Pak.j.stat.oper.res. Vol.21 No.4 2025 pp 631-652 DOI: http://dx.doi.org/10.18187/pjsor.v21i4.4910

Pakistan Journal of Statistics and Operation Research

Conditional quantile estimation under LTRC model with functional regressors

Hamida Chemerik[1,2]*, Zohra Guessoum[1,3]

1. Lab. MSTD, Faculty of Mathematics, USTHB, Algeria.
2. Faculty of Sciences, University of Boumerdes, UMBB, Algeria, h.chemerik@univ-boumerdes.dz
3. Dpt of Probability and Statictic, Faculty of Mathematics, USTHB, Algeria, zguessoum@usthb.dz

*. Corresponding author

Abstract

In this paper, we study a kernel-based estimator of the conditional quantile function for a response variable Y subject to left truncation and right censoring (LTRC), with a functional covariate X. We establish the consistency properties with the rate of this estimator when the observations are independent and identically distributed. The performance of the proposed estimator is illustrated through simulations and a real-data application.

Key Words: Conditional quantile, Convergence rate, Functional variable, Kernel estimator, LTRC model.

Mathematical Subject Classification: 60E05, 62E15.

1. Introduction

The conditional quantile, by its robustness, offers a good alternative to the conditional mean when describing the impact of the explanatory variables on the response variable, especially in the case of functional data (curves, functions, tables, etc.). These data are defined as realizations of a random variable taking its values in infinite dimensional space.
Over the last few decades, statistical modeling for functional data has developed considerably, thanks to its numerous applications in various fields such as biology, agronomy, and environmental science. Thus, the contribution of the scientific community rises to a rich statistical literature. We can refer to Bosq (2000) and Ramsay and Silverman (2005) for parametric models, and the monograph of Ferraty and Vieu (2006) for the prediction problems in functional nonparametric statistics. Recent advances in the functional data analysis can be found in Ferraty (2011) and Horvath and Kokoszka (2012).
In this article, we focus on the nonparametric estimation of the conditional quantile of a real-valued random variable Y, given a functional covariate X that takes values in a semi-metric space $\Im$.
For this purpose we consider the conditional distribution function of Y given X

$$F(y|x) = \mathbb{P}(Y \leq y|X = x) = \mathbb{E}[\mathbf{1}_{\{Y\leq y\}}|X = x],$$

and the conditional quantile of order p, $p \in (0, 1)$, defined for $x \in \Im$ by

$$Q_p(x) := \inf\{y, F(y|x) \geq p\}.$$

When $F(.|.)$ is continuous, $Q_p(x)$ satisfies

$$F(Q_p(x)|x) = p.$$

Recall that the most commonly used quantiles are the median ($p = 0.5$), the first and last quartiles ($p = 0.25$, $p = 0.75$), and the first and last deciles ($p = 0.1$, $p = 0.9$).
The nonparametric estimation of conditional quantile is an important topic in statistics, and it is used to construct predictive intervals, determine reference curves, or as a means of prediction. There exists a large literature on the conditional quantile function estimation of a scalar response given a scalar (or multivariate) covariate. Historically, the first result was obtained by Roussas (1969), who proved the convergence and asymptotic normality of kernel estimates of conditional quantile under Markov assumptions. For independent and identically distributed (i.i.d.) random variables, Stone (1977) showed the weak convergence of kernel estimates. Xiang (1996) studied the almost sure convergence and asymptotic normality of a kernel estimator of the conditional quantile.
In the case of complete data, when the response variable is real and the covariate is functional, Ferraty and Vieu (2006) presented some asymptotic results linked with nonparametric kernel estimation of the regression, the conditional median, the conditional mode, and the conditional quantile. Whereas, Farraty et al. (2006) constructed the double kernel estimator of the conditional distribution function and established the almost complete convergence of this estimator when the observations are i.i.d. Their study has been extended to the dependent case by other authors, for example, Ferraty et al. (2005) and Dabo-Niang and Laksaci (2012).
An important characteristic of survival data is the existence of incomplete observations.An important characteristic of survival data is the existence of incomplete observations. Indeed, the data can be partially collected because of censoring and/or truncation mechanisms. Right censoring arises if the individual did not experience the event during his last observation. Left truncation occurs if an individual's event of interest is only observable if its time of occurrence is greater than a certain value. In finite-dimensional space, for censored data, several authors considered the greater problem of conditional quantile estimation, such as Ould Said and Sadki (2008) and Djelladj and Tatachak (2019), etc. For truncated data, we refer to Lemdani et al. (2009) and Adjoudj and Tatachak (2019). Strong representation for conditional quantile estimator with left-truncated and right-censored (LTRC) data has been studied by some authors, notably Iglesias-Pérez (2003), Liang and Miao (2015), and others.
In infinite-dimensional space, under the censoring model, the kernel estimation of the conditional quantile has been considered by Horrigue and Ould Said (2011). They proved the uniform strong convergence with rate in the dependent case, while Kadiri et al. (2018) investigated the kernel estimation of the conditional quantiles when the observations are linked with a single-index relationship. One year later, Leulmi (2019) introduced a new local linear estimator of the conditional quantile and extended the results of Messaci et al. (2015) to censored data.
Under the left truncation model, Helal and Ould Said (2016) studied the kernel conditional quantile estimator and established the almost sure convergence with rate in the i.i.d. case.
In this work, our main objective is to define an estimator of the conditional distribution function $F(y|x)$ and, consequently, an estimator of the conditional quantile function $Q_p(x)$ for the LTRC model, and to study its almost sure convergence with rate under i.i.d. data with functional covariate. In this context, we propose a new estimator of the conditional distribution, which is not simply a combination of the work of Horrigue and Ould Said (2011) for right-censored data and that of Helal and Ould Said (2016) for left-truncated data. Indeed, our estimator accounts for censoring not only in the response variable Y but also in the covariate X (see Remark 3.1).
This paper is structured as follows: in Section 2, we present our estimation model; in Section 3, we define new estimators of the conditional distribution function and the conditional quantile; in Section 4, we give the assumptions on the proposed model under which we state our main results; in Section 5, we conduct a simulation study; and in Section 6, we demonstrate the effectiveness of our method on a real dataset. The proofs of the main results are given in Section 7, while the appendix contains some technical lemmas with their proofs.

2. LTRC model and estimate

Let $\{(Y_i, T_i, C_i), i = 1, 2, ..., N\}$ be a sequence of random vectors from (Y, T, C), where Y is the variable of interest (lifetime) with continuous distribution function (d.f.) F. The variables T and C represent the random left truncation and the random right censoring times with (d.f's) L and G, respectively. In LTRC model, we observe (Z, T, δ) only if $Z \geq T$, where $Z = min(Y, C)$ and $\delta = \mathbf{1}_{\{Y \leq C\}}$ is the censoring indicator. When $Z < T$ nothing is observed. In the sequel, we suppose that Y, T and C are independent each of others, then the (d.f.) V of Z is given by $V = 1-(1-F)(1-G)$. As a consequence of truncation, the size of the actually observed sample $n := \sum_{i=1}^{N} \mathbf{1}_{\{Z_i \geq T_i\}}$ is a binomial $\mathfrak{B}(N, \theta)$ random variable, with $\theta := \mathbb{P}(Z \geq T)$. Note that if $\theta = 0$, nothing is observable; hence, we can suppose that $\theta > 0$. Denote by $\{(Z_i, T_i, \delta_i), i = 1, 2, ..., n; (n \leq N)\}$ the i.i.d observed sample. Without loss of generality, we assume that Y, T and C are nonnegative random variables, as usual in survival analysis.

Here $\mathbb{P}$ denotes the probability related to the N-sample, and $\mathbf{P}$ the conditional probability related to the actually observed n-sample. As well as $\mathbb{E}$ and $\mathbf{E}$ are the expectation operators related to $\mathbb{P}$ and $\mathbf{P}$ respectively. For any df R we denote the left and right endpoints of its support by $a_R = \inf\{y : R(y) > 0\}$ and $b_R = \sup\{y : R(y) < 1\}$ respectively. Obviously $a_V = min(a_F, a_G)$ and $b_V = min(b_F, b_G)$.
For LTRC data, Tsai et al. (1987) proposed a product limit estimator (PLE), referred to as the TJW estimator F_n of the distribution function F, given by

$$F_n(y) = 1 - \prod_{i:Z_i \le y} \left(1 - \frac{1}{nC_n(Z_i)}\right)^{\delta_i},$$

where $C_n(y) = n^{-1} \sum_{i=1}^{n} \mathbf{1}_{\{T_i \le y \le Z_i\}}$ is the empirical estimator of $C(y) := \mathbf{P}\{T \le y \le Z\}$.
Note that F_n reduces to the Kaplan and Meier PLE when there is no left truncation, and to the Lynden-Bell PLE when there is no right censoring. For independent data, the properties of F_n have been studied by several authors. Tsai et al. (1987) treated asymptotic normality of the estimator, Gu and Lai (1990) and Lai and Ying (1991) obtained a functional law of the iterated logarithm and strong approximations results for a slightly modified form of the PLE F_n using martingale theory. Gijbels and Wang (1993) provided a strong i.i.d. approximation for F_n with a remainder term of order $O(n^{-1} \log n)$ almost surely, when $a_L < a_V$. Zhou (1996) established almost sure representation of the estimator and proved that the rates of the remainder terms are of order $O(n^{-1} \log^{1+\varepsilon} n)$, for $\varepsilon > \frac{1}{2}$, when $a_L = a_V$ and under condition $\int_{a_V}^{\infty} \frac{dF}{L^2} < \infty$. In 1999, Zhou and Yip improved this rate to $O(n^{-1} \log \log n)$ almost surely.
Let us write

$$C(y) = \mathbb{P}\{T \le y \le Z | Z \ge T\} = \theta^{-1} L(y)(1 - V(y)) = \theta^{-1} L(y)(1 - F(y))(1 - G(y)).$$

Then

$$\theta = \frac{L(y)(1 - F(y))(1 - G(y))}{C(y)}. \tag{1}$$

According to the idea of He and Yang (1998), the proposed estimator of θ is given by

$$\theta_n = \frac{L_n(y)(1 - F_n(y))(1 - G_n(y))}{C_n(y)}, \tag{2}$$

for all y such that $C_n(y) \neq 0$, where L_n is the Lynden-Bell estimator of the distribution function L defined by

$$L_n(y) = \prod_{T_i > y} \left(1 - \frac{1}{nC_n(T_i)}\right), \tag{3}$$

and $G_n(y)$ is the TJW- type estimator of the distribution function G defined by

$$1 - G_n(y) =: \overline{G}_n(y) = \prod_{Z_i \le y} \left(1 - \frac{1}{nC_n(Z_i)}\right)^{1-\delta_i}. \tag{4}$$

He and Yang (1998) showed that θ_n is independent of y and its value can be obtained for any y such that $C_n(y) \neq 0$.
Throughout this study, we suppose that $a_L < a_V,\ b_L \le b_V$ and

$$\begin{cases} (T, C) \text{ are independent of } Y \\ T \text{ is independent of } C \end{cases} \tag{5}$$

Remark 2.1. *To ensure the identifiability of the model, we have assumed $a_L < a_V$ to alleviate the assumptions and avoid the additional condition of integrability.*
The independence required in condition (5) will be needed to study the bias expressions of some pseudo estimators (see Lemma 9.3 and Lemma 9.5).

In what follows, we use the notations $\mathbf{V}_1^*(y|x)$ and $\mathbf{v}_1^*(y|x)$ to denote the conditional sub-distribution of Z given X

and its derivative, and we have

$$\mathbf{v}_1^*(y|x)d\mathbf{P}^X(x) = \theta^{-1}L(y)\overline{G}(y)f(y|x)d\mathbb{P}^X(x), \tag{6}$$

where $f(y|x)$ is the derivative with respect to y of the conditional distribution function $F(y|x)$.

3. Conditional quantile and distribution estimation

Let (X, Y) be a couple of random variables in $\Im \times \mathbb{R}$, where $\Im$ is a semi-metric space. Denoting by $d(.,.)$ a semi-metric on $\Im$. In the case of complete data, it is well known that an estimator of $F(y|x)$ is defined by Farraty et al. (2006) as

$$\widehat{F}_n(y|x) = \frac{\sum_{i=1}^n K\left(\frac{d(x,X_i)}{h_{n,K}}\right) H\left(\frac{y-Y_i}{h_{n,H}}\right)}{\sum_{i=1}^n K\left(\frac{d(x,X_i)}{h_{n,K}}\right)}$$

where K is a kernel function, H is a distribution function, and $h_{n,K} =: h_K$ (resp. $h_{n,H} =: h_H$) is a sequence of positive real numbers that goes to zero as n goes to infinity.
Under the left truncation model, Helal and Ould Said (2016) introduced a kernel estimator of $F(y|x)$ based on the n actually observed data$(X_i, Y_i, T_i)_{i=1:n}$ in the normed space ($\Im$, $d(.,.)$).

$$\widehat{F}_{n_T}(y|x) = \frac{\sum_{i=1}^n \frac{1}{L_n(Y_i)} K\left(\frac{d(x,X_i)}{h_K}\right) H\left(\frac{y-Y_i}{h_H}\right)}{\sum_{i=1}^n \frac{1}{L_n(Y_i)} K\left(\frac{d(x,X_i)}{h_K}\right)}.$$

Recall that $L_n(.)$ is the Lynden-Bell estimator of $L(.)$ defined in (3).
In the right censoring model, Horrigue and Ould Said (2011) proposed the following estimator of $F(y|x)$

$$\widehat{F}_{n_C}(y|x) = \frac{\sum_{i=1}^n \frac{\delta_i}{\overline{\widehat{G}}_n(Z_i)} K\left(\frac{d(x,X_i)}{h_K}\right) H\left(\frac{y-Z_i}{h_H}\right)}{\sum_{i=1}^n K\left(\frac{d(x,X_i)}{h_K}\right)}, \tag{7}$$

where $\widehat{G}_n(.)$ is the Kaplan and Meier estimator of G given by

$$\widehat{G}_n(.) = 1 - \prod_{i=1}^n \left(1 - \frac{1-\delta_{(i)}}{n-i+1}\right)^{\{Z_{(i)} \leq y\}},$$

with $Z_{(1)}, Z_{(2)}, .., Z_{(n)}$ are the order statistics of $(Z_i)_{1\leq i \leq n}$ and $\delta_{(i)}$ is the concomitant of $Z_{(i)}$.
Combining the proposed estimators of Helal and Ould Said (2016) and Horrigue and Ould Said (2011), we define a new estimator in the LTRC model that considers the effect of censoring and truncation.

$$F_n(y|x) = \frac{\sum_{i=1}^n \frac{\delta_i}{L_n(Z_i)\overline{G}_n(Z_i)} K\left(\frac{d(x,X_i)}{h_K}\right) H\left(\frac{y-Z_i}{h_H}\right)}{\sum_{i=1}^n \frac{\delta_i}{L_n(Z_i)\overline{G}_n(Z_i)} K\left(\frac{d(x,X_i)}{h_K}\right)}$$

where $G_n(.)$ is the TJW estimator of G defined in (4), and the sum is taken only for i such that $L_n(Z_i)\overline{G}_n(Z_i) \neq 0$.
Then, the estimator of $Q_p(.)$ is given by

$$Q_{p,n}(x) = \inf\{y, F_n(y|x) \geq p\}, \tag{8}$$

which satisfies

$$F_n(Q_{p,n}(x)|x) = p.$$

Remark 3.1. *Remark that $F_n(y|x)$ reduces to the estimator $\widehat{F}_n(y|x)$ when there is no right censoring and left truncation and to the estimator $\widehat{F}_{n_T}(y|x)$ when there is no right censoring. In the absence of truncation-meaning under right* *censoring, we obtain*

$$F_{n_C}(y|x) = \frac{\sum_{i=1}^{n} \frac{\delta_i}{\widehat{\overline{G}}_n(Z_i)} K\left(\frac{d(x,X_i)}{h_K}\right) H\left(\frac{y-Z_i}{h_H}\right)}{\sum_{i=1}^{n} \frac{\delta_i}{\widehat{\overline{G}}_n(Z_i)} K\left(\frac{d(x,X_i)}{h_K}\right)}.$$

This last estimator differs from the one proposed by Horrigue and Ould Said (2011) in (7). Indeed, in its current form, the denominator in (7) does not account for the right censoring associated with the covariate X_i. This may seem paradoxical, since the numerator includes only the actually observed Y_i values. However, this is not the case for our estimator.

4. Assumptions

For $x \in \Im$, let $B(x,h) = \{x' \in \Im, d(x,x') \leq h\}$ be the ball of center x and radius h. Let Ω be a compact subset of $\Im$ and $\aleph_x$ be a neighborhood of x. Denote by $[a_V, b]$ a compact set, where b is a positive real number satisfying $b < b_V$. For any continuous function W, set $\|W\|_\infty = \sup_x |W(x)|$. Throughout this paper, $c_1, c_2, .., c_8$ denote some arbitrary positive constants and c, c' some generic constants. Moreover, we state the following assumptions required to formulate the main results.

A1. $\forall x \in \Omega, \forall h > 0$, $0 < c_1\phi(h) \leq \mathbb{P}(X \in B(x,h)) =: \phi_x(h) \leq c_2\phi(h)$, where $\phi(h)$ is a positive real function such that $\lim_{h\to 0} \phi_(h) = 0$.

A2. The kernel K is

- (i) a continuous function with compact support [0,1] and satisfies $c_3 \leq K(.) \leq c_4$;
- (ii) of classe C^1 on $(0,1)$ and $c_5 \leq K'(.) \leq c_6$, where K' denotes the first derivative of K.

A3. The conditional distribution function $F(.|.)$ is such that

- (i) $\exists\beta_1 > 0, \beta_2 > 0, \forall(y_1,y_2) \in \mathbb{R}^2; \forall(x_1,x_2) \in \aleph_x \times \aleph_x$, we have

$$|F(y_1|x_1) - F(y_2|x_2)| \leq c_7(d(x_1,x_2)^{\beta_1} + |y_1 - y_2|^{\beta_2})$$

- (ii) $F(.|.)$ admits a positive and bounded first derivative $f(.|.)$ and there exists $\tau > 0$ such that $f(y|x) > \tau$, for all $(x,y) \in \Omega \times [a_V, b]$;

A4. The distribution function H is Lipschitz continuous function, with bounded derivative H' such that $\int_{\mathbb{R}} |s|^{\beta_2} H'(s)ds < \infty$.

A5. (i) $\lim_{n\to\infty} h_K = 0$ and $\lim_{n\to\infty} n^\beta h_K = +\infty$, for $\beta > 0$;
- (ii) $\lim_{n\to\infty} \frac{\log n}{n\phi(h_K)} = 0$;
- (iii) $\lim_{n\to\infty} h_H = 0$ and $\lim_{n\to\infty} n^\eta h_H = +\infty$, for $\eta > 0$.

Comments on the Assumptions

Assumption **A1** concerns the concentration property of the functional variable X. This condition is used in several existing works dealing with this kind of variable. In this case, the functional variable does not have the density function as in the real or vectorial case. So an alternative to this problem is the probability measures of small balls defined in A1 (see for more details Ferraty and Vieu (2006) and Farraty et al. (2006). Assumptions **A2** and **A4** are technical conditions imposed for the proofs. Assumption **A3** means that $F(.|.)$ satisfies a Holder condition with respect to each variable and is needed to evaluate the bias term in our asymptotic results. Assumption **A5** concerns the choice of the bandwidth which is closely linked to the small ball probability.

Remark 4.1. *The condition $b < b_V$ is essential to obtain the consistency of the proposed estimator (see Lemma 9.1 and Lemma 9.2). Indeed, when $b = b_V = min(b_F, b_G)$, the last value of the observations can be censored and will thus not be observed.*

Proposition 4.1. *Under Assumptions A1-A5, we have*

$$\sup_{x\in\Omega}\sup_{a_V\leq y\leq b}|F_n(y|x)-F(y|x)| = O\left(h_K^{\beta_1}+h_H^{\beta_2}\right)+O\left(\sqrt{\frac{\log n}{n\phi(h_K)}}\right) \text{ a.s. as } n\to\infty$$

Theorem 4.1. *Under Assumptions of Proposition 4.1 we have*

$$\sup_{x\in\Omega}|Q_{p,n}(x)-Q_p(x)| = O\left(h_K^{\beta_1}+h_H^{\beta_2}\right)+O\left(\sqrt{\frac{\log n}{n\phi(h_K)}}\right) \text{ a.s. as } n\to\infty$$

.

Remark 4.2. *We obtain the same convergence rate as those established by Farraty et al. (2006) in the complete case and Helal and Ould Said (2016) in the left truncation case. If we take $\beta_2 = 1$ our rate becomes similar to that provided by Horrigue and Ould Said (2011) in the right censoring case.*

5. Simulation study

This section is devoted to studying the behavior of the conditional quantile estimator $Q_{p,n}(.)$ defined in (8). The simulation is conducted to assess the performance of our estimate. For this aim, we consider the following functional regression model.

$$Y = \tilde{R}(X)+\varepsilon \tag{9}$$

where X and ε are independent, the error ε is normally distributed as $\mathcal{N}(0, 0.1)$ and $\tilde{R}(X) = \int_0^1 (X(t))^2 dt$. For the model (9), we fix the sample size n and we proceed with the following algorithm:

Step 1: The functional covariate $X(t)$ is generated on $[0, 1]$ by

$$X_i(t) = A_i(2-\cos(\pi t S_i)) + (1-A_i)\ \cos(\pi t S_i)\,,\ i\geq 1,\ t\in[0,1]\,,$$

where $S_i \hookrightarrow U[0,1]$ uniformly distributed on $[0, 1]$ and A_i is a Bernoulli random variable. $A_i \hookrightarrow \mathfrak{B}(0.5)$.

Step 2: We calculate

$$Y_i = \tilde{R}(X_i)+\varepsilon_i,\ i\geq 1$$

Note that under this model, the conditional random variable $(Y|X = x)$ follows $\mathcal{N}(\tilde{R}(x), 0.1)$.

Step 3: Generate i.i.d rv's C_i such that C_i is distributed according to the exponential law with parameter a_1, (a_1 allows to control the percentage of censoring (CR)). Take $Z_i = min(Y_i, C_i)$ and $\delta_i = \mathbf{1}_{\{Y_i\leq C_i\}}$.

Step 4: Generate the random variable $T_1 \hookrightarrow \mathcal{N}(\mu, 1)$, where μ is adapted in order to get different percentage of truncation (TR). If $Z_1 \geq T_1$, enter the observation $(X_1(t), Z_1, T_1, \delta_1)$ in the final sample. Otherwise, reject the observation $(X_1(t), Z_1, T_1, \delta_1)$ and go back to step 1. By repeating the processes above, the observed data $(X_i(t), Z_i, T_i, \delta_i)$, $i = 1, .., n$ are obtained. At the end of this procedure we get the deterministic variable N, which permits to get the truncation percentage.

Step 5:: We calculate the Lynden Bell and the concomitant TJW estimators L_n and G_n based on the observed data (Z_i, T_i, δ_i), $i = 1, 2, .., n$.

Step 6: We divide our sample into two randomly chosen sets of observations I and J

• training sample $(X_j, Z_j, T_j, \delta_j)_{j\in J}$.

• test sample $(X_i, Z_i, T_i, \delta_i)_{i\in I}$.

Here, the test sample size has been fixed to 30 regardless of the total sample size n (that means that $Card(I) = 30$ and $Car(J) = n - 30$).

Step 7: Using the training sample, we calculate the value of the conditional quantile estimator $Q_{p,n}(X_i)$ for each $i \in I$. For that we apply an iterative dichotomy method by choosing an initial interval $[Q_p(X_i) - a_2, Q_p(X_i) + a_2]$, where a_2 is some constant properly chosen. We use the quadratic kernel $K(x) = 1.5(1 - x^2)\mathbf{1}_{[0,1]}(x)$ and the distribution function $H(y) = \int_{-\infty}^{y} 0.75(1 - t^2)\mathbf{1}_{[-1,1]}(t)dt$. The selection of the optimal values h_{Kopt} and h_{Hopt} was performed by varying h_K over the interval $[0.01, 1.2]$ and h_H over the interval $[0.01, 1.3]$, with steps sizes of 0.02 and 0.03, respectively. The values minimizing the global mean square error ($GMSE$) defined in (10) below were then selected. This procedure was repeated for each combination of n, CR and TR.

We also choose the distance $d(., .)$ between the first derivatives of the curves described in Farraty et al. (2006) as a semi-metric in $\Im$.

$$d(x_1, x_2) = \left(\int_0^1 (x_1'(t) - x_2'(t))^2 dt \right)^{\frac{1}{2}} \quad x_1, x_2 \in \Im,$$

where x' denote the first derivative of x.

5.1. Simulation results

To visually assess the efficacy of our estimator, we start by plotting graphs that represent the true values $Q_p(.)$ (in blue) versus the predicted values $Q_{p,n}(.)$ (in red), for various combinations of the sample size n, the value p, the censoring rate (CR), and the truncation rate (TR), as illustrated in Figure 1, Figure 2, Figure 3 and Figure 4 (On the $y-$axis, the value of the conditional quantile for the $i-$th observation, $i = 1, ..., 30$).

1. Effect of sample size: For $p = 0.5$, $CR \approx 30\%$ and $TR \approx 30\%$, we take $n = 50$, 100, and 500. Figure 1, shows that the estimator exhibits good behavior as the sample size increases.

2. Effect of censoring rate: For $p = 0.5$, $n = 100$ and $TR \approx 20\%$, we take $CR \approx 0\%,\ 30\%$ and 50%. From Figure 2, we observe that the quality of fit improves as the percentage of censoring decreases. Indeed, the estimate is better for a lower censoring percentage, as confirmed in Table 1.

3. Effect of truncation rate: For $p = 0.5$, $n = 100$ and $CR \approx 30\%$, we choose $TR \approx 0\%,\ 20\%$ and 60%. From Figure 3, the quality of fit is slightly affected by the truncation percentage.

4. Effect of probability p: For $n = 100,\ CR \approx 30\%$ and $TR \approx 30\%$, we take $p = 0.25,\ 0.5$ and 0.75. Figure 4 illustrates that, the quality of fit for both $p = 0.25$ and $p = 0.75$ values is as good as for $p = 0.5$.

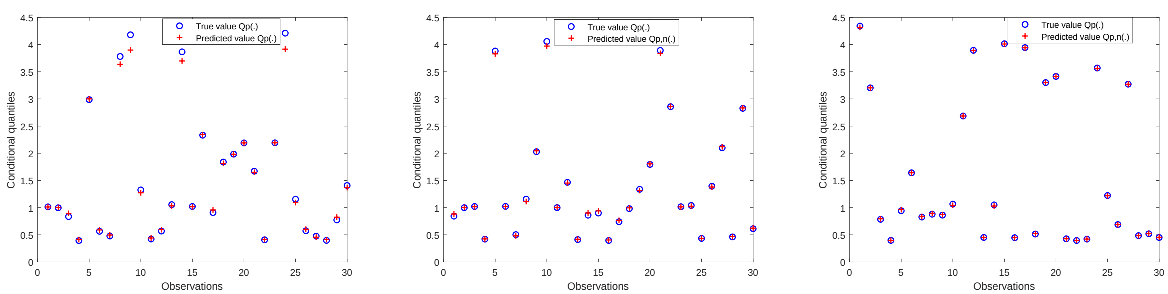

Figure 1: True (blue) and predicted (red) values for $p = 0.5$, $CR \approx 30\%$, $TR \approx 30\%$, $n = 50$, 100 and 500 respectively.

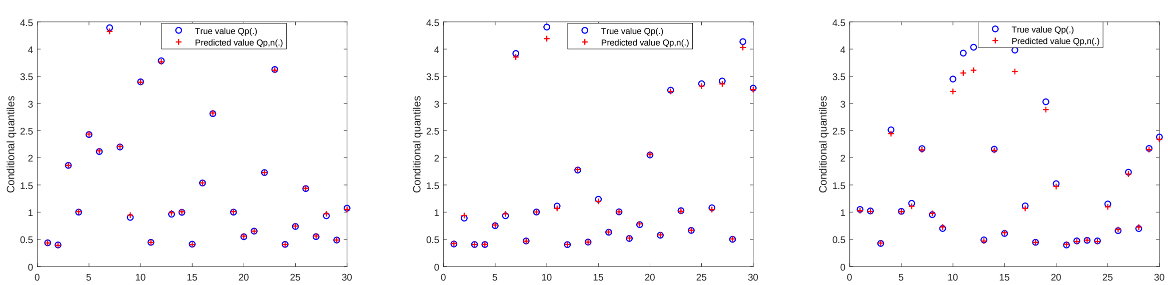

Figure 2: True (blue) and predicted (red) values for $p = 0.5$, $n = 100$, $TR \approx 20\%$, $CR \approx 0\%$, 30%, 50% respectively.

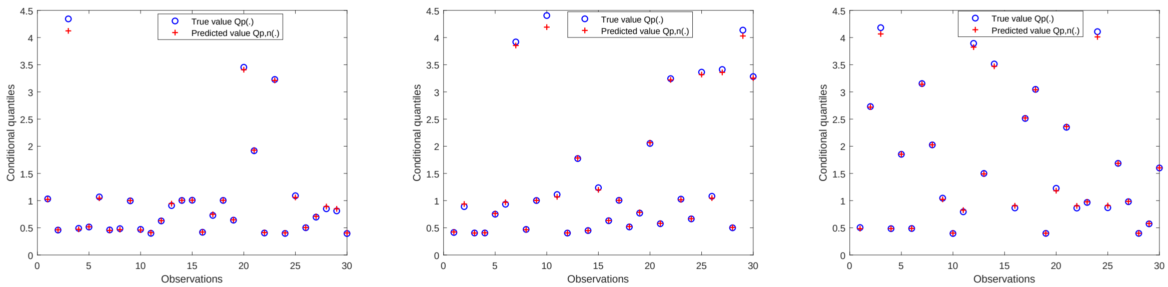

Figure 3: True (blue) and predicted (red) values for $p = 0.5$, $n = 100$, $CR \approx 30\%$, $TR \approx 0\%$, 20%, and 60% respectively.

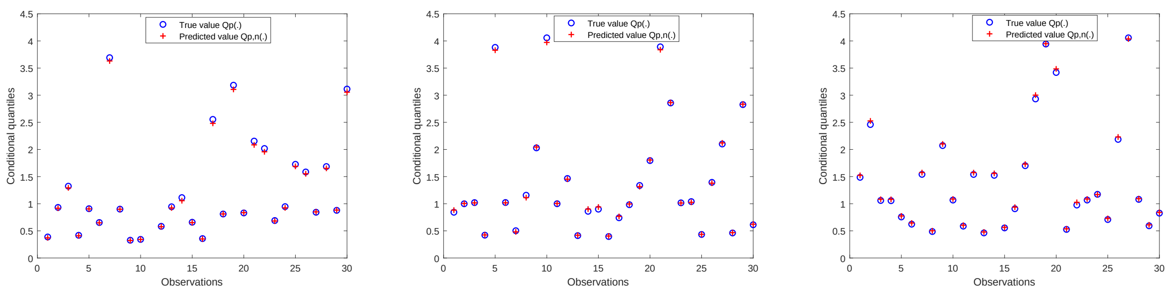

Figure 4: True (blue) and predicted (red) values for $n = 100$, $CR \approx 30\%$, $TR \approx 30\%$ and $p = 0.25, 0.5,$ and 0.75 respectively.

Next, we generate the data with observed sample size $n = 50,\ 100$ and 500, considering different percentage of censoring and truncation. In Table 1, we present the Global Mean Square Error ($GMSE$) at $p = 0.25,\ 0.5$ and 0.75, computed along $B = 100$ Monte Carlo trials. The $GMSE$ for the estimator $Q_{p,n}(.)$ of $Q_p(.)$ is defined as

$$GMSE = \frac{1}{IB}\sum_{i=1}^{I}\sum_{k=1}^{B}(Q_{p,n,k}(X_i) - Q_p(X_i))^2, \qquad (10)$$

where $Q_{p,n,k}(X_i)$ is the value of $Q_{p,n}(X_i)$ at iteration k, and I is the size of the test sample.
The results, reported in Table 1 show that:

1. The $GMSE$ decreases along with an increasing sample size n, so the estimator performs better for high sample size.
2. For a fixed truncation rate, the quality of fit is affected by the rate of censoring, and the estimator performance deteriorates as the percentage of censoring increases.
3. A high truncation rate slightly affects the quality of the fit.
4. The estimation quality is more affected by a high percentage of censoring than truncation. In fact, the presence of censoring serves to mitigate the impact of truncation on the estimate.
5. The $GMSE$ results reveal that the estimate performs well for the cases of $p = 0.25$ and $p = 0.75$ only with small sample sizes. Otherwise, for $p = 0.5$, the estimation is better for $n = 100$ and $n = 500$.

6. Application to real-world data

In this section, we present a real data application in which we evaluate the performance of the proposed conditional quantile estimator. The dataset comprising 8784 records of hourly energy consumption for the year 2016 (measured in

Table 1: The $GMSE$'s values for $Q_{p,n}$.

p	$TR\%$	$CR\%$	$GMSE$		
			$n = 50$	$n = 100$	$n = 500$
0.25	20	0	0.0136	0.0061	0.0014
	20	30	0.0293	0.0074	0.0019
	20	50	0.1182	0.0655	0.0053
	60	0	0.0101	0.0066	0.0013
	60	30	0.0331	0.0076	0.0018
	60	50	0.1334	0.0279	0.0031
0.5	20	0	0.0440	0.0034	0.0006
	20	30	0.0633	0.0049	0.0007
	20	50	0.1271	0.0396	0.0031
	60	0	0.0439	0.0037	0.0007
	60	30	0.0459	0.0039	0.0008
	60	50	0.1385	0.0247	0.0025
0.75	20	0	0.0120	0.0043	0.0013
	20	30	0.0607	0.0076	0.0021
	20	50	0.1901	0.0673	0.0047
	60	0	0.0319	0.0074	0.0015
	60	30	0.0513	0.0089	0.0017
	60	50	0.1132	0.0527	0.0048

MWh), sourced from the smart metering device of a consumer-type shopping center (a large hypermarket), and the corresponding historical hourly temperature (measures in degrees Celsius), recorded by weather sensors at a specialized institute. (for a more detailed description of this dataset, see Pîrjan et al. (2017) and Fetitah et al. (2020). We are interested in the estimation of the conditional quantile prediction of daily peak consumption of energy, Y_i, with respect to the hourly temperature $X_i\left(t_j\right)_{j=1..24}$ for the day i. The full dataset of 366 days is divided into a training set consisting of the first 351 days, and a test set comprising the remaining 15 days. From the training set, we restrict our analysis to the peak demand electricity exceeding the one-third quantile consumption, that is we consider only $Y_i > 0.848\ MWh$, $i = 1...351$, this choice amounts to truncating the training dataset approximately $TR = 34\%$, this phenomenon can be justified by a loss of information caused by technical failures. Additionally, an electricity shortage can occur on certain days and last for several hours. In such cases, we only have a lower limit for peak demand, but not the exact value. On these censored days, electricity consumption is only observed up to a certain time $t_c \in [1, 24]$, we thus provide a censorship rate around $CR = 22\%$, for further insight and illustrative examples of this censorship phenomenon, see Fetitah et al. (2020) or Boucetta et al. (2024).
Consequently the observed training sample is constructed as $(X_i, Z_i, T_i, \delta_i)_{i=1,\ldots,232}$, where $\delta_i = 1$ if $Z_i = Y_i$ and $\delta_i = 0$ if $Z_i = C_i$, C_i are the censored data, and $T_i = 0.848$, $i = 1, \ldots, 232$. Figure 5 illustrates clearly these data.

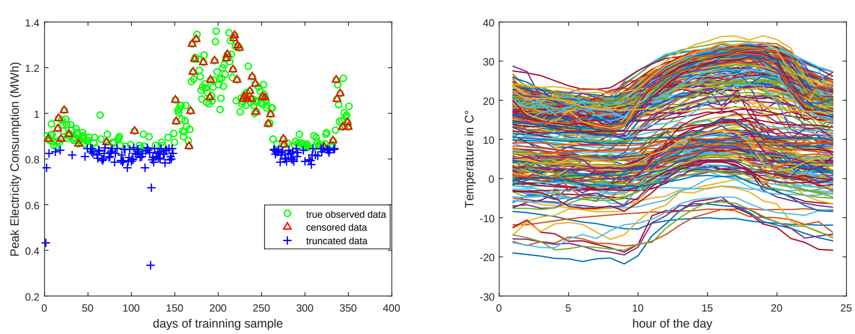

Figure 5: The peak electricity consumption and their corresponding observed temperature daily curves, for the training sample.

Conditional Quantiles estimators are built using the same quadratic kernel $K(.)$ and the same distribution $H(.)$ as in the simulation study. The choice of the semi-metric is based on the data, in our case we choose the Empirical L^2 distance defined by

$$d(X_i, X_j) = \left(\frac{1}{24} \sum_{k=1}^{24} (X_i(t_k) - X_j(t_k))^2 \right)^{1/2}.$$

The estimators are derived by selecting both the optimal bandwidths (h_{Kopt}, h_{Hopt}) that minimizes the Mean Quantile Loss function based on the Pinball Loss $\rho_p(u) = u\left(p - \mathbf{1}_{\{u<0\}}\right)$, calculated on the true observed training data in the grid over h_H and h_K, given by

$$(h_{Kopt}, h_{Hopt}) = arg \min_{h_H, h_K} \left\{ \frac{1}{232} \sum_{j=1}^{232} \delta_j (Y_j - Q_{p,n}^{(j)}(X_j)) \left(p - \mathbf{1}_{\left\{\delta_j (Y_j - Q_{p,n}^{(j)}(X_j))<0\right\}} \right) \right\}.$$

where $Q_{p,n}^{(j)}(.)$ are the conditional quantiles estimators calculated with the cross-validation technique. The results are given in Figure 6, which displays the prediction of the conditional quantile of peak electricity consumption for the 15 days of December, against the true peak demand (Y test), for $p = 0.25$, 0.5 and 0.75, respectively. It is important not to confuse these results with those from the simulation study, in which we overlaid the estimated quantile with the true quantile in order to evaluate the performance of our estimator. Here, we do not have the true quantile, but only the observed sample; therefore, the analysis follows a different approach.

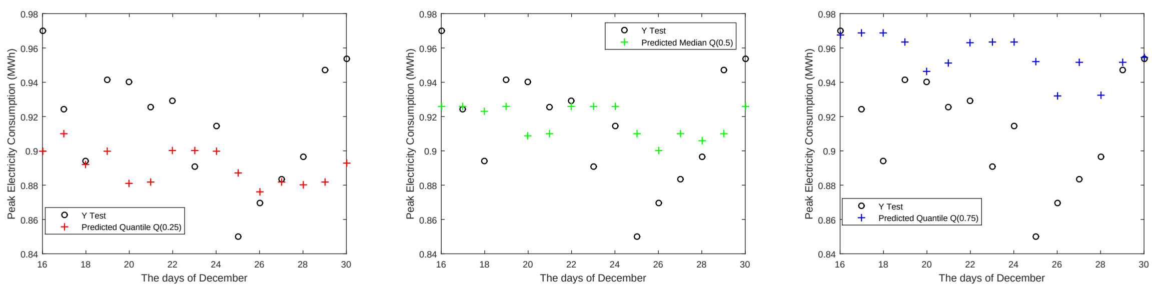

Figure 6: Predicted Quantile peak electricity consumption against the true observed consumption, for p=0.25, 0.5 and 0.75, respectively

When interpreting quantile-prediction, it's important not to mistake them for inaccurate predictions. Unlike traditional regression models that aim to estimate the precise value of a target variable, quantile regression aim to provide a range that captures the possible values of the data, as shown in Figure 7

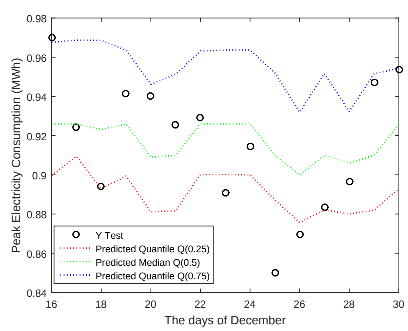

Figure 7: Predicted Interquartile range for peak electricity consumption against the true observed consumption,

To assess the quality of quantile predictions, we used two key criterion:

1. Mean Quantile Loss (MQL), measures how far our predictions are from actual values, helping us evaluate

accuracy, calculated on the Test sample

$$MQL(Y_i, Q_{p,n}(X_i)) = \frac{1}{15}\sum_{i=1}^{15}(Y_i - Q_{p,n}(X_i))\left(p - \mathbf{1}_{\{(Y_i - Q_{p,n}(X_i))<0\}}\right).$$

2. Empirical Coverage Probability (E.C.P), an approximation of the coverage probability which indicates how well our predictions capture the observed outcomes, which is vital for model reliability

$$E.C.P = \frac{1}{15}\sum_{i=1}^{15}\mathbf{1}_{\{Y_i \le Q_{p,n}(X_i)\}}.$$

The results are given in the Table 2

p	E.C.P	MQL
0.25	0.2667	0.0096
0.5	0.5333	0.0126
0.75	0.8665	0.0102

Table 2: Empirical Coverage Probability and Mean Quantile Loss for $p = 0.25, 0.5,$ and 0.75

By examining both criteria, it is evident that the losses associated with the three predicted quantiles are low, indicating that the model effectively captures the asymmetric nature of the quantile functions. The empirical coverage probabilities for the first and second conditional quartiles are well-calibrated (very close to the theoretical coverage 25% and 50%, respectively). For the third quartile, even though it appears to be overestimated, the result remains satisfactory.

7. Proofs of the main results

Note that the estimator $F_n(y|x)$ can be rewritten as $F_n(y|x) =: \frac{\Psi_n(x,y)}{\ell_n(x)}$
where

$$\Psi_n(x,y) := \frac{\theta_n}{n\mathbb{E}\left[K\left(\frac{d(x,X_1)}{h_K}\right)\right]}\sum_{i=1}^{n}\frac{\delta_i}{L_n(Z_i)\overline{G}_n(Z_i)}K\left(\frac{d(x,X_i)}{h_K}\right)H\left(\frac{y-Z_i}{h_H}\right), \tag{11}$$

and

$$\ell_n(x) := \frac{\theta_n}{n\mathbb{E}\left[K\left(\frac{d(x,X_1)}{h_K}\right)\right]}\sum_{i=1}^{n}\frac{\delta_i}{L_n(Z_i)\overline{G}_n(Z_i)}K\left(\frac{d(x,X_i)}{h_K}\right). \tag{12}$$

Similarly to (11) and (12) we define the following pseudo estimators

$$\tilde{\Psi}_n(x,y) := \frac{\theta}{n\mathbb{E}\left[K\left(\frac{d(x,X_1)}{h_K}\right)\right]}\sum_{i=1}^{n}\frac{\delta_i}{L(Z_i)\overline{G}(Z_i)}K\left(\frac{d(x,X_i)}{h_K}\right)H\left(\frac{y-Z_i}{h_H}\right),$$

and

$$\tilde{\ell}_n(x) := \frac{\theta}{n\mathbb{E}\left[K\left(\frac{d(x,X_1)}{h_K}\right)\right]}\sum_{i=1}^{n}\frac{\delta_i}{L(Z_i)\overline{G}(Z_i)}K\left(\frac{d(x,X_i)}{h_K}\right),$$

and we consider the decomposition

$$\begin{aligned} F_n(y|x) - F(y|x) &= \frac{1}{\ell_n(x)}\left\{\left(\Psi_n(x,y) - \tilde{\Psi}_n(x,y)\right) + \left(\tilde{\Psi}_n(x,y) - \mathbf{E}[\tilde{\Psi}_n(x,y)]\right)\right. \\ &+ \left.\left(\mathbf{E}[\tilde{\Psi}_n(x,y)] - F(y|x)\right)\right\} - \frac{F(y|x)}{\ell_n(x)}\left\{\left(\ell_n(x) - \tilde{\ell}_n(x)\right)\right. \\ &+ \left.\left(\tilde{\ell}_n(x) - \mathbf{E}[\tilde{\ell}_n(x)]\right) + \left(\mathbf{E}[\tilde{\ell}_n(x)] - 1\right)\right\}. \end{aligned} \tag{13}$$

Proof of Proposition 4.1. The proof of Proposition 4.1 is deduced from the decomposition (13), indeed

$$\begin{aligned}\sup_{x\in\Omega}\sup_{a_V\le y\le b}|F_n(y|x)-F(y|x)| \quad\le\quad & \frac{1}{\inf_{x\in\Omega}|\ell_n(x)|}\left\{\sup_{x\in\Omega}\sup_{a_V\le y\le b}\left|\Psi_n(x,y)-\tilde{\Psi}_n(x,y)\right|\right.\\ & +\sup_{x\in\Omega}\sup_{a_V\le y\le b}\left|\tilde{\Psi}_n(x,y)-\mathbf{E}[\tilde{\Psi}_n(x,y)]\right|\\ & +\sup_{x\in\Omega}\sup_{a_V\le y\le b}\left|\mathbf{E}[\tilde{\Psi}_n(x,y)]-F(y|x)\right|\\ & \left.+\sup_{x\in\Omega}\sup_{a_V\le y\le b}|F(y|x)|\sup_{x\in\Omega}|1-\ell_n(x)|\right\},\end{aligned}$$

hence, the result is a direct consequence of Lemmas 9.1–9.7 and Corollary 9.1.

Proof of Theorem 4.1.

Let $x\in\Omega$, $F(.|x)$ and $F_n(.|x)$ are continuous, we have $F(Q_p(x)|x)=F_n(Q_{p,n}(x)|x)=p$. So

$$\begin{aligned}|F(Q_{p,n}(x)|x)-F(Q_p(x)|x)| \quad&\le\quad |F(Q_{p,n}(x)|x)-F_n(Q_{p,n}(x)|x)|+|F_n(Q_{p,n}(x)|x)-F(Q_p(x)|x)|\\ &\le\quad |F(Q_{p,n}(x)|x)-F_n(Q_{p,n}(x)|x)|\\ &\le\quad \sup_{a_V\le y\le b}|F_n(y|x)-F(y|x)|\,. \end{aligned}\tag{14}$$

A Taylor's expansion of the function $F(Q_{p,n}(x)|x)$ in a neighborhood of $Q_p(x)$, implies that

$$F(Q_{p,n}(x)|x)-F(Q_p(x)|x)=(Q_{p,n}(x)-Q_p(x))f(Q_p^*(x)|x),$$

where $Q_p^*(x)$ is between $Q_p(x)$ and $Q_{p,n}(x)$. Then using the inequality (14) we obtain

$$\sup_{x\in\Omega}|Q_{p,n}(x)-Q_p(x)||f(Q_p^*(x)|x)|\le\sup_{x\in\Omega}\sup_{a_V\le y\le b}|F_n(y|x)-F(y|x)|. \tag{15}$$

From (15), Assumption A3(ii) and Proposition 4.1 we conclude the proof of the Theorem 4.1.

8. Conclusion

It is worth recalling that the construction of the conditional quantile estimator from the conditional distribution function follows the same principle in the case of complete data (Ferraty and Vieu, 2006), right-censored data (Horrigue and Ould Said, 2011), and left-truncated data (Hellal and Ould Said, 2016). The fundamental difference between these settings lies in the estimation of the conditional distribution function $F(.|x)$. In this paper our main contribution was to simultaneously address two forms of incomplete data, namely right censoring and left truncation. As highlighted in Remark 3.1, the estimator we propose is not merely a combination of the methods developed for the censored and truncated cases. Nevertheless, the results obtained are comparable to those reported in the three aforementioned settings (see Remark 4.2) and are supported by an extensive simulation study and a real-data application.

9. Appendix

Lemma 9.1. *Under Assumptions A1, A2 and A5, we have*

$$\sup_{x\in\Omega}\left|\ell_n(x)-\tilde{\ell}_n(x)\right|=O\left(\sqrt{\frac{\log\log n}{n}}\right)\ a.s.\ as\ n\to\infty.$$

Proof. We write

$$\begin{aligned}\ell_n(x)-\tilde{\ell}_n(x) &= \frac{1}{n\mathbb{E}\left[K\left(\frac{d(x,X_1)}{h_K}\right)\right]}\sum_{i=1}^{n}\left(\frac{\delta_i\theta_n}{L_n(Z_i)\overline{G}_n(Z_i)}K\left(\frac{d(x,X_i)}{h_K}\right)\right)-\frac{\delta_i\theta}{L(Z_i)\overline{G}(Z_i)}K\left(\frac{d(x,X_i)}{h_K}\right)\\ &= \frac{1}{n\mathbb{E}\left[K\left(\frac{d(x,X_1)}{h_K}\right)\right]}\sum_{i=1}^{n}\delta_iK\left(\frac{d(x,X_i)}{h_K}\right)\left(\frac{\theta_n}{L_n(Z_i)\overline{G}_n(Z_i)}-\frac{\theta}{L(Z_i)\bar{G}(Z_i)}\right)\end{aligned}$$

Replacing θ and θ_n by their expressions defined in (1) and (2) respectively, we get

$$\begin{aligned}\left|\ell_n(x)-\tilde{\ell}_n(x)\right| &\leq \frac{1}{n\mathbb{E}\left[K\left(\frac{d(x,X_1)}{h_K}\right)\right]}\sum_{i=1}^{n}\delta_iK\left(\frac{d(x,X_i)}{h_K}\right)\left(\left|\frac{\overline{F}_n(Z_i)}{C_n(Z_i)}-\frac{\overline{F}(Z_i)}{C(Z_i)}\right|\right)\\ &\leq \frac{1}{n\mathbb{E}\left[K\left(\frac{d(x,X_1)}{h_K}\right)\right]}\sum_{i=1}^{n}\delta_iK\left(\frac{d(x,X_i)}{h_K}\right)\left(\left|\frac{F_n\left(Z_i\right)-F\left(Z_i\right)}{C_n\left(Z_i\right)}\right|\right.\\ &+ \left.(1-F(Z_i))\left|\frac{1}{C_n(Z_i)}-\frac{1}{C(Z_i)}\right|\right)\\ &\leq \left(\sup_{a_V\leq y\leq b}\left|\frac{F_n(y)-F(y)}{C_n(y)}\right|+\sup_{a_V\leq y\leq b}\left|\frac{1-F(y)}{C_n(y)C(y)}(C_n(y)-C(y))\right|\right)\\ &\times \left(\frac{1}{n\mathbb{E}\left[K\left(\frac{d(x,X_1)}{h_K}\right)\right]}\sum_{i=1}^{n}\delta_iK\left(\frac{d(x,X_i)}{h_K}\right)\right)\\ &= :J_1\times J_2.\end{aligned}$$

On the one hand, following Zhou and Yip (1999), we have $\sup_{a_V\leq y\leq b}|F_n(y)-F(y)|=O\left(\sqrt{\frac{\log\log n}{n}}\right)$ a.s., and the Law of the Iterated Logarithm (LIL) stated for the empirical d.f in the independent case gives $\sup_{a_V\leq y\leq b}|C_n(y)-C(y)|=O\left(\sqrt{\frac{\log\log n}{n}}\right)$ a.s. as $n\to\infty$. Furthermore, since $C(y)>0$, for $a_V\leq y\leq b$, then there exists a $\tau_0>0$ such that $C(y)\geq\tau_0>0$ and as $C_n(y)\to C(y)$a.s, we obtain $J_1=O\left(\sqrt{\frac{\log\log n}{n}}\right)$ .a.s. In the other hand, we have

$$\begin{aligned}J_2 &= \frac{1}{n\mathbb{E}\left[K\left(\frac{d(x,X_1)}{h_K}\right)\right]}\sum_{i=1}^{n}\left(\delta_iK\left(\frac{d(x,X_i)}{h_K}\right)-\mathbf{E}\left[\delta_iK\left(\frac{d(x,X_i)}{h_K}\right)\right]\right)\\ &+ \frac{1}{\mathbb{E}\left[K\left(\frac{d(x,X_1)}{h_K}\right)\right]}\mathbf{E}\left[\delta_1K\left(\frac{d(x,X_1)}{h_K}\right)\right]\\ &= :J_{21}+J_{22}.\end{aligned}\tag{16}$$

For J_{21}, we set $\Delta_i(x)=\frac{1}{\mathbb{E}\left[K\left(\frac{d(x,X_1)}{h_K}\right)\right]}\left(\delta_iK\left(\frac{d(x,X_i)}{h_K}\right)-\mathbf{E}\left[\delta_iK\left(\frac{d(x,X_i)}{h_K}\right)\right]\right)$. Obviously, $J_{21}=\frac{1}{n}\sum_{i=1}^{n}\Delta_i(x)$.

The main tool here is the use of an exponential inequality given by corollary A8(i) in Ferraty and Vieu (2006). To do that, we have to evaluate the term $\mathbf{E}[|\Delta_1(x)|^m]$, for all $m\geq 2$. By Assumptions A1 and A2(i) we have

$$c_3\mathbf{1}_{B(x,h_K)}(X_1)\leq K\left(\frac{d(x,X_1)}{h_K}\right)\leq c_4\mathbf{1}_{B(x,h_K)}(X_1).$$

Therefore

$$c'\phi(h_K)\leq\mathbb{E}\left[K\left(\frac{d(x,X_1)}{h_K}\right)\right]\leq c\phi(h_K),\tag{17}$$

and for $l \geq 1$

$$\mathbb{E}\left[K^l\left(\frac{d(x,X_1)}{h_K}\right)\right] \leq c_4^l \phi_x(h_K) \leq c_4^l c_2 \phi(h_K).$$

Then

$$\begin{aligned}
\mathbf{E}\left[\delta_1 K^l\left(\frac{d(x,X_1)}{h_K}\right)\right] &= \int\int K^l\left(\frac{d(x,u)}{h_K}\right)\mathbf{v}_1^*(t|u)d\mathbf{P}^{X_1}(u)dt \\
&= \int\int K^l\left(\frac{d(x,u)}{h_K}\right)\frac{L(t)\overline{G}(t)}{\theta}f(t|u)d\mathbb{P}^{X_1}(u)dt \\
&\leq \frac{1}{\theta}\mathbb{E}\left[K^l\left(\frac{d(x,X_1)}{h_K}\right)\right] \\
&\leq \frac{c_4^l c_2}{\theta}\phi(h_K) =: O(\phi(h_K)). \qquad (18)
\end{aligned}$$

The Newton's binomial expansion, combined with (17) and (18) allows us to write

$$\begin{aligned}
\mathbf{E}\left[|\Delta_1(x)|^m\right] &= \mathbf{E}\left[\left|\frac{1}{\mathbb{E}^m\left[K\left(\frac{d(x,X_1)}{h_K}\right)\right]}\sum_{j=0}^m C_m^j \delta_1 K^j\left(\frac{d(x,X_1)}{h_K}\right)\mathbf{E}^{m-j}\left[\delta_1 K\left(\frac{d(x,X_1)}{h_K}\right)\right](-1)^{m-j}\right|\right] \\
&\leq \frac{1}{\mathbb{E}^m\left[K\left(\frac{d(x,X_1)}{h_K}\right)\right]}\sum_{j=0}^m C_m^j \mathbf{E}\left[\delta_1\left|K^j\left(\frac{d(x,X_1)}{h_K}\right)\right|\right]\left|\mathbf{E}^{m-j}\left[\delta_1 K\left(\frac{d(x,X_1)}{h_K}\right)\right]\right| \\
&\leq c\max_{0\leq j\leq m}\phi^{1-j}(h_K) \\
&\leq c\phi^{1-m}(h_K).
\end{aligned}$$

Now, we can apply corollary A8(i) in Ferraty and Vieu (2006) with $a = (\phi(h_K))^{-\frac{1}{2}}$, to get for all $\varepsilon > 0$

$$\begin{aligned}
\mathbf{P}\left(|J_{21}| > \varepsilon\right) &= \mathbf{P}\left(\frac{1}{n}\left|\sum_{i=1}^n \Delta_i(x)\right| > \varepsilon\right) \\
&\leq 2\exp\left\{-\frac{n\varepsilon^2}{2a^2(1+\varepsilon)}\right\}.
\end{aligned}$$

Taking $\varepsilon = \varepsilon_0\sqrt{\frac{\log n}{n\phi(h_K)}}$, for all $\varepsilon_0 > 0$, we obtain

$$\mathbf{P}\left(|J_{21}| > \varepsilon_0\sqrt{\frac{\log n}{n\phi(h_K)}}\right) \leq 2\exp\left\{-\frac{\varepsilon_0^2\log n}{2\left(1+\varepsilon_0\sqrt{\frac{\log n}{n\phi(h_K)}}\right)}\right\}.$$

Under A5(ii), for n large enough, we have

$$\begin{aligned}
\mathbf{P}\left(|J_{21}| > \varepsilon_0\sqrt{\frac{\log n}{n\phi(h_K)}}\right) &\leq 2\exp\{-c\varepsilon_0^2\log n\} \\
&\leq 2n^{-c\varepsilon_0^2}. \qquad (19)
\end{aligned}$$

For a suitable choice of ε_0 (i.e. $\varepsilon_0 > \sqrt{\frac{1}{c}}$), the term on the right hand side of (19) is the general term of a convergent series, then Borel-Cantelli's lemma gives

$$\sup_{x\in\Omega}|J_{21}| = O\left(\sqrt{\frac{\log n}{n\phi(h_K)}}\right) \text{ a.s. as } n\to\infty. \qquad (20)$$

For J_{22}, using the same argument as in (18), we have

$$\begin{aligned}
\frac{1}{\mathbb{E}\left[K\left(\frac{d(x,X_1)}{h_K}\right)\right]}\mathbf{E}\left[\delta_1 K\left(\frac{d(x,X_1)}{h_K}\right)\right] &= \frac{1}{\mathbb{E}\left[K\left(\frac{d(x,X_1)}{h_K}\right)\right]}\int\int K\left(\frac{d(x,u)}{h_K}\right)\mathbf{v}_1^*(t|u)d\mathbf{P}^{X_1}(u)dt \\
&\leq \frac{1}{\theta\mathbb{E}\left[K\left(\frac{d(x,X_1)}{h_K}\right)\right]}\mathbb{E}\left[K\left(\frac{d(x,X_1)}{h_K}\right)\right] \\
&= O(1),
\end{aligned}$$

which together with (16) and (20) give

$$\sup_{x\in\Omega}|J_2| = O(1) \text{ a.s. as } n\to\infty.$$

So the Lemma 9.1 is proved.

Lemma 9.2. *Under Assumptions A1, A2 and A5, we have*

$$\sup_{x\in\Omega}\left|\tilde{\ell}_n(x)-\mathbf{E}[\tilde{\ell}_n(x)]\right| = O\left(\sqrt{\frac{\log n}{n\phi(h_K)}}\right) \textit{ a.s. as } n\to\infty.$$

Proof. The compact Ω can be covered by a finite number l_n of balls $B(x_k, r_n)$ of radius r_n and centered at x_k, where $\Omega\subset\bigcup_{k=1}^{l_n}B(x_k,r_n)$ such that $l_n r_n \leq M_0$ where M_0 is a positive finite constant. For any $x\in\Omega$, we set $k_0(x) = \arg\min_{1\leq k\leq l_n} d(x_k, x)$ and consider the following decomposition

$$\left|\tilde{\ell}_n(x)-\mathbf{E}[\tilde{\ell}_n(x)]\right| \leq \left|\tilde{\ell}_n(x)-\tilde{\ell}_n(x_{k_0(x)})\right| + \left|\tilde{\ell}_n(x_{k_0(x)})-\mathbf{E}[\tilde{\ell}_n(x_{k_0(x)})]\right| + \left|\mathbf{E}[\tilde{\ell}_n(x_{k_0(x)})]-\mathbf{E}[\tilde{\ell}_n(x)]\right|.$$

We have

$$\begin{aligned}
\sup_{x\in\Omega}\left|\tilde{\ell}_n(x)-\mathbf{E}[\tilde{\ell}_n(x))]\right| &\leq \max_{1\leq k\leq l_n}\sup_{x\in B(x_k,r_n)}\left|\tilde{\ell}_n(x))-\tilde{\ell}_n(x_{k_0(x)})\right| + \max_{1\leq k\leq l_n}\left|\tilde{\ell}_n(x_{k_0(x)})-\mathbf{E}[\tilde{\ell}_n(x_{k_0(x)})]\right| \\
&+ \max_{1\leq k\leq l_n}\sup_{x\in B(x_k,r_n)}\left|\mathbf{E}[\tilde{\ell}_n(x_{k_0(x)})]-\mathbf{E}[\tilde{\ell}_n(x)]\right| \\
&=: I_1+I_2+I_3.
\end{aligned}$$

For I_1, we use similar arguments as in lemma 4.2 of Horrigue and Ould Said (2011).

$$\begin{aligned}
\left|\tilde{\ell}_n(x)-\tilde{\ell}_n(x_{k_0(x)})\right| &= \left|\frac{1}{n}\sum_{i=1}^n\frac{\theta\delta_i}{L(Z_i)\overline{G}(Z_i)}\left(\frac{K\left(\frac{d(x,X_i)}{h_K}\right)}{\mathbb{E}\left[K\left(\frac{d(x,X_1)}{h_K}\right)\right]}-\frac{K\left(\frac{d(x_{k_0(x)},X_i)}{h_K}\right)}{\mathbb{E}\left[\frac{d(x_{k_0(x)},X_1)}{h_K}\right]}\right)\right| \\
&\leq \frac{1}{n}\sum_{i=1}^n\frac{\theta}{L(Y_i)\bar{G}(Y_i)}\underbrace{\left|\frac{K\left(\frac{d(x,X_i)}{h_K}\right)-K\left(\frac{d(x_{k_0(x)},X_i)}{h_K}\right)}{\mathbb{E}\left[K\left(\frac{d(x,X_1)}{h_K}\right)\right]}\right|}_{I_1^*} \\
&+ \underbrace{\left|\frac{K\left(\frac{d(x_{k_0(x)},X_i)}{h_K}\right)}{\mathbb{E}\left[K\left(\frac{d(x,X_1)}{h_K}\right)\right]}-\frac{K\left(\frac{d(x_{k_0(x)},X_i)}{h_K}\right)}{\mathbb{E}\left[K\left(\frac{d(x_{k_0(x)},X_1)}{h_K}\right)\right]}\right|}_{I_2^*}.
\end{aligned}$$

On the one hand, since Assumption A2 implies that K is Lipschitz, we obtain

$$\begin{aligned} I_1^* &\leq \frac{c_8}{\mathbb{E}\left[K\left(\frac{d(x,X_1)}{h_K}\right)\right] h_K} \left|d(x,X_i) - d(x_{k_0(x)},X_i)\right| \\ &\leq \frac{c_8 d(x,x_{k_0(x)})}{\mathbb{E}\left[K\left(\frac{d(x,X_1)}{h_K}\right)\right] h_K}, \end{aligned} \tag{21}$$

On the other hand, there exists $x^* \in B(x_{k_0(x)}, h_K)$, and under Assumption A2 we have

$$\begin{aligned} I_2^* &= \left| \frac{K\left(\frac{d(x_{k_0(x)},X_i)}{h_K}\right)\left(\mathbb{E}\left[K\left(\frac{d(x_{k_0(x)},X_1)}{h_K}\right)\right] - \mathbb{E}\left[K\left(\frac{d(x,X_1)}{h_K}\right)\right]\right)}{\mathbb{E}\left[K\left(\frac{d(x,X_1)}{h_K}\right)\right]\mathbb{E}\left[K\left(\frac{d(x_{k_0(x)},X_1)}{h_K}\right)\right]} \right| \\ &\leq \frac{\left|K\left(\frac{d(x_{k_0(x)},\ X_i)}{h_K}\right)\right|}{\mathbb{E}\left[K\left(\frac{d(x,X_1)}{h_K}\right)\right]\mathbb{E}\left[K\left(\frac{d(x_{k_0(x)},\ X_1)}{h_K}\right)\right]}\mathbb{E}\left[\left|K\left(\frac{d(x_{k_0(x)},\ X_1)}{h_K}\right) - K\left(\frac{d(x,X_1)}{h_K}\right)\right|\right] \\ &\leq \frac{||K||_\infty\,\mathbb{E}\left[\left|K'\left(\frac{d(x^*,\ X_1)}{h_K}\right)\right|\right]}{\mathbb{E}\left[K\left(\frac{d(x,X_1)}{h_K}\right)\right]\mathbb{E}\left[K\left(\frac{d(x_{k_0(x)},\ X_1)}{h_K}\right)\right] h_K} d\left(x, x_{k_0(x)}\right). \end{aligned}$$

Moreover, for any $x \in \Im$, the term $\frac{d(x,X_i)}{h_K}$ is a real random variable and denote its distribution by $\mathbb{P}^{\frac{d(x,X_i)}{h_K}}(u) = \mathbb{P}\left(\frac{d(x,X_i)}{h_K} \leq u\right)$. We have

$$\begin{aligned} \mathbb{E}\left[\left|K'\left(\frac{d(x^*,\ X_1)}{h_K}\right)\right|\right] &= \int_0^1 |K'(t)|\, d\mathbb{P}^{\frac{d(x^*,\ X_1)}{h_K}}(t) \\ &\leq ||K'||_\infty \left[\mathbb{P}\left(\frac{d(x^*,\ X_1)}{h_K} \leq t\right)\right]_0^1 \\ &\leq \|K'\|_\infty \left[\mathbb{P}(d(x^*,X_1) \leq h_K) - \mathbb{P}(d(x^*,X_1) \leq 0)\right] \\ &\leq \|K'\|_\infty \left[\phi_{x^*}(h_K) - \phi_{x^*}(0)\right] \\ &\leq \|K'\|_\infty \phi_{x^*}(h_K), \quad \text{since} \ \ \phi_{x^*}(0) = 0. \end{aligned}$$

Then

$$I_2^* \leq \frac{\|K'\|_\infty \|K\|_\infty \phi_{x^*}(h_K)}{\mathbb{E}\left[K\left(\frac{d(x,X_1)}{h_K}\right)\right]\mathbb{E}\left[K\left(\frac{d(x_{k_0(x)},X_1)}{h_K}\right)\right] h_K} d(x, x_{k_0(x)}). \tag{22}$$

From (21) and (22), we have

$$\begin{aligned} \left|\tilde{\ell}_n(x) - \tilde{\ell}_n(x_{k_0(x)})\right| \leq \frac{\theta}{L(a_V)\overline{G}(b)} &\left(\frac{c_8}{\mathbb{E}\left[K\left(\frac{d(x,X_1)}{h_K}\right)\right] h_K} d(x, x_{k_0(x)}) \right. \\ &\left. + \frac{\|K'\|_\infty \|K\|_\infty \phi_{x^*}(h_K)}{\mathbb{E}\left[K\left(\frac{d(x,X_1)}{h_K}\right)\right]\mathbb{E}\left[K\left(\frac{d(x_{k_0(x)},X_1)}{h_K}\right)\right] h_K} d(x, x_{k_0(x)}) \right). \end{aligned}$$

Hence, from (17) and Assumption A1, we get

$$\begin{aligned}\left|\tilde{\ell}_n(x)-\tilde{\ell}_n(x_{k_0(x)})\right| &\leq \frac{\theta}{L(a_V)\overline{G}(b)}\left(\frac{c_8}{c'h_K\phi\left(h_K\right)}+\frac{c_2\|K'\|_\infty\|K\|_\infty\phi(h_K)}{c'^2h_K(\phi(h_K)^2}\right)d(x,x_{k_0(x)})\\ &\leq \frac{\theta}{L(a_V)\overline{G}(b)}\left(\frac{c_8}{c'h_K\phi(h_K)}+\frac{c_2\|K\|_\infty\|K'\|_\infty}{c'^2h_K\phi(h_K)}\right)d(x,x_{k_0(x)})\\ &\leq \frac{\theta\left(c_8c'+c_2\|K\|_\infty\|K'\|_\infty\right)}{c'^2L(a_V)\overline{G}(b)h_K\phi(h_K)}r_n.\end{aligned}$$

Taking $r_n = h_K\sqrt{\frac{\phi(h_K)}{n}}$ and from Assumption A2, we obtain that

$$\begin{aligned}\left|\tilde{\ell}_n(x)-\tilde{\ell}_n(x_{k_0(x)})\right| &\leq \frac{\theta\left(c_8c'+c_2\|K\|_\infty\|K'\|_\infty\right)}{c'^2L(a_V)\overline{G}(b)}\sqrt{\frac{1}{n\phi(h_K)}}\\ &= O\left(\sqrt{\frac{1}{n\phi(h_K)}}\right).\end{aligned}$$

Therefore

$$I_1=\max_{1\leq k\leq l_n}\sup_{x\in B(x_k,r_n)}\left|\tilde{\ell}_n(x))-\tilde{\ell}_n(x_{k_0(x)})\right| = O\left(\sqrt{\frac{1}{n\phi(h_K)}}\right)\text{ a.s. as } n\to\infty. \tag{23}$$

For I_3, and from (23) it comes that

$$\begin{aligned}\left|\mathbf{E}[\tilde{\ell}_n(x_{k_0(x)})]-\mathbf{E}[\tilde{\ell}_n(x)]\right| &\leq \mathbf{E}\left[\left|\tilde{\ell}_n(x_{k_0(x)})-\tilde{\ell}_n(x)\right|\right]\\ &= O\left(\sqrt{\frac{1}{n\phi(h_K)}}\right).\end{aligned}$$

Hence

$$I_3=\max_{1\leq k\leq l_n}\sup_{x\in B(x_k,r_n)}\left|\mathbf{E}[\tilde{\ell}_n(x_{k_0(x)})]-\mathbf{E}[\tilde{\ell}_n(x)]\right| = O\left(\sqrt{\frac{1}{n\phi(h_K)}}\right)\text{ a.s. as } n\to\infty. \tag{24}$$

Now, in order to study the term I_2, we set

$$\tilde{\ell}_n(x_{k_0(x)})-\mathbf{E}[\tilde{\ell}_n(x_{k_0(x)})]=\frac{1}{n}\sum_{i=1}^{n}\Gamma_i(x_{k_0(x)}),$$

where

$$\begin{aligned}\Gamma_i(x_{k_0(x)}) &= \frac{1}{\mathbb{E}\left[K\left(\frac{d(x_{k_0(x)},\,X_1}{h_K}\right)\right]}\left(\frac{\theta\delta_i}{L(Z_i)\overline{G}(Z_i)}K\left(\frac{d(x_{k_0(x)},\ X_i}{h_K}\right)\right.\\ &- \left.\mathbf{E}\left[\frac{\theta\delta_1}{L(Z_1)\overline{G}(Z_1)}K\left(\frac{d(x_{k_0(x)},\ X_1}{h_K}\right)\right]\right).\end{aligned}$$

The application of the exponential inequality of corollary A8 in Ferraty and Vieu (2006) is based on

$$\begin{aligned}\mathbf{E}\left[\left|\Gamma_1(x_{k_0(x)})\right|^m\right] &= \frac{1}{\left|\mathbb{E}^m\left[K\left(\frac{d(x_{k_0(x)},X_1)}{h_K}\right)\right]\right|}\mathbf{E}\left|\left[\frac{\theta\delta_1}{L(Z_1)\overline{G}(Z_1)}K\left(\frac{d(x_{k_0(x)},\ X_1)}{h_K}\right)\right]\right.\\ &- \left.\mathbf{E}\left[\frac{\theta\delta_1}{L(Z_1)\overline{G}(Z_1)}K\left(\frac{d(x_{k_0(x)},\ X_1)}{h_K}\right)\right]\right|^m,\end{aligned}$$

by Newton's binomial formula, we obtain for $q \leq m$

$$\begin{aligned}\mathbf{E}\left[\left|\Gamma_1(x_{k_0(x)})\right|^m\right] &\leq \frac{1}{\left|\mathbb{E}^m\left[K\left(\frac{d(x_{k_0(x)},X_1)}{h_K}\right)\right]\right|}\left\{\sum_{q=0}^m C_m^q \mathbf{E}\left[\left|\frac{\theta}{L(Y_1)\overline{G}(Y_1)}K\left(\frac{d(x_{k_0(x)},X_1)}{h_K}\right)\right|^q\right]\right.\\ &\qquad \left.\times\left|\mathbf{E}\left[\frac{\theta}{L(Y_1)\overline{G}(Y_1)}K\left(\frac{d(x_{k_0(x)},X_1)}{h_K}\right)\right]\right|^{m-q}\right\}.\end{aligned}$$

Using the same arguments as those invoked in proof of Lemma 9.1, we get for $m \geq 2$

$$\mathbf{E}\left[\left|\Gamma_1(x_{k_0(x)})\right|^m\right] = O(\max_{0\leq q\leq m}\left(\phi\left(h_K\right)\right)^{1-q}) = O\left(\left(\phi\left(h_K\right)\right)^{1-m}\right).$$

The corollary A8(i) of Ferraty and Vieu (2006) with $a = (\phi(h_K))^{-\frac{1}{2}}$ gives

$$\begin{aligned}\mathbf{P}\left(\max_{1\leq k\leq l_n}\left|\tilde{\ell}_n(x_{k_0(x)}) - \mathbf{E}[\tilde{\ell}_n(x_{k_0(x)})]\right| > \varepsilon\right) &\leq \sum_{k=1}^{l_n}\mathbf{P}\left(\left|\tilde{\ell}_n(x_{k_0(x)}) - \mathbf{E}[\tilde{\ell}_n(x_{k_0(x)})]\right| > \varepsilon\right)\\ &\leq l_n\mathbf{P}\left(\sum_{i=1}^n\left|\Gamma_i(x_{k_0(x)})\right| > n\varepsilon\right)\\ &\leq 2l_n exp\left(-\frac{n\varepsilon^2}{2a^2(1+\varepsilon)}\right),\end{aligned}$$

hence, from Assumption A5 and for $\varepsilon = \varepsilon_0\sqrt{\frac{\log n}{n\phi(h_K)}}$ and $r_n = h_K\sqrt{\frac{\phi(h_K)}{n}}$, we get

$$\begin{aligned}\mathbf{P}\left(\max_{1\leq k\leq l_n}\left|\tilde{\ell}_n(x_{k_0(x)}) - \mathbf{E}[\tilde{\ell}_n(x_{k_0(x)})]\right| > \varepsilon\right) &\leq \frac{2M_0}{r_n}exp\left\{-\frac{\varepsilon_0^2\log n}{2\left(1+\varepsilon_0\left(\frac{\log n}{n\phi(h_K)}\right)^{1/2}\right)}\right\}\\ &\leq \frac{2M_0 n^{\frac{1}{2}}}{h_K(\phi(h_K))^{\frac{1}{2}}}n^{-c\varepsilon_0^2} \leq \frac{2M_0}{h_K(\phi(h_K))^{\frac{1}{2}}n^{\frac{3}{2}}}n^{2-c\varepsilon_0^2}\\ &\leq \frac{2M_0}{(n\phi(h_K))^{\frac{1}{2}}(nh_K)}n^{2-c\varepsilon_0^2},\end{aligned}$$

for $\varepsilon_0 > \sqrt{\frac{3}{c}}$, the term $n^{2-c\varepsilon_0^2}$ is the general term of a convergent series. Therefore, Borel-Cantelli's lemma yields that

$$I_2 = \max_{1\leq k\leq l_n}\left|\tilde{\ell}_n(x_{k_0(x)}) - \mathbf{E}[\tilde{\ell}_n(x_{k_0(x)})]\right| = O\left(\sqrt{\frac{\log n}{n\phi(h_K)}}\right) \text{ a.s. as } n\to\infty. \tag{25}$$

Thus, the results (23), (24) and (25) conclude the proof.

Lemma 9.3. *Under Assumptions A1, A2 and A5, we have*

$$\mathbf{E}[\tilde{\ell}_n(x)] = 1$$

Proof. Using the conditional expectation, we have

$$\begin{aligned}
\mathbf{E}[\tilde{\ell}_n(x)] &= \frac{1}{\mathbb{E}\left[K\left(\frac{d(x,X_1)}{h_K}\right)\right]}\mathbf{E}\left[\frac{\theta\delta_1}{L(Z_1)\overline{G}(Z_1)}K\left(\frac{d(x,X_1)}{h_K}\right)\right] \\
&= \frac{1}{\mathbb{E}\left[K\left(\frac{d(x,X_1)}{h_K}\right)\right]}\mathbf{E}\left[K\left(\frac{d(x,X_1)}{h_K}\right)\mathbf{E}\left[\frac{\theta\delta_1}{L(Z_1)\overline{G}(Z_1)}|X_1\right]\right] \\
&= \frac{1}{\mathbb{E}\left[K\left(\frac{d(x,X_1)}{h_K}\right)\right]}\int\int K\left(\frac{d(x,u)}{h_K}\right)\frac{\theta}{L(t)\overline{G}(t)}\mathbf{v}_1^*(t|u)d\mathbf{P}^{X_1}(u)dt.
\end{aligned}$$

From (6), we have

$$\begin{aligned}
\mathbf{E}[\tilde{\ell}_n(x)] &= \frac{1}{\mathbb{E}\left[K\left(\frac{d(x,X_1)}{h_K}\right)\right]}\int\int K\left(\frac{d(x,u)}{h_K}\right)f(t|u)d\mathbb{P}^{X_1}(u)dt \\
&= \frac{1}{\mathbb{E}\left[K\left(\frac{d(x,X_1)}{h_K}\right)\right]}\int K\left(\frac{d(x,u)}{h_K}\right)d\mathbb{P}^{X_1}(u)
\end{aligned}$$

and since $\mathbb{E}\left[K\left(\frac{d(x,X_1)}{h_K}\right)\right] = \int K\left(\frac{d(x,u)}{h_K}\right)d\mathbb{P}^{X_1}(u)$ we have the expected result.

Lemma 9.4. *Under the same assumptions as those of Lemma 9.1, we have*

$$\sup_{x\in\Omega}\sup_{a_V\le y\le b}\left|\Psi_n(x,y)-\tilde{\Psi}_n(x,y)\right| = O\left(\sqrt{\frac{\log\log n}{n}}\right) \text{ a.s. as } n\to\infty.$$

Proof. As H is a distribution function, and from Lemma 9.1 we obtain

$$\left|\Psi_n(x,y)-\tilde{\Psi}_n(x,y)\right| \le \left|\ell_n(x)-\tilde{\ell}_n(x)\right| = O\left(\sqrt{\frac{\log\log n}{n}}\right) \text{ a.s.}$$

This conclude the proof.

Lemma 9.5. *Under Assumptions A1, A2, A3 and A4, we have*

$$\sup_{x\in\Omega}\sup_{a_V\le y\le b}\left|\mathbf{E}[\tilde{\Psi}_n(x,y)]-F(y|x)\right| = O\left(h_K^{\beta_1}\right)+O\left(h_H^{\beta_2}\right) \text{ a.s. as } n\to\infty.$$

Proof. From (6), we have

$$\begin{aligned}
\mathbf{E}[\tilde{\Psi}_n(x,y)] &= \frac{1}{\mathbb{E}\left[K\left(\frac{d(x,X_1)}{h_K}\right)\right]}\mathbf{E}\left[\frac{\theta\delta_1}{L(Z_1)\overline{G}(Z_1)}K\left(\frac{d(x,X_1)}{h_K}\right)H\left(\frac{y-Z_1}{h_H}\right)\right] \\
&= \frac{1}{\mathbb{E}\left[K\left(\frac{d(x,X_1)}{h_K}\right)\right]}\int\int\frac{\theta}{L(t)\overline{G}(t)}K\left(\frac{d(x,u)}{h_K}\right)H\left(\frac{y-t}{h_H}\right)\mathbf{v}_1^*(t|u)d\mathbf{P}^{X_1}(u)dt \\
&= \frac{1}{\mathbb{E}\left[K\left(\frac{d(x,X_1)}{h_K}\right)\right]}\int K\left(\frac{d(x,u)}{h_K}\right)\left(\int H\left(\frac{y-t}{h_H}\right)f(t|u)dt\right)d\mathbb{P}^{X_1}(u) \\
&= \frac{1}{\mathbb{E}\left[K\left(\frac{d(x,X_1)}{h_K}\right)\right]}\mathbb{E}\left[K\left(\frac{d(x,X_1)}{h_K}\right)\mathbb{E}\left[H\left(\frac{y-Y_1}{h_H}\right)|X_1\right]\right].
\end{aligned}$$

Remark that

$$\mathbb{E}\left[H\left(\frac{y-Y_1}{h_H}\right)|X_1\right] = \int_{\mathbb{R}} H\left(\frac{y-t}{h_H}\right)dF(t|X_1),$$

then, using an integration by part we get

$$\mathbb{E}\left[H\left(\frac{y-Y_1}{h_H}\right)|X_1\right] = \frac{1}{h_H}\int_{\mathbb{R}} H'\left(\frac{y-t}{h_H}\right)F(t|X_1)dt,$$

by applying a change of variables, we obtain

$$\mathbb{E}\left[H\left(\frac{y-Y_1}{h_H}\right)|X_1\right] = \int_{\mathbb{R}} H'(s)F(y-sh_H|X_1)ds.$$

Thus, we have

$$\left|\mathbb{E}\left[H\left(\frac{y-Y_1}{h_H}\right)|X_1\right] - F(y|x)\right| \leq \int_{\mathbb{R}} H'(s)\left|F(y-sh_H|X_1) - F(y|x)\right| ds,$$

and, the condition A3 allows us to write

$$\left|\mathbb{E}\left[H\left(\frac{y-Y_1}{h_H}\right)X_1\right] - F(y|x)\right| \leq c_7\int_{\mathbb{R}} H'(s)\left(d(x,X_1)^{\beta_1} + |sh_H|^{\beta_2}\right) ds.$$

Making use of A4, we get

$$\begin{aligned}\left|\mathbf{E}(\tilde{\Psi}_n(x,y)) - F(y|x)\right| &= \left|\frac{1}{\mathbb{E}\left[K\left(\frac{d(x,X_1)}{h_K}\right)\right]}\mathbb{E}\left[K\left(\frac{d(x,X_1)}{h_K}\right)\right] \times \left(\mathbb{E}\left[H\left(\frac{y-Y_1}{h_H}\right)|X_1\right] - F(y|x)\right)\right| \\ &\leq c_7\left(h_K^{\beta_1} + h_H^{\beta_2}\right) = O(h_K^{\beta_1} + h_H^{\beta_2}).\end{aligned}$$

Which completes the proof of the Lemma.

Lemma 9.6. *Under Assumptions A1-A5, we have*

$$\sup_{x\in\Omega}\sup_{a_V\leq y\leq b}\left|\tilde{\Psi}_n(x,y) - \mathbf{E}[\tilde{\Psi}_n(x,y)]\right| = O\left(\sqrt{\frac{\log n}{n\phi(h_K)}}\right) \text{ a.s. as } n\to\infty.$$

Proof. We follow step by step the proof of Lemma 9.2, replacing $\tilde{\ell}_n(x)$ by $\tilde{\Psi}_n(x,y)$ and the fact that H is a distribution function yields the result.

Lemma 9.7. *Under Assumptions A1, A2 and A5, we have*

$$\sup_{x\in\Omega}|1-\ell_n(x)| = O\left(\sqrt{\frac{\log n}{n\phi(h_K)}}\right) \text{ a.s. as } n\to\infty.$$

Proof. We have the following decomposition:

$$\ell_n(x) = \left(\ell_n(x) - \tilde{\ell}_n(x)\right) + \left(\tilde{\ell}_n(x) - \mathbf{E}[\tilde{\ell}_n(x)]\right) + \mathbf{E}[\tilde{\ell}_n(x)].$$

By Lemma 9.3, $\mathbf{E}[\tilde{\ell}_n(x)] = 1$, then we have

$$\sup_{x\in\Omega}|1-\ell_n(x)| \leq \sup_{x\in\Omega}|\tilde{\ell}_n(x) - \ell_n(x)| + \sup_{x\in\Omega}|\mathbf{E}[\tilde{\ell}_n(x)] - \tilde{\ell}_n(x)|. \tag{26}$$

Lemmas 9.1 and 9.2 imply that

$$\sup_{x\in\Omega}|1-\ell_n(x)| = O\left(\sqrt{\frac{\log n}{n\phi(h_K)}}\right) \text{ a.s. as } n\to\infty.$$

Corollary 9.1. *Under assumptions of Lemma 9.7 we have*

$$\sum_{n=1}^{\infty} \mathbf{P}\left(\inf_{x\in\Omega} |\ell_n(x)| < \frac{1}{2}\right) < \infty.$$

Proof. Clearly, we have

$$\inf_{x\in\Omega} |\ell_n(x)| < \frac{1}{2} \Rightarrow \exists x \in \Omega, \text{ such that } 1 - \ell_n(x) > \frac{1}{2} \Rightarrow \sup_{x\in\Omega} |1 - \ell_n(x)| > \frac{1}{2}.$$

We deduce from Lemma 9.7 and Borel-Cantelli's lemma that

$$\mathbf{P}\left(\inf_{x\in\Omega} |\ell_n(x)| < \frac{1}{2}\right) \leq \mathbf{P}\left(\sup_{x\in\Omega} |1 - \ell_n(x)| > \frac{1}{2}\right),$$

thus, we obtain

$$\sum_{n=1}^{\infty} \mathbf{P}\left(\inf_{x\in\Omega} |\ell_n(x)| < \frac{1}{2}\right) < \infty.$$

The proof of the Corrollary is ended.

Acknowledgments

The authors are indebted to the Editor and the anonymous reviewers for their pertinent and constructive comments, which have significantly enhanced the quality of this manuscript. Many thanks also to Dr. Hassiba Benseradj, from the Department of Mathematics at UMBB, for their availability and valuable suggestions, which were essential to the development of the practical part of this work.

References

Adjoudj, L. and Tatachak, A. (2019). Conditional quantile estimation for truncated and associated data. *Communications in Statistics. Theory and Methods*, 48(18):4598–4641.

Bosq, D. (2000). *Linear processes in function spaces, theory and applications*. Springer-Verlag, New York.

Boucetta, A., Guessoum, Z., and Ould Said, E. (2024). *Strong convergence of a nonparametric relative error regression estimator under missing data with functional predictors.*

Dabo-Niang, S. and Laksaci, A. (2012). Nonparametric quantile regression estimation for functional dependent data. *Communications in statistics. Theory and Methods*, 41(7):1254–1268.

Djelladj, W. and Tatachak, A. (2019). Strong uniform consistency of a kernel conditional quantile estimator for censored and associated data. *Hacettepe Journal of Mathematics and Statistics*, 48(1):290–311.

Farraty, F., Laksaci, A., and Vieu, P. (2006). Estimating some characteristics of the conditional distribution in nonparametric functional models. *Statistical Inference for Stochastic Processes*, 9:47–76.

Ferraty, F. (2011). *Recent advances in functional data analysis and related topics*. Springer-Verlag, New York.

Ferraty, F., Rabhi, A., and Vieu, P. (2005). Conditional quantiles for dependent functional data with application to the climatic el niño phenomenon. *Sankhyā*, 67(2):378–398.

Ferraty, F. and Vieu, P. (2006). *Nonparametric functional data analysis*. Springer-Verlag, New York.

Fetitah, O., Almanjahie, I. M., Attouch, M. K., and Righi, A. (2020). Strong convergence of the functional nonparametric relative error regression estimator under right censoring. *Mathematica Slovaca*, 70(6):1469–1490.

Gijbels, I. and Wang, J. L. (1993). Strong representations of the survival function estimator for truncated and censored data with applications. *Journal of Multivariate Analysis*, 47(2):210–229.

Gu, M. G. and Lai, T. L. (1990). Functional laws of the iterated logarithm for the product-limit estimator of a distribution function under random censorship or truncation. *The Annals of Probability*, 18(1):160–189.

He, S. and Yang, G. (1998). Estimation of the truncation probability in the random truncation model. *The Annals of Statistics*, 26(3):1011–1027.

Helal, N. and Ould Said, E. (2016). Kernel conditional quantile estimator under left truncation for functional regressor. *Opuscula Mathematica*, 36(1):25–48.

Horrigue, W. and Ould Said, E. (2011). Strong uniform consistency of a nonparametric estimator of a conditional quantile for censored dependent data and functional regressors. *Random Operators and Stochastic Equations*, 19:131–156.

Horvath, L. and Kokoszka, P. (2012). *Inference for functional data with applications*. Springer-Verlag, New York.

Iglesias-Pérez, M. C. (2003). Strong representation of a conditional quantile function estimator with truncated and censored data. *Statistics & probability letters*, 65(2):79–91.

Kadiri, N. L., Rabhi, A., and Bouchentouf, A. A. (2018). Strong uniform consistency rates of conditional quantile estimation in the single functional index model under random censorship. *Dependence Modeling*, 6(1):197–227.

Kaplan, E. L. and Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, 53:457–481.

Lai, T. L. and Ying, Z. (1991). Estimating a distribution with truncated and censored data. *The Annals of Statistics*, 19(1):417–442.

Lemdani, M., Ould Said, E., and Poulin, N. (2009). Asymptotic properties of a conditional quantile estimator with randomly truncated data. *Journal of Multivariate Analysis*, 100(3):546–559.

Leulmi, S. (2019). Local linear estimation of the conditional quantile for censored data and functional regressors. *Communications in Statistics. Theory and Methods*, 50(14):3286–3300.

Liang, H. Li, D. and Miao, T. (2015). Conditional quantile estimation with truncated, censored and dependent data. *Annals of Mathematics*, 36(6):969–990.

Lynden-Bell, D. (1971). A method of allowing for known observational selection in small sample applied to 3cr quasars. *Monthly Notices of the Royal Astronomy Society*, 155(1):95–118.

Messaci, Nemouchi, N., Ouassou, I., and Rachdi, M. (2015). Local polynomial modelling of the conditional quantile for functional data. *Statistical Methods & Applications*, 24:597–622.

Ould Said, E. and Sadki, O. (2008). Prediction via the conditional quantile for right censorship model. *Far East Journal. Theoretical Statistics*, 25:145–179.

Pîrjan, A., Oprea, S. V., Căruțașu, G., Petroșanu, D. M., Bâra, A., and Coculescu, C. (2017). Devising hourly forecasting solutions regarding electricity consumption in the case of commercial center type consumers. *Energies*, 10(11):1727.

Ramsay, J. O. and Silverman, B. W. (2005). *Functional data analysis*. Springer, New York.

Roussas, G. G. (1969). Nonparametric estimation of the transition distribution function of a markov process. *The Annals of Mathematical Statistics*, 40(4):1386–1400.

Stone, C. (1977). Consistent nonparametric regression. *The annals of statistics*, 5:595–620.

Tsai, W. Y., Jewell, N. P., and Wang, M. C. (1987). A note on the product-limit estimator under right censoring and left truncation. *Biometrika*, 74(4):883–886.

Xiang, X. (1996). A kernel estimator of a conditional quantile. *Journal of Multivariate Analysis*, 59(2):206–216.

Zhou, Y. (1996). A note on the tjw product-limit estimator for truncated and censored data. *Statistics & Probability Letters*, 26(4):381–387.

Zhou, Y. and Yip, P. F. (1999). A strong representation of the product limit estimator for left truncated and right censored data. *Journal of Multivariate Analysis*, 69(2):261–280.