where $Y_i^- = (Y_i - t)\mathbb{1}_{\{(Y_i - t)\leq 0\}}$. We have

$$\begin{aligned}
\mathbb{E}\left[K_1(x)\left(G^1(t;X_1) - G^1(t;x)\right)\right] &= \mathbb{E}\left(K\left(\frac{d(x,X_1)}{h_n}\right)\left(G^1(t;X_1) - G^1(t;x)\right)\right) = \int K\left(\frac{d(x,z)}{h_n}\right)\left(G^1(t;z) - G^1(t;x)\right)d\mathbb{P}^{X_1} \\
&= \int_{d(x,z)\leq h_n} K\left(\frac{d(x,z)}{h_n}\right)\left(G^1(t;z) - G^1(t;x)\right)d\mathbb{P}^{X_1} = \int_{B(x,h_n)} K\left(\frac{d(x,z)}{h_n}\right)\left(G^1(t;z) - G^1(t;x)\right)d\mathbb{P}^{X_1} \\
&= \mathbb{E}\left[K_1(x)\left(G^1(t;X_1) - G^1(t;x)\right)\mathbb{1}_{B(x,h_n)}\right].
\end{aligned}$$

By using the condition **(A1)**(for $i = 1$), we obtain that

$$\left|\mathbb{E}\left[\widehat{G}_{n,h_n;-}(t;x)\right] - G^1(t;x)\right| = \frac{1}{\mathbb{E}\left[K_1(x)\right]}\left|\mathbb{E}\left[K_1(x)\left(G^1(t;X_1) - G^1(t;x)\right)\mathbb{1}_{B(x,h_n)}\right]\right| \leq C_1\mathbb{1}_{B(x,h_n)}d^{k_1}(x,X_1) \leq C_1h_n^{k_1}.$$

This readily implies that

$$\left|\mathbb{E}\left[\widehat{G}_{n,h_n;-}(t;x)\right] - G^1(t;x)\right| = O\left(h_n^{k_1}\right). \tag{41}$$

By combining the results (41) and (40), we establish the almost complete convergence of $\widehat{G}_{n,h_n,-}(t;x)$ to $G^1(t;x)$. Thus, for all $\epsilon > 0$, we have

$$\sum_{n\geq 1}\mathbb{P}\left(\left|\widehat{G}_{n,h_n;-}(t;x) - G^1(t;x)\right| > \epsilon\right) < \infty. \tag{42}$$

Finally, by Lemma 2 and 1, we get

$$\left|\widehat{G}_{n,h_n;-}(t;x) - G^1(t;x)\right| = O\left(h_n^{k_1}\right) + O_{a.co.}\left(\sqrt{\frac{\log n}{n\phi_x(h_n)}}\right). \tag{43}$$

Hence the proof is complete. □

Proof of Lemma 4. We observe that we have

$$\mathbb{E}\left[\widehat{G}_{n,h_n;+}(t;x)\right] - G^2(t;x) = \mathbb{E}\left[\frac{1}{n\mathbb{E}\left[K_1(x)\right]}\sum_{i=1}^{n} K_i(x)Y_i^+\right] - \mathbb{E}\left[Y^+ \mid X = x\right] = \frac{1}{\mathbb{E}\left[K_1(x)\right]}\left\{\mathbb{E}\left[K_1(x)\left(G_2^{X_1}(t) - G^2(t;x)\right)\right]\right\},$$

where

$$Y_i^+ = (Y_i - t)\mathbb{1}_{\{(Y_i - t) > 0\}}.$$

Then, under assumption **(A1)**(ii), we get:

$$\left|\mathbb{E}\left[\widehat{G}_{n,h_n;+}(t;x)\right] - G^2(t;x)\right| = O\left(h_n^{k_2}\right). \tag{44}$$

By combining the results (35) and (44), we obtain the almost complete convergence of $\widehat{G}_{n,h_n,+}(t;x)$ to $G^2(t;x)$. Thus for all $\epsilon > 0$, we have

$$\sum_{n\geq 1}\mathbb{P}\left(\left|\widehat{G}_{n,h_n;+}(t;x) - G^2(t;x)\right| > \epsilon\right) < \infty. \tag{45}$$

Finally, the application of Lemmas 4 and 3 yields to

$$\left|\widehat{G}_{n,h_n;+}(t;x) - G^2(t;x)\right| = O\left(h_n^{k_2}\right) + O_{a.co.}\left(\sqrt{\frac{\log n}{n\phi_x(h_n)}}\right). \tag{46}$$

Hence the proof is complete. □

Proof of Theorem 2. Let us introduce

$$Z_n = \left(\frac{n\phi_x(h)}{\sigma_p^2(x)}\right)^{\frac{1}{2}}\left(\widehat{\boldsymbol{\theta}}_{n,h_n}(p;x) - \boldsymbol{\theta}(p;x)\right).$$

For $z \in \mathbb{R}$, we set

$$\delta_p(z,x) = \boldsymbol{\theta}(p;x) + z(n\phi_x(h))^{-\frac{1}{2}}\sigma_p(x).$$

We have the following decomposition

$$\begin{aligned}\mathbb{P}\left(Z_n \leq z\right) &= \mathbb{P}\left(\widehat{\boldsymbol{\theta}}_{n,h_n}(p;x) \leq \delta_p(z,x)\right) = \mathbb{P}\left(\left\{\widehat{\boldsymbol{\theta}}_{n,h_n}(p;x) \leq \delta_p(z,x)\right\} \cap \left\{\widehat{G}_{n,h_n;+}\left(\delta_p(z,x)\right) = 0\right\}\right)\\ &+\mathbb{P}\left(\left\{\widehat{\boldsymbol{\theta}}_{n,h_n}(p;x) \leq \delta_p(z,x)\right)\right\} \cap \left\{\widehat{G}_{n,h_n;+}\left(\delta_p(z,x)\right) \neq 0\right\}\right) = I_1(x) + I_2(x).\end{aligned} \tag{47}$$

The term $I_2(x)$ can be written as follows

$$\widehat{\boldsymbol{\theta}}_{n,h_n}(p;x) \leq \delta_p(z,x) \Longleftrightarrow \frac{p}{1-p} \leq \widehat{G}_{n,h_n}\left(\delta_p(z,x);x\right) \Longleftrightarrow \frac{p}{1-p} \leq \frac{\widehat{G}_{n,h_n;-}\left(\delta_p(z,x);x\right)}{\widehat{G}_{n,h_n;+}\left(\delta_p(z,x);x\right)}.$$

Using the fact that the function $\widehat{G}_{n,h_n}(\cdot;x)$ is an increasing function and $\widehat{\boldsymbol{\theta}}_{n,h_n}(p;x)$ is the unique solution of the equation $\widehat{G}_{n,h_n}(\boldsymbol{\theta};x) = \frac{p}{1-p}$, we readily obtain

$$I_2(x) = \mathbb{P}\left(\frac{p}{1-p} \leq \frac{\widehat{G}_{n,h_n;-}\left(\delta_p(z,x);x\right)}{\widehat{G}_{n,h_n;+}\left(\delta_p(z,x);x\right)}\right) = \mathbb{P}\left(\left(\frac{p}{1-p}\right)\widehat{G}_{n,h_n;+}\left(\delta_p(z,x);x\right) - \widehat{G}_{n,h_n;-}\left(\delta_p(z,x);x\right) \leq 0\right).$$

We have

$$I_2(x) = \mathbb{P}\left(\widehat{L}_p(z,x) - \mathbb{E}\left[\widehat{L}_p(z,x)\right] \leq \mathbb{E}\left[-\widehat{L}_p(z,x)\right]\right), \tag{48}$$

where

$$\widehat{L}_p(z,x) = \left(\frac{p}{1-p}\right)\widehat{G}_{n,h_n;+}\left(\delta_p(z,x)\right) - \widehat{G}_{n,h_n;-}\left(\delta_p(z,x)\right).$$

In order to complete the proof of Theorem 2, we need first to prove the following Lemmas. Lemma 6 shows that $I_1(x)$ converges to zero, and the Lemmas 7, 8 show the convergence in distribution of $I_2(x)$ to a standard normal random variable, as n tends to infinity.

Lemma 6. *Assume that the hypotheses **(A3)**, **(A5)** and **(B3)** are satisfied. We have, as $n \to \infty$,*

$$I_1(x) = o(1), \tag{49}$$

where $I_1(x)$ is defined in the equation (47).

Due to the lack of space, the proof of Lemma 6 will be omitted.

Lemma 7. *Assume that the hypotheses of Theorem 2 are satisfied. We have, as $n \to \infty$,*

$$\left(\frac{n\phi_x(h)}{\sigma_p^2(x)}\right)^{\frac{1}{2}} \left(\Lambda_p(\boldsymbol{\theta}(p;x);x)\right)^{-1} \mathbb{E}\left[-\widehat{L}_p(z,x)\right] = z + o(1). \tag{50}$$

Lemma 8. *Suppose that the hypotheses of Theorem 2 hold true. We have, as $n \to \infty$,*

$$(n\phi_x(h))\, \mathbb{V}ar\left(\sum_{i=1}^{n} \mu_i(x)\right) \longrightarrow \frac{\alpha_2(x)\gamma_p(\boldsymbol{\theta}(p;x);x)}{\alpha_1^2(x)}, \tag{51}$$

and

$$Z_n = \left(\frac{n\phi_x(h)}{\sigma_p^2(x)}\right)^{\frac{1}{2}} \left(\Lambda_p(\boldsymbol{\theta}(p;x);x)\right)^{-1} \left(\widehat{L}_p(z,x) - \mathbb{E}\left[\widehat{L}_p(z,x)\right]\right) \xrightarrow{\mathcal{D}} N(0,1), \tag{52}$$

where

$$\mu_i(x) = \frac{1}{n\mathbb{E}\left[K_1(x)\right]} K_i(x)\left(Y_i - \delta_p(z,x)\right)\left(\frac{p}{1-p}\mathbb{1}_{A_i} + \mathbb{1}_{A_i^c}\right),$$

with $A_i = \{Y_i > \delta_p(z,x)\}$ and $A_i^c = \{Y_i \leq \delta_p(z,x)\}$.

Proof of Lemma 7. Keep in mind the definition

$$\widehat{L}_p(z,x) = \left(\frac{p}{1-p}\right)\widehat{G}_{n,h_n;+}\left(\delta_p(z,x)\right) - \widehat{G}_{n,h_n;-}\left(\delta_p(z,x)\right),$$

and

$$G(\boldsymbol{\theta}(p;x);x) = \frac{G^1(\boldsymbol{\theta}(p;x);x)}{G^2(\boldsymbol{\theta}(p;x);x)}.$$

By the fact that $\boldsymbol{\theta}(p;x)$ is the conditional expectile of $G(\cdot;x)$ of order p, i.e.,

$$G(\boldsymbol{\theta}(p;x);x) = \frac{p}{1-p}.$$

We have

$$\begin{aligned}
\mathbb{E}\left[-\widehat{L}_p(z,x)\right] &= \mathbb{E}\left[\widehat{G}_{n,h_n;-}(\delta_p(z,x);x)\right] - \frac{p}{1-p}\mathbb{E}\left[\widehat{G}_{n,h_n;+}(\delta_p(z,x);x)\right] \\
&= \left(\mathbb{E}\left[\widehat{G}_{n,h_n;-}(\delta_p(z,x);x)\right] - G^1(\delta_p(z,x);x)\right) + \frac{p}{1-p}\left(G^2(\delta_p(z,x);x) - \mathbb{E}\left[\widehat{G}_{n,h_n;+}(\delta_p(z,x);x)\right]\right) \\
&\quad + \left(G^1(\delta_p(z,x);x) - G^1(\boldsymbol{\theta}(p;x);x)\right) - \frac{p}{1-p}\left(G^2(\delta_p(z,x);x) - G^2(\boldsymbol{\theta}(p;x);x)\right) + \left(G^1\left(\boldsymbol{\theta}(p;x)\right) - \frac{p}{1-p}G^2\left(\boldsymbol{\theta}(p;x);x\right)\right) \\
&= J_1(x) + J_2(x) + J_3(x) + J_4(x) + J_5(x).
\end{aligned} \tag{53}$$

By the fact that $\boldsymbol{\theta}(p;x)$ is the unique solution of the equation (6), we infer that the term $J_5(x)$, in the last equation, is equal to zero. We next evaluate the term $J_1(x)$ in the right side of (53). Taking into consideration the conditions **(A1)**(for $i = 1$), **(A5)** and **(B3)**, we have the following bound

$$\left|J_1(x)\right| = \left|\mathbb{E}\left[\widehat{G}_{n,h_n;-}(\delta_p(z,x))\right] - G^1(\delta_p(z,x);x)\right| \leq C_1 h_n^{k_1}.$$

Following similar steps for $J_2(x)$ ans using the hypotheses **(A1)**(for $i = 2$), **(A5)** and **(B3)**, we show that

$$\left|J_2(x)\right| = \frac{p}{1-p}\left|\left[\mathbb{E}\left[\widehat{G}_{n,h_n;+}(\delta_p(z,x))\right] - G^2(\delta_p(z,x);x)\right]\right| \leq C_2 h_n^{k_2},$$

which implies, for $i = 1, 2$,

$$\left|J_i(x)\right| = O(h_n^{k_i}).$$

In order to evaluate the terms $J_3(x)$ and $J_4(x)$ in the right side of (53), we start with the Taylor's expansion of the function $G^1(\cdot; x)$ to obtain

$$J_3(x) = G^1\left(\delta_p(z,x); x\right) - G^1\left(\boldsymbol{\theta}(p;x); x\right) = z\sigma_p(x)\left(n\phi_x(h)\right)^{-\frac{1}{2}} G^{1'}(\boldsymbol{\theta}(p;x); x) + o\left(\left(n\phi_x(h)\right)^{-\frac{1}{2}}\right),$$

and the similar arguments for $J_4(x)$ yield to

$$J_4(x) \quad = \quad \left(-\frac{p}{1-p}\right) G^2(\delta_p(z,x); x) - G^2(\boldsymbol{\theta}(p;x); x) = z\sigma_p(x)\left(n\phi_x(h)\right)^{-\frac{1}{2}} \left(\left(-\frac{p}{1-p}\right) G^{2'}(\boldsymbol{\theta}(p;x); x)\right) + o\left(\left(n\phi_x(h)\right)^{-\frac{1}{2}}\right)$$

Then, by setting

$$\Lambda_p\left(\boldsymbol{\theta}(p;x); x\right) = G^{1'}(\boldsymbol{\theta}(p;x); x) - \left(\frac{p}{1-p}\right) G^{2'}(\boldsymbol{\theta}(p;x); x),$$

one can see that

$$\left|\sum_{i=1}^{5} J_i(x)\right| = O(h_n^{\eta}) + z\sigma_p(x)\left(n\phi_x(h)\right)^{-\frac{1}{2}} \left(\Lambda_p(\boldsymbol{\theta}(p;x); x)\right) + o\left(\left(n\phi_x(h)\right)^{-\frac{1}{2}}\right).$$

Under the assumption **(B3)**, for $\eta > \min(k_1, k_2)$, we get the following result

$$\left(\frac{n\phi_x(h)}{\sigma_p^2(x)}\right)^{\frac{1}{2}} \left(\Lambda_p(\boldsymbol{\theta}(p;x); x)\right)^{-1} \left|\mathbb{E}\left(-\widehat{L}_p(z,x)\right)\right| \quad = \quad O\left(nh_n^{2\eta}\phi_x(h)\right)^{\frac{1}{2}} + z + o(1) = z + o(1).$$

Thus the proof is complete. □

Proof of Lemma 8. In what follows, we set

$$Z_n = \left(\frac{n\phi_x(h)}{\sigma_p^2(x)}\right)^{\frac{1}{2}} \left(\Lambda_p(\boldsymbol{\theta}(p;x); x)\right)^{-1} \left(\widehat{L}_p(z,x) - \mathbb{E}\left[\widehat{L}_p(z,x)\right]\right).$$

We can apply the Lyapunov CLT to conclude that, as $n \longrightarrow \infty$,

$$Z_n = \frac{1}{S_n}\sum_{i=1}^{n}\left(\mu_i(x) - \mathbb{E}\left[\mu_i(x)\right]\right) \xrightarrow{\mathcal{D}} N(0,1), \quad S_n = \left(\sum_{i=1}^{n} \mathbb{V}ar(\mu_i(x))\right)^{\frac{1}{2}}. \tag{54}$$

In order to demonstrate that equation (54) holds, we start by computing the asymptotic variance of $\widehat{L}_p(z,x)$. We keep in mind the following equations

$$\begin{aligned}
\widehat{L}_p(z,x) \quad &= \quad \widehat{L}_p(z,x) = \left(\frac{p}{1-p}\right)\widehat{G}_{n,h_n;+}\left(\delta_p(z,x); x\right) - \widehat{G}_{n,h_n;-}\left(\delta_p(z,x); x\right), \\
\delta_p(z,x) \quad &= \quad \delta_p(z,x) = \boldsymbol{\theta}(p;x) + z(n\phi_x(h))^{-\frac{1}{2}}\sigma_p(x).
\end{aligned}$$

We first write $\widehat{L}_p(z,x)$ in the following form

$$\begin{aligned}
\widehat{L}_p(z,x) \quad &= \quad \frac{p}{1-p}\left[\frac{1}{n\mathbb{E}\left[K_1(x)\right]}\sum_{i=1}^{n} K_i(x)\left(Y_i - \delta_p(z,x)\right)\mathbb{1}_{(Y_i-\delta_p(z,x))>0}\right] - \frac{1}{n\mathbb{E}\left[K_1(x)\right]}\left[\sum_{i=1}^{n} K_i(x)\left(Y_i - \delta_p(z,x)\right)\mathbb{1}_{(Y_i-\delta_p(z,x))\leq 0}\right] \\
&= \quad \frac{1}{n\mathbb{E}\left[K_1(x)\right]}\sum_{i=1}^{n} K_i(x)\left(Y_i - \delta_p(z,x)\right)\left[\frac{p}{1-p}\mathbb{1}_{A_i} + \mathbb{1}_{A_i^c}\right] = \sum_{i=1}^{n}\mu_i(x),
\end{aligned}$$

where we recall that $\mu_i(x)$, A_i and A_i^c are defined in Lemma 8. Thus we can simplify the problem of obtaining the asymptotic variance of $\widehat{L}_p(z,x)$ by computing instated the variance of the sum of $\mu_i(x)$. We can therefore write the following equalities

$$\begin{aligned}
(n\phi_x(h))\,\mathbb{V}ar\left(\widehat{L}_p(z,x)\right) \quad &= \quad (n\phi_x(h))\,\mathbb{V}ar\left(\sum_{i=1}^{n}\mu_i(x)\right) = \frac{\phi_x(h)}{\mathbb{E}^2\left[K_1(x)\right]}\mathbb{V}ar\left(K_1(x)\left(Y_1 - \delta_p(z,x)\right)\left(\frac{p}{1-p}\mathbb{1}_{A_1} + \mathbb{1}_{A_1^c}\right)\right) \\
&= \quad \phi_x(h)\frac{\mathbb{E}\left[K_1^2(x)\right]}{\mathbb{E}^2\left[K_1(x)\right]}\mathbb{E}\left[\frac{K_1^2(x)(Y_1 - \delta_p(z,x))^2\left(\frac{p}{1-p}\mathbb{1}_{A_1} + \mathbb{1}_{A_1^c}\right)^2}{\mathbb{E}\left[K_1^2(x)\right]}\right] \\
&\quad -\phi_x(h)\mathbb{E}^2\left[\frac{K_1(x)\left(Y_1 - \delta_p(z,x)\right)\left(\frac{p}{1-p}\mathbb{1}_{A_1} + \mathbb{1}_{A_1^c}\right)}{\mathbb{E}\left[K_1(x)\right]}\right] \\
&= \quad T_0(x)T_1(x) + T_2(x).
\end{aligned} \tag{55}$$

By conditioning with respect to X_1, we readily obtain

$$\begin{aligned}\mathbb{E}\left[K_1(x)\left(Y_1-\delta_p(z,x)\right)\left(\frac{p}{1-p}\mathbb{1}_{A_1}+\mathbb{1}_{A_1^c}\right)\right] &= \mathbb{E}\left[K_1(x)\left\{\left(\frac{p}{1-p}\right)G^2\left(\delta_p(z,x),X_1\right)-G^1(\delta_p(z,x),X_1)\right\}\right]\\ &= \mathbb{E}\left[K_1(x)\left\{\left(\frac{p}{1-p}\right)M_2(x)-M_1(x)\right\}\right],\end{aligned}$$

where, $M_i(x)=G^i(\delta_p(z,x),X_1)-G^i(\boldsymbol{\theta}(p;x);x)$, for $i\in\{1,2\}$. Under the assumption **(A1)**, it follows immediately from the triangle inequality that, for $i\in\{1,2\}$,

$$\mathbb{1}_{d(x,X_1)\leq h}\left|M_i(x)\right|\leq C_i h^{k_i}+z\left(n\phi_x(h)\right)^{-\frac{1}{2}}\sigma_p(x)G^{i'}(\boldsymbol{\theta}(p;x);x).$$

From which we deduce that

$$\mathbb{E}\left[\frac{K_1(x)\left(Y_1-\delta_p(z,x)\right)\left(\frac{p}{1-p}\mathbb{1}_{A_1}+\mathbb{1}_{A_1^c}\right)}{\mathbb{E}\left[K_1(x)\right]}\right]=o(1).$$

This, implies that

$$T_2(x)=o(1). \tag{56}$$

We next evaluate the term $T_1(x)$ in the right side of (55). We set

$$\lambda_p(\delta_p;x)=\left(\frac{p}{1-p}\right)^2 R_+\left(\delta_p(z,x);x\right)+R_-\left(\delta_p(z,x);x\right),$$

and

$$\gamma_p(\boldsymbol{\theta}(p;x);x)=\left(\frac{p}{1-p}\right)^2 R_+\left(\boldsymbol{\theta}(p;x);x\right)+R_-\left(\boldsymbol{\theta}(p;x);x\right).$$

By the same previous arguments used to treat the $T_2(x)$, we have

$$T_1(x) = \mathbb{E}\left[\frac{K_1^2(x)\left(Y_1-\delta_p(z,x)\right)^2\left(\frac{p}{1-p}\mathbb{1}_{A_1}+\mathbb{1}_{A_1^c}\right)^2}{\mathbb{E}\left[K_1^2(x)\right]}\right]=\mathbb{E}\left[\frac{K_1^2(x)\lambda_p\left(\delta_p,X_1\right)}{\mathbb{E}\left[K_1^2(x)\right]}\right]=\mathbb{E}\left[\frac{K_1^2(x)\left\{\left(\frac{p}{1-p}\right)^2 N_2(x)+N_1(x)\right\}}{\mathbb{E}\left[K_1^2(x)\right]}\right]+\gamma_p\left(\boldsymbol{\theta}(p;x);x\right),$$

where

$$N_1(x)=R_-\left(\delta_p(z,x);X_1\right)-R_-\left(\boldsymbol{\theta}(p;x);x\right),\quad N_2(x)=R_+\left(\delta_p(z,x);X_1\right)-R_+(\boldsymbol{\theta}(p;x);x).$$

An application of the triangle inequality implies readily that

$$\begin{aligned}\mathbb{1}_{d(x,X_1)\leq h}\left|N_1(x)\right| &\leq C_1h^{r_1}+z\left(n\phi_x(h)\right)^{-\frac{1}{2}}\sigma_p(x)R'_-(\boldsymbol{\theta}(p;x);x),\\ \mathbb{1}_{d(x,X_1)\leq h}\left|N_2(x)\right| &\leq C_2h^{r_2}+z\left(n\phi_x(h)\right)^{-\frac{1}{2}}\sigma_p(x)R'_+(\boldsymbol{\theta}(p;x);x).\end{aligned}$$

By combining the assumptions **(A3)** and **(B1)**, **(B3)**, we infer that

$$T_1(x)=\gamma_p\left(\boldsymbol{\theta}(p;x);x\right)+o(1). \tag{57}$$

We evaluate the term $T_0(x)$ in the right side of (55). Using similar arguments as in [24], we can show that

$$\mathbb{E}\left[K_1^j(x)\right]=K^j(1)\phi_x(h)-\int_0^1\left((K^j)\right)'(s)\phi_x(sh)ds+o\left(\phi_x(h)\right)\ \text{ for } j\in\{1,2\}. \tag{58}$$

This when combined with (15) gives that, as $n\longrightarrow\infty$,

$$\frac{\phi_x(h)\mathbb{E}\left[K_1^2(x)\right]}{\mathbb{E}^2\left[K_1(x)\right]}\longrightarrow\frac{\alpha_2(x)}{\alpha_1^2(x)}. \tag{59}$$

By combining (56), (57) and (59) we obtain the statement (51). To derive the asymptotic normality, we will show the Lyapunov conditions for the independent variables $\mu_i(x)$. Since the observations are i.i.d. and making use of the C_r-inequality (see [42, p.155]), this implies, for some $\eta>0$, that

$$\left(n\phi_x(h)\right)^{\frac{2+\eta}{2}}\sum_{i=1}^n\mathbb{E}\left[\left|\mu_i(x)-\mathbb{E}\left[\mu_i(x)\right]\right|^{2+\eta}\right]\leq 2^{\eta+1}n\left(n\phi_x(h)\right)^{\frac{2+\eta}{2}}\mathbb{E}\left[\left|\mu_1(x)\right|^{2+\eta}\right]+2^{\eta+1}n\left(n\phi_x(h)\right)^{\frac{2+\eta}{2}}\left|\mathbb{E}\left[\mu_1(x)\right]\right|^{2+\eta}=A_1(x)+A_2(x).$$

Through conditioning with respect to X_1 and using the conditions **(B2)** and **(B3)**, we obtain

$$\begin{aligned} A_1(x) &= 2^{\eta+1}\left(n\phi_x(h)\right)^{\frac{-\eta}{2}}\left(\frac{\phi_x(h)}{\mathbb{E}\left[K_1(x)\right]}\right)^{2+\eta}\times\mathbb{E}\left[\frac{K_1^{2+\eta}(x)}{\phi_x(h)}\Big|Y_1-\delta_p(z,x)\Big|^{2+\eta}\Big|\frac{p}{1-p}\mathbb{1}_A+\mathbb{1}_{A^c}\Big|^{2+\eta}\right] \\ &\leq \left(\max\left(1,\frac{p}{1-p}\right)\right)^{2+\eta}2^{\eta+1}\left(n\phi_x(h)\right)^{\frac{-\eta}{2}}\left(\frac{\phi_x(h)}{\mathbb{E}\left[K_1(x)\right]}\right)^{2+\eta}\times\mathbb{E}\left[\frac{K_1^{2+\eta}(x)}{\phi_x(h)}\mathbb{E}\left(\Big|Y_1-\delta_p(z,x)\Big|^{2+\eta}\mid X_1\right)\right]\leq C\left(n\phi_x(h)\right)^{\frac{-\eta}{2}}. \end{aligned}$$

This, implies that

$$A_1(x)\longrightarrow 0 \text{ as } n\longrightarrow\infty. \tag{60}$$

In order to evaluate $A_2(x)$, we use the same arguments as those involved in the proof of $T_2(x)$. We get

$$\begin{aligned} A_2(x) &= 2^{\eta+1}n\left(n\phi_x(h)\right)^{\frac{2+\eta}{2}}\Big|\mathbb{E}\left[\mu_1(x)\right]\Big|^{2+\eta}=2^{\eta+1}n^{-\frac{\eta}{2}}\left(\phi_x(h)\right)^{\frac{2+\eta}{2}}\left|\mathbb{E}\left[\frac{K_1(x)(Y_1-\delta_p(z,x))(\frac{p}{1-p}\mathbb{1}_A+\mathbb{1}_{A^c})}{\mathbb{E}\left[K_1(x)\right]}\right]\right|^{2+\eta} \\ &= 2^{\eta+1}n^{-\frac{\eta}{2}}\left(\phi_x(h)\right)^{\frac{2+\eta}{2}}\Big|T_2(x)\Big|^{2+\eta}. \end{aligned}$$

By combing the condition (56) with the assumption **(A3)**, we obtain

$$A_2(x)=o(1). \tag{61}$$

Consequently, from (60), (61) and (51), for some $\eta>0$, we infer that, as $n\longrightarrow\infty$,

$$\frac{1}{\left(n\phi_x(h)\right)^{\frac{2+\eta}{2}}S_n^{2+\eta}}\left(n\phi_x(h)\right)^{\frac{2+\eta}{2}}\sum_{i=1}^{n}\mathbb{E}\left[\Big|\mu_i(x)-\mathbb{E}\left[\mu_i(x)\right]\Big|^{2+\eta}\right]\longrightarrow 0. \tag{62}$$

We conclude from (62) that the Lyapunov condition is verified which, implies the convergence in distribution of Z_n to a standard normal $N(0,1)$. We therefore obtain, as $n\longrightarrow\infty$,

$$\left(n\phi_x(h)\right)\mathbb{V}ar\left(\widehat{L}_p(z,x)\right)\longrightarrow\sigma_p^2(x)\Lambda_p^2\left(\boldsymbol{\theta}(p;x);x\right). \tag{63}$$

Combining (51) and (63), we infer

$$\sigma_p^2(x)\Lambda_p^2\left(\boldsymbol{\theta}(p;x);x\right)=\frac{\alpha_2(x)\gamma_p(\boldsymbol{\theta}(p;x);x)}{\alpha_1^2(x)}.$$

Thus by (47) and (49) whith (48), we get

$$\mathbb{P}(Z_n\leq z)\leq o(1)+\mathbb{P}\left(Z_n\leq U_n.\mathbb{E}\left[-\widehat{L}_p(z,x)\right]\right), \tag{64}$$

where

$$U_n=\left(\frac{n\phi_x(h)}{\sigma_p^2(x)}\right)^{\frac{1}{2}}\left(\Lambda_p(\boldsymbol{\theta}(p;x);x)\right)^{-1}.$$

Finally, in order to prove Theorem 2, we combine (64) with the equations (50), (51) and (52). □

Proof of Theorem 3. The proof is based on the same arguments as those used in the proof of Theorem 1. However the bias term is not affected by the dependence assumption. So, it remains to evaluate the asymptotic behaviors of the variance terms of $\widehat{G}_{n,h_n;-}(t;x)$ and $\widehat{G}_{n,h_n;+}(t;x)$ to prove the convergence in probability of the estimator $\widehat{\boldsymbol{\theta}}_{n,h_n}(p;x)$ stated in Theorem 3.

Lemma 9. *Under the hypotheses of the Theorem 3, we have, as $n\to\infty$,*

$$\mathrm{Var}\left[\widehat{G}_{n,h_n;-}(t;x)\right]=O\left(\frac{n}{\phi_x(h_n)}\right), \tag{65}$$

and

$$\mathrm{Var}\left[\widehat{G}_{n,h_n;+}(t;x)\right]=O\left(\frac{n}{\phi_x(h_n)}\right). \tag{66}$$

Proof of Lemma 9. The proof of this lemma is omitted, it follows the same lines of the proof of Lemma 11.5 in [24]. □

7. Appendix

7.1. Demonstration of Remark 1: the function $G(\cdot; x)$ is increasing.

Recall that

$$G(t;x) = \frac{G^1(t;x)}{G^2(t;x)}, \tag{67}$$

where

$$G^1(t;x) = -\mathbb{E}\left[(Y-t)\mathbb{1}_{\{(Y-t)\leq 0\}} \mid X = x\right], G^2(t;x) = \mathbb{E}\left[(Y-t)\mathbb{1}_{\{(Y-t)>0\}} \mid X = x\right].$$

If $t_1 > t_2$, we get, $-(y - t_1) \geq -(y - t_2)$ and

$$\int_{-\infty}^{t_2} -(y-t_2)F_Y^x(y) \quad \leq \quad \int_{-\infty}^{t_2} -(y-t_1)dF_Y^x(y),$$

where $F_Y^x(\cdot)$ is the distribution function of Y conditional on $\{X = x\}$: then

$$\int_{-\infty}^{t_2} -(y-t_2)dF_Y^x(y) \quad \leq \quad \int_{-\infty}^{t_1} -(y-t_1)dF_Y^x(y),$$

which implies that

$$\mathbb{E}\left(-(Y-t_2)\mathbb{1}_{(Y-t_2)\leq 0} \mid X = x\right) \quad \leq \quad \mathbb{E}\left(-(Y-t_1)\mathbb{1}_{(Y-t_1)\leq 0} \mid X = x\right).$$

Then, $G^1(\cdot; x)$ is an increasing function. Similarly, if $t_1 > t_2$ for $G^2(\cdot; x)$

$$\int_{t_2}^{+\infty} (y-t_2)dF_Y^x(y) \quad \geq \quad \int_{t_2}^{+\infty} (y-t_1)dF_Y^x(y) \geq \int_{t_1}^{+\infty} (y-t_1)dF_Y^x(y),$$

This implies that

$$\mathbb{E}\left((Y-t_2)\mathbb{1}_{(Y-t_2)>0} \mid X = x\right) \quad \geq \quad \mathbb{E}\left((Y-t_1)\mathbb{1}_{(Y-t_1)>0} \mid X = x\right).$$

We readily infer that $G^2(\cdot; x)$ is a decreasing function. Therefore we obtain that $G(\cdot; x)$ is an increasing function. □

7.2. Bernstein's inequality.

Let $Z_1, \ldots, Z_n$ be independent real zero mean random variables. If for all $m \geq 2$, there exists a constant $C_m > 0$, such that

$$\mathbb{E}|Z_1^m| \leq C_m a^{2(m-1)},$$

we have for $\epsilon > 0$,

$$\mathbb{P}\left(\left|\sum_{i=1}^{n} Z_i\right| > \epsilon n\right) \leq 2\exp\left(\frac{-\epsilon^2 n}{2a^2(1+\epsilon)}\right). \tag{68}$$

7.3. Lyapunov CLT.

In this variant of the central limit theorem, the random variables X_i have to be independent, but not necessarily identically distributed. The theorem also requires that random variables have moments of order $(2 + \delta)$, $\delta > 0$.

Theorem 4. *Suppose $\{X_1, X_2, \ldots\}$ is a sequence of independent random variables, each with a finite expected value μ_i and variance σ_i^2. Define $s_n^2 = \sum_{i=1}^{n} \sigma_i^2$. If for some $\delta > 0$, the Lyapunov's condition*

$$\lim_{n\longrightarrow\infty} \frac{1}{s_n^{2+\delta}} \sum_{i=1}^{n} \mathbb{E}\left[\mid X_i - \mu_i|^{2+\delta}\right] = 0,$$

is satisfied, then we have, as $n \to \infty$,

$$\frac{1}{s_n}\sum_{i=1}^{n}(X_i - \mu_i) \xrightarrow{\mathcal{D}} N(0,1).$$

Acknowledgement

The authors are indebted to the Editor-in-Chief Professor Dietrich von Rosen, the Associate Editor and the three referees for their very valuable comments and suggestions which led to a considerable improvement of the manuscript. They also extend their appreciation to the Deanship of Scientific Research at King Khalid University for funding this work through research groups program under the project number R.G.P.2/68/41.

References

[1] B. Abdous, K. Ghoudi, B. Rémillard, Nonparametric weighted symmetry tests, Canad. J. Statist. 31 (2003) 357–381.
[2] B. Abdous, B. Rémillard, Relating quantiles and expectiles under weighted-symmetry, Ann. Inst. Statist. Math. 47 (1995) 371–384.
[3] D. J. Aigner, T. Amemiya, D. J. Poirier, On the estimation of production frontiers: maximum likelihood estimation of the parameters of a discontinuous density function, Internat. Econom. Rev. 17 (1976) 377–396.
[4] F. Al-Awadhi, Z. Kaid, A. Laksaci, I. Ouassou, M. Rachdi, Functional data analysis: local linear estimation of the l1-conditional quantiles, Statistical Methods and Applications 28 (2019) 217–240.
[5] G. Aneiros, R. Cao, R. Fraiman, C. Genest, P. Vieu, Recent advances in functional data analysis and high-dimensional statistics, J. Multivariate Anal. 170 (2019) 3–9.
[6] G. Aneiros, R. Cao, P. Vieu, Editorial on the special issue on functional data analysis and related topics, Computational Statistics 34 (2019) 447–450.
[7] G. Aneiros, P. Vieu, Partial linear modelling with multi-functional covariates, Comput. Statist. 30 (2015) 647–671.
[8] P. Artzner, F. Delbaen, J.-M. Eber, D. Heath, Coherent measures of risk, Math. Finance 9 (1999) 203–228.
[9] F. Bellini, E. D. Bernardino, Risk management with expectiles, The European Journal of Finance 23 (2017) 487–506.
[10] F. Bellini, V. Bignozzi, On elicitable risk measures, Quant. Finance 15 (2015) 725–733.
[11] F. Bellini, V. Bignozzi, G. Puccetti, Conditional expectiles, time consistency and mixture convexity properties, Insurance Math. Econom. 82 (2018) 117–123.
[12] F. Bellini, B. Klar, A. Müller, E. Rosazza Gianin, Generalized quantiles as risk measures, Insurance Math. Econom. 54 (2014) 41–48.
[13] C. Crambes, L. Delsol, A. Laksaci, Robust nonparametric estimation for functional data, J. Nonparametr. Stat. 20 (2008) 573–598.
[14] S. Dabo-Niang, A. Laksaci, Nonparametric quantile regression estimation for functional dependent data, Comm. Statist. Theory Methods 41 (2012) 1254–1268.
[15] A. Daouia, I. Gijbels, G. Stupfler, Extremiles: A new perspective on asymmetric least squares, Journal of the American Statistical Association 114 (2019) 1366–1381.
[16] A. Daouia, S. Girard, G. Stupfler, Estimation of tail risk based on extreme expectiles., J. R. Stat. Soc., Ser. B, Stat. Methodol. 80 (2018) 263–292.
[17] B. Efron, Regression percentiles using asymmetric squared error loss, Statist. Sinica 1 (1991) 93–125.
[18] W. Ehm, T. Gneiting, A. Jordan, F. Krüger, Of quantiles and expectiles: consistent scoring functions, choquet representations and forecast rankings, Journal of the Royal Statistical Society: Series B (Statistical Methodology) 78 (2016) 505–562.
[19] S. Emmer, M. Kratz, D. Tasche, What is the best risk measure in practice? a comparison of standard measures., Journal of Risk 18 :2 (2015) 31–60.
[20] M. Ezzahrioui, E. Ould-Saïd, Asymptotic results of a nonparametric conditional quantile estimator for functional time series, Comm. Statist. Theory Methods 37 (2008) 2735–2759.
[21] F. Ferraty, A. Laksaci, A. Tadj, P. Vieu, Rate of uniform consistency for nonparametric estimates with functional variables, J. Statist. Plann. Inference 140 (2010) 335–352.
[22] F. Ferraty, A. Mas, P. Vieu, Nonparametric regression on functional data: inference and practical aspects, Aust. N. Z. J. Stat. 49 (2007) 267–286.
[23] F. Ferraty, A. Rabhi, P. Vieu, Conditional quantiles for dependent functional data with application to the climatic El Niño phenomenon, Sankhyā 67 (2005) 378–398.
[24] F. Ferraty, P. Vieu, Nonparametric functional data analysis, Springer Series in Statistics, Springer, New York, 2006. Theory and practice.
[25] T. Gneiting, Making and evaluating point forecasts, J. Amer. Statist. Assoc. 106 (2011) 746–762.
[26] A. Goia, P. Vieu, An introduction to recent advances in high/infinite dimensional statistics [Editorial], J. Multivariate Anal. 146 (2016) 1–6.
[27] J. Hahn, Bootstrapping quantile regression estimators, Econometric Theory 11 (1995) 105–121.
[28] H. Holzmann, B. Klar, Expectile asymptotics, Electron. J. Stat. 10 (2016) 2355–2371.
[29] L. Horváth, P. Kokoszka, Inference for functional data with applications, Springer Series in Statistics, Springer, New York, 2012.
[30] T. Hsing, R. Eubank, Theoretical foundations of functional data analysis, with an introduction to linear operators, Wiley Series in Probability and Statistics, John Wiley & Sons, Ltd., Chichester, 2015.
[31] M. C. Jones, Expectiles and M-quantiles are quantiles, Statist. Probab. Lett. 20 (1994) 149–153.
[32] L.-Z. Kara, A. Laksaci, M. Rachdi, P. Vieu, Data-driven kNN estimation in nonparametric functional data analysis, J. Multivariate Anal. 153 (2017) 176–188.
[33] R. Koenker, G. Bassett, Jr., Regression quantiles, Econometrica 46 (1978) 33–50.
[34] R. Koenker, G. Bassett, Jr., Robust tests for heteroscedasticity based on regression quantiles, Econometrica 50 (1982) 43–61.
[35] P. Kokoszka, M. Reimherr, Introduction to functional data analysis, Texts in Statistical Science Series, CRC Press, Boca Raton, FL, 2017.
[36] V. Krätschmer, H. Zähle, Statistical inference for expectile-based risk measures., Scand. J. Stat. 44 (2017) 425–454.
[37] C.-M. Kuan, J.-H. Yeh, Y.-C. Hsu, Assessing value at risk with care, the conditional autoregressive expectile models, Journal of Econometrics 150 (2009) 261 – 270. Recent Development in Financial Econometrics.

[38] A. Laksaci, M. Lemdani, E. Ould-Saïd, A generalized L^1-approach for a kernel estimator of conditional quantile with functional regressors: consistency and asymptotic normality, Statist. Probab. Lett. 79 (2009) 1065–1073.
[39] A. Laksaci, M. Lemdani, 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.
[40] N. Ling, S. Meng, P. Vieu, Uniform consistency rate of kNN regression estimation for functional time series data, J. Nonparametr. Stat. 31 (2019) 451–468.
[41] N. Ling, P. Vieu, Nonparametric modelling for functional data: selected survey and tracks for future, Statistics 52 (2018) 934–949.
[42] M. Loève, Probability theory, Third edition, D. Van Nostrand Co., Inc., Princeton, N.J.-Toronto, Ont.-London, 1963.
[43] V. Maume-Deschamps, D. Rullière, K. Said, Multivariate extensions of expectiles risk measures, Depend. Model. 5 (2017) 20–44.
[44] W. K. Newey, J. L. Powell, Asymmetric least squares estimation and testing, Econometrica 55 (1987) 819–847.
[45] S. Novo, G. Aneiros, P. Vieu, Automatic and location-adaptive estimation in functional single-index regression, J. Nonparametr. Stat. 31 (2019) 364–392.
[46] P. Raña, G. Aneiros, J. Vilar, P. Vieu, Bootstrap confidence intervals in functional nonparametric regression under dependence, Electron. J. Stat. 10 (2016) 1973–1999.
[47] M. Rachdi, P. Vieu, Nonparametric regression for functional data: automatic smoothing parameter selection, J. Statist. Plann. Inference 137 (2007) 2784–2801.
[48] J. O. Ramsay, B. W. Silverman, Functional Data Analysis, Second Edition, Springer, New York, NY, 2005.
[49] M. Rosenblatt, A central limit theorem and a strong mixing condition, Proc. Nat. Acad. Sci. U.S.A. 42 (1956) 43–47.
[50] H. L. Shang, Bayesian bandwidth estimation for a functional nonparametric regression model with mixed types of regressors and unknown error density, J. Nonparametr. Stat. 26 (2014) 599–615.
[51] J. W. Taylor, Estimating value at risk and expected shortfall using expectiles, Journal of Financial Econometrics 6 (2008) 231–252.
[52] V. A. Volkonskiĭ, Y. A. Rozanov, Some limit theorems for random functions. I, Teor. Veroyatnost. i Primenen 4 (1959) 186–207.
[53] L. S. Waltrup, F. Sobotka, T. Kneib, G. Kauermann, Expectile and quantile regression—david and goliath?, Statistical Modelling 15 (2015) 433–456.
[54] Q. Yao, H. Tong, Asymmetric least squares regression estimation: a nonparametric approach, J. Nonparametr. Statist. 6 (1996) 273–292.
[55] J. F. Ziegel, Coherence and elicitability., Math. Finance 26 (2016) 901–918.