Chin. Ann. Math.
36B(6), 2015, 969–990
DOI: 10.1007/s11401-015-0905-9

Chinese Annals of Mathematics, Series B

Conditional Quantile Estimation with Truncated, Censored and Dependent Data*

Hanying LIANG[1] Deli LI[2] Tianxuan MIAO[2]

Abstract This paper deals with the conditional quantile estimation based on left-truncated and right-censored data. Assuming that the observations with multivariate covariates form a stationary α-mixing sequence, the authors derive the strong convergence with rate, strong representation as well as asymptotic normality of the conditional quantile estimator. Also, a Berry-Esseen-type bound for the estimator is established. In addition, the finite sample behavior of the estimator is investigated via simulations.

Keywords Berry-Esseen-type bound, Conditional quantile estimator, Strong representation, Truncated and censored data, α-mixing
2000 MR Subject Classification 62N02, 62G20

1 Introduction

In medical follow-up or engineering life testing studies, one may not be able to observe the variable of interest, referred to hereafter as the lifetime. In this paper, we focus on the lifetime data with multivariate covariates which are subject to both left truncation and right censorship. Let $(\mathbf{X}, Y, T, W)$ be a random vector, where Y is the random lifetime with the distribution function (df) F, T is the random left truncation time with the df L, W denotes the random right censoring time with df G and $\mathbf{X}$ is an $\mathbb{R}^d$-valued random vector of covariates related with Y. Assume that $\mathbf{X}$ admits the df $M(\cdot)$ and density $m(\cdot)$.

In the random left truncation and the right censoring (LTRC) model, one observes $(\mathbf{X}, Z, T, \delta)$ if $Z \geq T$, where $Z = \min(Y, W)$ and $\delta = I(Y \leq W)$; when $Z < T$, nothing is observed. Clearly, if Y is independent of W, then Z has df $H = 1 - (1 - F)(1 - G)$. Taking $\theta = P(T \leq Z)$, then necessarily, we assume $\theta > 0$. If $(\mathbf{X}_i, Z_i, T_i, \delta_i)$, for $i = 1, 2, \cdots, n$, is a stationary random sample from $(\mathbf{X}, Z, T, \delta)$ which one observes, then $(T_i \leq Z_i, \forall i)$. Without loss of generality, we assume that Y, T and W are nonnegative random variables as are usual in survival analysis. Following the idea of Iglesias-Pérez and González-Manteiga [13], we define a generalized product-limit estimator (GPLE) of the conditional distribution function $F(y|\mathbf{x})$ of Y, given

Manuscript received September 19, 2013. Revised May 26, 2014.
[1]Department of Mathematics, Tongji University, Shanghai 200092, China.
E-mail: hyliang@tongji.edu.cn
[2]Department of Mathematical Sciences, Lakehead University, 955 Oliver Road, Thunder Bay, Ontario, Canada P7B 5E1. E-mail: dli@lakeheadu.ca; tmiao@lakeheadu.ca
*This work was supported by the National Natural Science Foundation of China (No. 11271286), the Specialized Research Fund for the Doctor Program of Higher Education of China (No. 20120072110007), and a grant from the Natural Sciences and Engineering Research Council of Canada.

$\mathbf{X} = \mathbf{x}$ for the LTRC data by

$$\widehat{F}_n(y|\mathbf{x}) = 1 - \prod_{i=1}^{n} \Big(1 - \frac{I(Z_i \le y)\delta_i B_{ni}(\mathbf{x})}{\sum_{j=1}^{n} I(T_j \le Z_i \le Z_j) B_{nj}(\mathbf{x})}\Big),$$

where $B_{ni}(\mathbf{x}) = \frac{K(\frac{\mathbf{x}-\mathbf{X}_i}{h_n})}{\sum_{j=1}^{n} K(\frac{\mathbf{x}-\mathbf{X}_j}{h_n})}$, $K(\cdot)$ denotes a kernel function on $\mathbb{R}^d$, and $0 < h_n \to 0$ is the bandwith parameter. Note that the GPLE reduces to the estimator for left truncated data when there is no right censoring ($\delta = 1$, $Z = Y$) (see [1]), and to the estimator for right censored data when there is no left truncation ($T = 0$) (see [2, 9]).

One characteristic of the conditional distribution function $F(y|\mathbf{x})$ that is of interest is the conditional quantile function. It plays an important role in various statistical applications, especially in data modeling, reliability, and medical studies. Let $\xi_p(\mathbf{x}) = \inf\{y : F(y|\mathbf{x}) \ge p\}$ for $p \in (0, 1)$ be the conditional quantile function of $F(y|\mathbf{x})$. We focus here on estimating $\xi_p(\mathbf{x})$ based on the LTRC data. A natural estimator of $\xi_p(\mathbf{x})$ is given by $\widehat{\xi}_{pn}(\mathbf{x}) = \inf\{y : \widehat{F}_n(y|\mathbf{x}) \ge p\}$. Iglesias-Pérez [14] first derived an almost sure representation and the asymptotic normality of $\widehat{\xi}_{pn}(\mathbf{x})$ under i.i.d. assumptions and the case $d = 1$.

Asymptotic properties for different quantile estimators with censored and/or truncated data have been studied by many authors. In the absence of covariables, representations of the product-limit quantile estimator were obtained by Lo and Singh [21] for censored data, by Gürler et al. [10] for truncated data; asymptotic normality and a Berry-Esseen-type bound for the kernel quantile estimator were derived by Zhou et al. [31] for jointly censored and truncated data. In the presence of covariables, we cite the representations derived by Dabrowska [6] and Van Keilegom and Veraverbeke [28] for conditional quantile estimators with censored data, the strong uniform convergence with rate for a kernel estimator of the conditional quantile established by Ould-Saïd [23] for censored data, and the asymptotic properties of the kernel conditional quantile estimator for the left-truncated model studied by Lemdani et al. [16]. In all of these papers, it is assumed that the observations are independent.

However, the dependent data scenario is an important one in a number of applications with survival data. When sampling clusters of individuals (family members, or repeated measurements on the same individual, for example), lifetimes within clusters are typically correlated (see [3, 15]). There has been some literature devoted to the study of the conditional quantile estimation under dependence. To mention some examples, Cai [4] investigated the asymptotic normality of a weighted Nadaraya-Watson conditional quantile estimator for the α-mixing time series. Honda [12] dealt with α-mixing processes and proved the uniform convergence and asymptotic normality of an estimate of $\xi_p(\mathbf{x})$ for the case $d = 1$ using the local polynomial fitting method. Ferraty et al. [8] considered quantile regression under dependence when the conditioning variable is infinite dimensional. A nice extension of the conditional quantile process theory to set-indexed processes under strong mixing was establish in [26]. Ould-Saïd et al. [24] recently discussed strong uniform convergence with rate of the kernel conditional quantile estimator with left-truncated and dependent data. Liang and de Uña-Álvarez [18] proved the strong uniform convergence and asymptotic normality for the kernel estimator of the conditional quantile under censored and dependent assumptions. The asymptotic normality of the conditional quantile estimator with auxiliary information for left-truncated and dependent data was discussed by Liang and de Uña-Álvarez [19]. However, to the best of our knowledge, the asymptotic properties of the conditional quantile estimator with dependent data for the LTRC model have not yet been investigated.

In this paper, we study the strong convergence with rate, strong representation as well as asymptotic normality of the conditional quantile estimator $\widehat{\xi}_{pn}(\mathbf{x})$ when the observations with multivariate covariates form a stationary α-mixing sequence. Also, a Berry-Esseen-type bound for the estimator is established; this result is new, even for independent data. The finite sample behavior of the estimator is also investigated via simulations.

In the sequel, $\{(\mathbf{X}_i, Z_i, T_i, \delta_i), 1 \le i \le n\}$ is assumed to be a stationary α-mixing sequence of random vectors. Recall that a sequence $\{\zeta_k, k \ge 1\}$ is said to be α-mixing if the α-mixing coefficient

$$\alpha(n) :\stackrel{\text{def}}{=} \sup_{k\ge1}\sup\{|P(AB) - P(A)P(B)| : A \in \mathcal{F}_{n+k}^{\infty}, B \in \mathcal{F}_1^k\}$$

converges to zero as $n \to \infty$, where $\mathcal{F}_l^m = \sigma\{\zeta_l, \zeta_{l+1}, \cdots, \zeta_m\}$ denotes the σ-algebra generated by $\zeta_l, \zeta_{l+1}, \cdots, \zeta_m$ with $l \le m$. Among various mixing conditions used in the literature, α-mixing is reasonably weak and known to be fulfilled for many stochastic processes including many time series models. Withers [29] derived conditions under which a linear process is α-mixing. In fact, under very mild assumptions, linear autoregressive and more generally bilinear time series models are strongly mixing with mixing coefficients decaying exponentially, i.e., $\alpha(k) = O(\rho^k)$ for some $0 < \rho < 1$. See [7, p. 99], for more details. We mention that α-mixing has been used in applications with clustered survival data; see, for instance, Cai and Kim [5].

In the sequel, for any df $Q(y) = P(\eta \le y)$, we denote its density function by $q(y)$, and the left and right support endpoints by $a_Q = \inf\{y : Q(y) > 0\}$ and $b_Q = \sup\{y : Q(y) < 1\}$, respectively. For $\mathbf{x} \in \mathbb{R}^d$, define $\theta(\mathbf{x}) = P(T \le Z \mid \mathbf{X} = \mathbf{x})$,

$$C(y|\mathbf{x}) = P(T \le y \le Z \mid \mathbf{X} = \mathbf{x}, T \le Z) \quad \text{and} \quad H_1^*(y|\mathbf{x}) = P(Z \le y, \delta = 1 \mid \mathbf{X} = \mathbf{x}, T \le Z).$$

Also, we define $Q(y|\mathbf{x}) = P(\eta \le y \mid \mathbf{X} = \mathbf{x})$ and $Q^*(y) = P(\eta \le y \mid T \le Z)$, while their density functions are denoted by $q(y|\mathbf{x})$ and $q^*(y)$, respectively. Thus $M^*(\mathbf{x}) = P(\mathbf{X} \le \mathbf{x} \mid T \le Z)$, and its density function is $m^*(\mathbf{x})$.

Remark 1.1 It is easy to verify that $m^*(\mathbf{x}) = \theta^{-1}\theta(\mathbf{x})m(\mathbf{x})$. Assuming that Y, T and W are conditionally independent at $\mathbf{X} = \mathbf{x}$, and $F(\cdot|\mathbf{x})$ and $G(\cdot|\mathbf{x})$ are continuous, then $C(y|\mathbf{x}) = \theta^{-1}(\mathbf{x})L(y|\mathbf{x})(1 - G(y|\mathbf{x}))(1 - F(y|\mathbf{x})) = \theta^{-1}(\mathbf{x})L(y|\mathbf{x})(1 - H(y|\mathbf{x}))$, and $H_1^*(y|\mathbf{x}) = \theta^{-1}(\mathbf{x})\int_0^y L(t|\mathbf{x})(1 - G(t|\mathbf{x}))f(t|\mathbf{x})\mathrm{d}t$, which gives $h_1^*(y|\mathbf{x}) = \theta^{-1}(\mathbf{x})L(y|\mathbf{x})(1 - G(y|\mathbf{x}))f(y|\mathbf{x})$.

Define estimators of $H_1^*(\cdot|\mathbf{x})$, $C(\cdot|\mathbf{x})$ and $m^*(\mathbf{x})$ respectively as follows:

$$\widehat{H}_{1n}^*(y|\mathbf{x}) = \sum_{i=1}^n I(Z_i \le y, \delta_i = 1)B_{ni}(\mathbf{x}), \quad \widehat{C}_n(y|\mathbf{x}) = \sum_{i=1}^n I(T_i \le y \le Z_i)B_{ni}(\mathbf{x})$$

and $\widehat{m}_n^*(\mathbf{x}) = \frac{1}{nh_n^d}\sum_{i=1}^n K(\frac{\mathbf{x}-\mathbf{X}_i}{h_n})$.

The rest of this paper is organized as follows. The main results are described in Section 2. A simulation study is presented in Section 3. All proofs are given in Section 4. Some preliminary lemmas, which are used in the proofs of the main results, are collected in Appendix.

2 The Main Results

Throughout this paper, $\mathbf{x} = (x_1, \cdots, x_d) \in \mathbb{R}^d$. For $(i,j) = (i_1, \cdots, i_d, j) \in \mathbb{N}^{d+1}$, put $f^{(i,j)}(y|\mathbf{x}) := \frac{\partial^{i_1+\cdots+i_d+j}}{\partial x_1^{i_1}\cdots\partial x_d^{i_d}\partial y^j} f(y|\mathbf{x})$. Let $C, C_1, \cdots$ and $c_0, c_1, \cdots$ denote generic finite positive constants, whose values may change from line to line, and let $\Phi(u)$ stand for the standard normal distribution function and $[t]$ be the integer part of t. The notation $A_n = O(B_n)$ means $|A_n| \le C|B_n|$, and $U(\mathbf{x})$ represents a neighborhood of $\mathbf{x}$. Let I be a compact set of $\mathbb{R}^d$, which is included in $D = \{\mathbf{x} \in \mathbb{R}^d \mid m(\mathbf{x}) > 0, \theta(\mathbf{x}) > 0\}$. Set $\mathbf{e} = (e_1, \cdots, e_d)$ for small $e_i > 0$, and $I_{\mathbf{e}} = \{\mathbf{x} \pm \mathbf{e}, \mathbf{x} \in I\}$ with $\inf\limits_{\mathbf{x}\in I_{\mathbf{e}}} \{m(\mathbf{x}), \theta(\mathbf{x})\} \ge \delta_0 > 0$.

Throughout this paper, we assume that $\alpha(k) = O(k^{-\lambda})$ for some $\lambda > 0$. We first list the following basic assumptions:

(A1) (i) $K(\cdot)$ is a Lipschitz-continuous density function with compact support on $\mathbb{R}^d$;

(ii) $\int_{\mathbb{R}^d} x_1^{i_1} \cdots x_d^{i_d} K(\mathbf{x})\mathrm{d}\mathbf{x} = 0$ for non-negative integers $i_1, \cdots, i_d$ with $i_1 + \cdots + i_d \le r_0 - 1$.

(A1') (i) $K(\cdot)$ is a bounded density function with compact support on $\mathbb{R}^d$;

(ii) $\int_{\mathbb{R}^d} x_1^{i_1} \cdots x_d^{i_d} K(\mathbf{x})\mathrm{d}\mathbf{x} = 0$ for non-negative integers $i_1, \cdots, i_d$ with $i_1 + \cdots + i_d \le r_0 - 1$.

(A2) (i) Y, T and W are conditionally independent at $\mathbf{X} = \mathbf{s}$ for $\mathbf{s} \in I_{\mathbf{e}}$;

(ii) τ_1 and τ_2 are two real numbers such that $a_{L(\cdot|\mathbf{x})} < \tau_1 \le \tau_2 < b_{H(\cdot|\mathbf{x})}$ and $a_{L(\cdot|\mathbf{x})} < a_{H(\cdot|\mathbf{x})}$ for $\mathbf{x} \in I$.

(A2') (i) Y, T and W are conditionally independent at $\mathbf{X} = \mathbf{s}$ for $\mathbf{s} \in U(\mathbf{x})$;

(ii) τ_1 and τ_2 are two real numbers such that $a_{L(\cdot|\mathbf{x})} < \tau_1 \le \tau_2 < b_{H(\cdot|\mathbf{x})}$ and $a_{L(\cdot|\mathbf{x})} < a_{H(\cdot|\mathbf{x})}$.

(A3) The first r_0 partial derivatives of functions $\theta(\mathbf{s})$ and $m(\mathbf{s})$ are bounded for $\mathbf{s} \in I_{\mathbf{e}}$, and the first r_0 partial derivatives with respect to $\mathbf{s}$ of functions $L(y|\mathbf{s})$, $G(y|\mathbf{s})$, $F(y|\mathbf{s})$, $l(y|\mathbf{s})$, $g(y|\mathbf{s})$ and $f(y|\mathbf{s})$ are bounded for $(\mathbf{s}, y) \in I_{\mathbf{e}} \times \mathbb{R}$.

(A3') The first r_0 partial derivatives of functions $\theta(\mathbf{s})$ and $m(\mathbf{s})$ are bounded for $\mathbf{s} \in U(\mathbf{x})$, and the first r_0 partial derivatives with respect to $\mathbf{s}$ of functions $L(y|\mathbf{s})$, $G(y|\mathbf{s})$, $F(y|\mathbf{s})$, $l(y|\mathbf{s})$, $g(y|\mathbf{s})$ and $f(y|\mathbf{s})$ are bounded for $(\mathbf{s}, y) \in U(\mathbf{x}) \times \mathbb{R}$.

(A4) For all integers $j \ge 1$, the joint conditional density $v_j^*(\cdot,\cdot)$ of $\mathbf{X}_1$ and $\mathbf{X}_{j+1}$ exists on $\mathbb{R}^d \times \mathbb{R}^d$ and satisfies $v_j^*(\mathbf{s}_1, \mathbf{s}_2) \le C_1$ for $(\mathbf{s}_1, \mathbf{s}_2) \in I_{\mathbf{e}} \times I_{\mathbf{e}}$.

(A4') For all integers $j \ge 1$, the joint conditional density $v_j^*(\cdot,\cdot)$ of $\mathbf{X}_1$ and $\mathbf{X}_{j+1}$ exists on $\mathbb{R}^d \times \mathbb{R}^d$ and satisfies $v_j^*(\mathbf{s}_1, \mathbf{s}_2) \le C_1$ for $(\mathbf{s}_1, \mathbf{s}_2) \in U(\mathbf{x}) \times U(\mathbf{x})$.

(A5) (i) $\sum\limits_{n=1}^{\infty} h_n^{-2d} \big(\frac{nh_n^d}{\ln(n)}\big)^{-\frac{\lambda-(2+d)}{2}} < \infty$; (ii) $h_n^{-2d} \big(\frac{nh_n^d}{\ln(n)}\big)^{-\frac{\lambda-(2+d)}{2}} = O(1)$.

(A5') (i) $\sum\limits_{n=1}^{\infty} h_n^{-d} \big(\frac{nh_n^d}{\ln(n)}\big)^{-\frac{\lambda-2}{2}} < \infty$; (ii) $h_n^{-d} \big(\frac{nh_n^d}{\ln(n)}\big)^{-\frac{\lambda-2}{2}} = O(1)$.

(B1) $\sup\limits_{(\mathbf{s},y)\in U(\mathbf{x})\times\mathbb{R}} \{|l^{(0,1)}(y|\mathbf{s})|, |g^{(0,1)}(y|\mathbf{s})|, |f^{(0,1)}(y|\mathbf{s})|\} < \infty$.

(B2) For all integers $j \ge 1$, the joint conditional density $f_j^*(\cdot,\cdot,\cdot,\cdot)$ of $(\mathbf{X}_1, \mathbf{X}_{j+1}, H_1^*(Z_1), H_1^*(Z_{j+1}))$ exists on $\mathbb{R}^d \times \mathbb{R}^d \times [0,1] \times [0,1]$ and satisfies $f_j^*(\mathbf{s}_1, \mathbf{s}_2, y_1, y_2) \le C_3$ for $(\mathbf{s}_1, \mathbf{s}_2, y_1, y_2) \in U(\mathbf{x}) \times U(\mathbf{x}) \times [0,1] \times [0,1]$.

(B3) (i) $\sum\limits_{n=2}^{\infty} \frac{n}{\ln(n)} \big(\frac{\ln(n)}{nh_n^d}\big)^{\frac{\lambda-2}{4}} < \infty$; (ii) $\frac{n}{\ln(n)} \big(\frac{\ln(n)}{nh_n^d}\big)^{\frac{\lambda-2}{4}} = O(1)$.

Remark 2.1 (a) (i) and (ii) in (B3) imply (i) and (ii) in (A5'), respectively.

(b) Similar conditions as (A1)–(A3), (A1')–(A3') and (B1) have been used commonly in the literature, see, e.g., Iglesias-Pérez and González-Manteiga [13] in the cases $d = 1$ and $r_0 = 2$. The role of condition $a_{L(\cdot|\mathbf{x})} < \tau_1 \le \tau_2 < b_{H(\cdot|\mathbf{x})}$ in (A2) and (A2') is to avoid the problem that the conditional function $C(y|\mathbf{x})$ may vanish. The conditions (A3) and (A3') allow us to apply Taylor expansions in the proofs to determine the order of convergence of the estimators. Conditions (A4), (A4') and (B2) are mainly technical, which are employed to simplify the calculations of covariances in the proofs, and are otherwise redundant for the independent setting.

(c) Assumptions (A5), (A5') and (B3) imply restrictions on the degree of dependence of the observable sequence; as we discuss now, the message under these assumptions is that one must prevent strongly dependent data. Indeed, all these conditions are satisfied by appropriately choosing the bandwidth h_n when λ is large enough. Note that, if the exponential decay $\alpha(k) = O(\rho^k)$ for some $0 < \rho < 1$, which has been used by some authors (see [7]), we replace $\alpha(k) = O(k^{-\lambda})$, and then λ can be arbitrarily large.

In order to give the strong convergence with rates of $\widehat{\xi}_{pn}(\mathbf{x})$, we need the following additional assumptions:

(D1) For each fixed $p \in (0,1)$, the function $\xi_p(\mathbf{x})$ satisfies that for any $\epsilon > 0$ and any function $\eta_p(\mathbf{x})$, there exists $\beta > 0$ such that $\sup\limits_{\mathbf{x}\in I}|\xi_p(\mathbf{x}) - \eta_p(\mathbf{x})| \geq \epsilon$ implies that $\sup\limits_{\mathbf{x}\in I}|F(\xi_p(\mathbf{x})|\mathbf{x}) - F(\eta_p(\mathbf{x})|\mathbf{x})| \geq \beta$.

(D1') For each fixed $p \in (0,1)$, the function $\xi_p(\mathbf{x})$ satisfies that for any $\epsilon > 0$ and any function $\eta_p(\mathbf{x})$, there exists $\beta > 0$ such that $|\xi_p(\mathbf{x}) - \eta_p(\mathbf{x})| \geq \epsilon$ implies that $|F(\xi_p(\mathbf{x})|\mathbf{x}) - F(\eta_p(\mathbf{x})|\mathbf{x})| \geq \beta$.

(D2) There exists $\gamma_1 > 0$ such that $\inf\limits_{(\mathbf{x},y)\in I\times[\tau_1,\tau_2]} f(y|\mathbf{x}) \geq \gamma_1$.

(D2') There exists $\gamma_1 > 0$ such that $\inf\limits_{y\in[\tau_1,\tau_2]} f(y|\mathbf{x}) \geq \gamma_1$.

Theorem 2.1 *Let $\alpha(n) = O(n^{-\lambda})$ for some $\lambda > 2$.*

(a) *Let $0 < p_0 \leq p_1 < 1$ be such that $\tau_1 < \xi_{p_0}(\mathbf{x}) \leq \xi_{p_1}(\mathbf{x}) < \tau_2$ for all $\mathbf{x} \in I$. Suppose that* (A1)–(A4) *and* (A5)(i) *are satisfied. If* (D1) *holds, then* $\lim\limits_{n\to\infty}\sup\limits_{\mathbf{x}\in I}|\widehat{\xi}_{pn}(\mathbf{x}) - \xi_p(\mathbf{x})| = 0$ a.s. *for $p \in [p_0, p_1]$. If* (D2) *holds, then*

$$\sup_{\mathbf{x}\in I}\sup_{p_0\leq p\leq p_1}|\widehat{\xi}_{pn}(\mathbf{x}) - \xi_p(\mathbf{x})| = O\Big(\max\Big\{\Big(\frac{\ln(n)}{nh_n^d}\Big)^{\frac{1}{2}}, h_n^{r_0}\Big\}\Big) \text{ a.s.}$$

(b) *Let $\mathbf{x} \in D$ and $0 < p_0 \leq p_1 < 1$ with $\tau_1 < \xi_{p_0}(\mathbf{x}) \leq \xi_{p_1}(\mathbf{x}) < \tau_2$. Suppose that* (A1′)–(A4′) *and* (A5′)(i) *are satisfied. If* (D1′) *holds, then* $\lim\limits_{n\to\infty}\widehat{\xi}_{pn}(\mathbf{x}) = \xi_p(\mathbf{x})$ a.s. *for $p \in [p_0, p_1]$. If* (D2′) *holds, then*

$$\sup_{p_0\leq p\leq p_1}|\widehat{\xi}_{pn}(\mathbf{x}) - \xi_p(\mathbf{x})| = O\Big(\max\Big\{\Big(\frac{\ln(n)}{nh_n^d}\Big)^{\frac{1}{2}}, h_n^{r_0}\Big\}\Big) \text{ a.s.}$$

In order to formulate the strong representation and asymptotic normality of $\widehat{\xi}_{pn}(\mathbf{x})$, we need to impose the following additional assumptions:

(E1) $f^{(0,1)}(y|\mathbf{x})$ is bounded for $y \in [\tau_1, \tau_2]$.

(E2) The sequence $\alpha(n)$ satisfies for positive integers $q := q_n$ that $q = o((nh_n^d)^{\frac{1}{2}})$ and $\lim\limits_{n\to\infty}(nh_n^{-d})^{\frac{1}{2}}\alpha(q) = 0$.

(E3) $nh_n^{d+2r_0} \to 0$, $\frac{(\ln(n))^3}{nh_n^d} \to 0$.

Theorem 2.2 *Set* $\xi(Z,T,\delta,y,\mathbf{x}) = \frac{I(Z\leq y,\delta=1)}{C(Z|\mathbf{x})} - \int_0^y \frac{I(T\leq t\leq Z)}{C^2(t|\mathbf{x})}\mathrm{d}H_1^*(t|\mathbf{x})$. *Let $\alpha(n) = O(n^{-\lambda})$ for some $\lambda > 6$, let $\mathbf{x} \in D$ and $0 < p_0 \leq p_1 < 1$ with $\tau_1 < \xi_{p_0}(\mathbf{x}) \leq \xi_{p_1}(\mathbf{x}) < \tau_2$, and let*

$p \in [p_0, p_1]$. *Suppose that* (A1′)–(A4′), (B1)–(B2), (D2′) *and* (E1) *are satisfied. If* $\tau_1 < a_{H(\cdot|\mathbf{x})}$ *and* $\frac{nh_n^{d+2r_0}}{\ln(n)} = O(1)$, *then*

$$\begin{aligned}\widehat{\xi}_{pn}(\mathbf{x}) - \xi_p(\mathbf{x}) &= \frac{p - \widehat{F}_n(\xi_p(\mathbf{x})|\mathbf{x})}{f(\xi_p(\mathbf{x})|\mathbf{x})} + R_{n1}(\xi_p(\mathbf{x})|\mathbf{x}) \\ &= -\frac{1-p}{f(\xi_p(\mathbf{x})|\mathbf{x})} \sum_{i=1}^{n} B_{ni}(\mathbf{x})\xi(Z_i, T_i, \delta_i, \xi_p(\mathbf{x}), \mathbf{x}) + R_{n2}(\xi_p(\mathbf{x})|\mathbf{x}),\end{aligned}$$

where for $i = 1, 2$, $\sup_{p_0 \le p \le p_1} |R_{ni}(\xi_p(\mathbf{x})|\mathbf{x})| = O\big(\frac{\ln(n)}{(nh_n^d)^{\frac{3}{4}}}\big)$ *a.s. when* (B3)(i) *holds;* $\sup_{p_0 \le p \le p_1} \cdot |R_{ni}(\xi_p(\mathbf{x})|\mathbf{x})| = O_p\big(\frac{\ln(n)}{(nh_n^d)^{\frac{3}{4}}}\big)$ *when* (A5′)(i) *and* (B3)(ii) *hold.*

Theorem 2.3 *Let* $\alpha(n) = O(n^{-\lambda})$ *for some* $\lambda > 6$, *let* $\mathbf{x} \in D$ *and* $0 < p_0 \le p_1 < 1$ *with* $\tau_1 < \xi_{p_0}(\mathbf{x}) \le \xi_{p_1}(\mathbf{x}) < \tau_2$, *and let* $p \in [p_0, p_1]$. *Suppose that* (A1′)–(A4′), (A5′)(i), (B1)–(B2), (B3)(ii), (D2′) *and* (E1)–(E3) *are satisfied. If* $\tau_1 < a_{H(\cdot|\mathbf{x})}$, *then for* $p \in [p_0, p_1]$, *we have*

$$(nh_n^d)^{\frac{1}{2}}(\widehat{\xi}_{pn}(\mathbf{x}) - \xi_p(\mathbf{x})) \xrightarrow{\mathcal{D}} N(0, \Delta^2(p|\mathbf{x})),$$

where $\Delta^2(p|\mathbf{x}) = \frac{\theta(1-p)^2}{m(\mathbf{x})f^2(\xi_p(\mathbf{x})|\mathbf{x})} \int_0^{\xi_p(\mathbf{x})} \frac{f(t|\mathbf{x})\mathrm{d}t}{L(t|\mathbf{x})(1-G(t|\mathbf{x}))(1-F(t|\mathbf{x}))^2} \int_{\mathbb{R}^d} K^2(\mathbf{s})\mathrm{d}\mathbf{s}$.

In order to give a Berry-Esseen-type bound for $\widehat{\xi}_{pn}(\mathbf{x})$ which will assess the quality of the normal approximation in Theorem 2.3, we need the following additional assumption.

(Q) $p := p_n$ and $q := q_n$ are positive integers such that $p + q \le n$, $\frac{p}{n} \to 0$ and $qp^{-1} \to 0$.

Put $\gamma_{1n} = \big(\frac{\ln^4(n)}{nh_n^d}\big)^{\frac{1}{6}}$, $\gamma_{2n} = (nh_n^{d+2r_0})^{\frac{1}{2}}$, $\gamma_{3n} = qp^{-1}h_n^{-\frac{d\delta}{2+\delta}}u(p)$, $\gamma_{4n} = (p/n)^{\beta}h_n^{-\frac{d\delta(1+\beta)}{2+\delta}}$, $\gamma_{5n} = np^{-1}\alpha(q)$ and $u(p) = \sum_{i=p}^{\infty}[\alpha(i)]^{\frac{\delta}{2+\delta}}$.

Theorem 2.4 *Let* $\alpha(n) = O(n^{-\lambda})$ *for some* $\lambda > \frac{2+\delta}{\delta}$ *with* $0 < \delta \le \frac{2}{5}$, *and let* $\mathbf{x} \in D$ *and* $0 < p_0 \le p_1 < 1$ *with* $\tau_1 < \xi_{p_0}(\mathbf{x}) \le \xi_{p_1}(\mathbf{x}) < \tau_2$. *Suppose that* (A1′)–(A4′), (B1)–(B2), (B3)(i), (D2′), (E1) *and* (Q) *are satisfied, and that* $\tau_1 < a_{H(\cdot|\mathbf{x})}$. *If* $\gamma_{in} \to 0$ $(i = 1, \cdots, 5)$ *for* $0 < 2\beta < \delta$ *and* $\beta \le \frac{\delta\lambda-(2+\delta)}{2\lambda+(2+\delta)}$, *then for* $\frac{20\lambda-1}{\lambda(10\lambda-1)} \le \rho < 1$ *and* $p \in [p_0, p_1]$, *we have*

$$\begin{aligned}&\sup_u \Big| P\Big(\frac{(nh_n^d)^{\frac{1}{2}}(\widehat{\xi}_{pn}(\mathbf{x}) - \xi_p(\mathbf{x}))}{\Delta(p|\mathbf{x})} \le u\Big) - \Phi(u)\Big| \\ &= O\big(h_n + (qp^{-1})^{\frac{1}{3}} + (pn^{-1})^{\frac{1}{3}} + h_n^{\frac{d(1-\rho)}{3}} + \gamma_{1n} + \gamma_{2n} + \gamma_{3n}^{\frac{1}{3}} + \gamma_{4n} + \gamma_{5n}^{\frac{1}{4}}\big).\end{aligned}$$

Remark 2.2 The assumptions $\gamma_{in} \to 0$ $(i = 1, \cdots, 5)$ in Theorem 2.4 can be satisfied by appropriate choice of h_n, p and q, when λ is large enough (note that if we replace $\alpha(n) = O(n^{-\lambda})$ by the exponential decay $\alpha(n) = O(\rho^n)$ for some $0 < \rho < 1$, then λ can be arbitrarily large). In particular, choosing $p = [n^s]$ and $q = [n^{2s-1}]$ for some $\frac{1}{2} < s < 1 - \frac{d\delta(1+\beta)}{\beta(2+\delta)(d+2r_0)}$, and $h_n^d = n^{-\frac{1}{M}}$ for $\frac{\delta(1+\beta)}{\beta(2+\delta)(1-s)} < M < 1 + \frac{2r_0}{d}$, if $\lambda > \max\{\frac{(2s-1)(2+\delta)}{s\delta} + \frac{1}{Ms}, \frac{1-s}{2s-1}, \frac{2+\delta}{\delta}, 10 + \frac{8}{M-1}\}$, then $\gamma_{in} \to 0$ $(i = 1, \cdots, 5)$, $qp^{-1} \to 0$, $pn^{-1} \to 0$, and (B3)(i) holds.

3 Simulation Study

In this section, we investigate with simulated data the finite sample performance of the proposed estimator $\widehat{\xi}_{pn}(\mathbf{x})$ with $p = 0.5$ in the case $d = 1$. In particular, we calculate the mean squared error (MSE), plot the Boxplots of the estimator $\widehat{\xi}_{pn}(x)$ at $x = 0.5$, and explore the estimator's graphical fit to the true underlying curve. We also investigate the goodness-of-fit to the normal distribution which is expected from our theoretical results in Section 2. At the same time, we check the influence of the dependence of the observations on the estimator. In order to obtain an α-mixing observed sequence $\{X_i, Z_i, T_i, \delta_i\}$, we generate the observed data as follows.

(1) Drawing of the first observation $(X_1, Z_1, T_1, \delta_1)$ in the final sample.

Step 1 Draw $e_1 \sim N(0, 1)$, and take $X_1 = 0.5e_1$;

Step 2 Compute Y_1 and W_1, respectively, from the model $Y_1 = \sin(\pi X_1) + \phi_1(1 + 0.3\cos(\pi X_1))\epsilon_1$, and $W_1 = \sin(\pi X_1) + 0.5\phi_2(1 + 0.3\cos(\pi X_1)) + \phi_3(1 + 0.3\cos(\pi X_1))\tilde{\epsilon}_1$, where both ϵ_1 and $\tilde{\epsilon}_1$ are $N(0, 1)$ random variables, ϵ_1, $\tilde{\epsilon}_1$ and X_1 are mutually independent, and ϕ_i $(i = 1, 2, 3)$ are chosen (see below) to control the percentage of censoring. Take $Z_1 = \min(Y_1, W_1)$, $\delta_1 = I(Y_1 \le W_1)$;

Step 3 Draw independently $T_1 \sim N(\mu, 1)$, where μ is adapted in order to get different values of θ. If $Z_1 < T_1$, reject the datum $(X_1, Z_1, T_1, \delta_1)$ and go back to Step 2; do this until $Z_1 \ge T_1$.

(2) Drawing of the second observation $(X_2, Z_2, T_2, \delta_2)$ in the final sample.

Step 4 Draw X_2 according to the AR(1) model $X_2 = \rho X_1 + 0.5e_2$, where $e_2 \sim N(0, 1)$ is independent of X_1, and $|\rho| < 1$ is some constant, which is chosen to control the dependence of the observations;

Step 5 Compute Y_2 and W_2, respectively, from the model $Y_2 = \sin(\pi X_2) + \phi_1(1 + 0.3\cos(\pi X_2))\epsilon_2$, and $W_2 = \sin(\pi X_2) + 0.5\phi_2(1 + 0.3\cos(\pi X_2)) + \phi_3(1 + 0.3\cos(\pi X_2))\tilde{\epsilon}_2$, where both ϵ_2 and $\tilde{\epsilon}_2$ are $N(0, 1)$ random variables, and ϵ_2, $\tilde{\epsilon}_2$ and X_2 are mutually independent. Take $Z_2 = \min(Y_2, W_2)$, and $\delta_2 = I(Y_2 \le W_2)$;

Step 6 Draw independently $T_2 \sim N(\mu, 1)$. If $Z_2 < T_2$, reject the datum $(X_2, Z_2, T_2, \delta_2)$ and go back to Step 5; do this until $Z_2 \ge T_2$.

By replicating the process (2) above, we generate the observed data $(X_i, Z_i, T_i, \delta_i)$, $i = 1, \cdots, n$. The generating process shows that $X_i = \rho X_{i-1} + 0.5e_i$, $Y_i = \sin(\pi X_i) + \phi_1(1 + 0.3\cos(\pi X_i))\epsilon_i$, $W_i = \sin(\pi X_i) + 0.5\phi_2(1 + 0.3\cos(\pi X_i)) + \phi_3(1 + 0.3\cos(\pi X_i))\tilde{\epsilon}_i$, $Z_i = \min(Y_i, W_i)$, and $\delta_i = I(Y_i \le W_i)$, where $e_i \sim N(0, 1)$, $\epsilon_i \sim N(0, 1)$, $\tilde{\epsilon}_i \sim N(0, 1)$, and $T_i \sim N(\mu, 1)$; everything is distributed conditionally on $Z_i \ge T_i$. Note that the α-mixing property of the observable X_i is immediately transferred to the $(X_i, Z_i, T_i, \delta_i)$. Also note that $Y|_{X=x} \sim N(\sin(\pi x), \phi_1^2(1 + 0.3\cos(\pi x))^2)$, which shows that the conditional quantile function $\xi_{0.5}(x) = \sin(\pi x)$. For the proposed estimators, we employ the kernel $K(x) = \frac{15}{16}(1 - x^2)^2 I(|x| \le 1)$.

In addition, the parameters ϕ_i $(i = 1, 2, 3)$ allow for the control of the percentage of censoring (PC) which is given by

$$\mathrm{PC} = P(Y_i > W_i \mid X_i = x) = 1 - \Phi\Big(\frac{0.5\phi_2}{\sqrt{\phi_1^2 + \phi_3^2}}\Big)$$

$$\stackrel{\phi_1=\phi_3=0.3}{=} 1-\Phi\Big(\frac{5\sqrt{2}}{6}\phi_2\Big) = \begin{cases} 10\%, & \text{when} \quad \phi_2 = 1.087, \\ 15\%, & \text{when} \quad \phi_2 = 0.8796, \\ 30\%, & \text{when} \quad \phi_2 = 0.445. \end{cases}$$

In the simulation below, we take $\phi_1 = \phi_3 = 0.3$.

3.1 Consistency

In this subsection, we draw random samples with sample sizes n =200, 350 and 500, respectively, and ρ =0.1, 0.3 and 0.5, respectively, from the above model. In Table 1, we report the MSE of the estimator $\widehat{\xi}_{pn}(x)$ with $p = 0.5$ at $x = 0.5$, for several truncation rates, percentage of censoring, and choice of bandwidth based on $M = 1000$ replications.

Table 1 Mean squared errors (MSEs) of $\widehat{\xi}_{pn}(x)$ with $p = 0.5$ at $x = 0.5$ along $M = 1000$ Monte Carlo trials, for several truncation rates and percentage of censoring (PC).

ρ	θ	PC	n	$h_n = 0.3$	$h_n = 0.35$	$h_n = 0.4$
0.1	30%	10%	200	0.7569×10^{-2}	1.0060×10^{-2}	1.3393×10^{-2}
			350	0.6330×10^{-2}	0.8714×10^{-2}	1.2988×10^{-2}
			500	0.5296×10^{-2}	0.8240×10^{-2}	1.2279×10^{-2}
		15%	200	0.7652×10^{-2}	1.0360×10^{-2}	1.3612×10^{-2}
			350	0.6173×10^{-2}	0.8544×10^{-2}	1.2915×10^{-2}
			500	0.5267×10^{-2}	0.8128×10^{-2}	1.1961×10^{-2}
		30%	200	0.8009×10^{-2}	1.0503×10^{-2}	1.3777×10^{-2}
			350	0.6018×10^{-2}	0.8430×10^{-2}	1.1996×10^{-2}
			500	0.5227×10^{-2}	0.7935×10^{-2}	1.1153×10^{-2}
	60%	10%	200	0.7349×10^{-2}	0.9612×10^{-2}	1.2418×10^{-2}
			350	0.5631×10^{-2}	0.8356×10^{-2}	1.1839×10^{-2}
			500	0.5183×10^{-2}	0.8090×10^{-2}	1.1085×10^{-2}
		15%	200	0.7542×10^{-2}	0.9762×10^{-2}	1.2800×10^{-2}
			350	0.5549×10^{-2}	0.8325×10^{-2}	1.1601×10^{-2}
			500	0.5096×10^{-2}	0.7807×10^{-2}	1.1025×10^{-2}
		30%	200	0.7628×10^{-2}	1.0204×10^{-2}	1.2688×10^{-2}
			350	0.5473×10^{-2}	0.8273×10^{-2}	1.1345×10^{-2}
			500	0.5028×10^{-2}	0.7543×10^{-2}	1.0329×10^{-2}
	90%	10%	200	0.7149×10^{-2}	0.9355×10^{-2}	1.1969×10^{-2}
			350	0.5471×10^{-2}	0.7639×10^{-2}	1.1414×10^{-2}
			500	0.4975×10^{-2}	0.7385×10^{-2}	1.0816×10^{-2}
		15%	200	0.7291×10^{-2}	0.9478×10^{-2}	1.2218×10^{-2}
			350	0.5509×10^{-2}	0.7886×10^{-2}	1.1160×10^{-2}
			500	0.5076×10^{-2}	0.7317×10^{-2}	1.0550×10^{-2}
		30%	200	0.7301×10^{-2}	0.9537×10^{-2}	1.2277×10^{-2}
			350	0.5722×10^{-2}	0.7453×10^{-2}	1.0503×10^{-2}
			500	0.4738×10^{-2}	0.7159×10^{-2}	0.9987×10^{-2}
0.3	90%	30%	200	0.7368×10^{-2}	0.9551×10^{-2}	1.2312×10^{-2}
			350	0.5830×10^{-2}	0.7719×10^{-2}	1.0861×10^{-2}
			500	0.5047×10^{-2}	0.7336×10^{-2}	1.0107×10^{-2}
0.5	90%	30%	200	0.7498×10^{-2}	0.9797×10^{-2}	1.2419×10^{-2}
			350	0.5983×10^{-2}	0.8019×10^{-2}	1.1099×10^{-2}
			500	0.5248×10^{-2}	0.7496×10^{-2}	1.0410×10^{-2}

From Table 1, it is seen that (i) the MSE decreases as the sample size n increases; (ii) the accuracy of the estimator is greatly affected by the choice of the bandwidth h_n, i.e., higher values for h_n give bad estimators; (iii) for the same sample size, the performance of the estimator is affected slightly by the percentage of truncated data $1-\theta$ and the percentage of censoring PC;

(iv) the values of the MSE become bigger as the dependence of the observations increases, i.e., the value of ρ increases.

In Figures 1–3, we plot the Boxplots of the MSE for the estimator $\widehat{\xi}_{pn}(x)$ with $p = 0.5$ and $h_n = 0.3$ at $x = 0.5$, along $M = 1000$ Monte Carlo trials, for $\theta = 90\%$, PC=30%, n =200, 350 and 500; $\theta = 90\%$, $n = 350$, PC=10%, 15% and 30%; PC=30%, $n = 350$, θ =30%, 60% and 90%, respectively.

Figure 1 shows that the quality of fit increases as the sample size n increases.

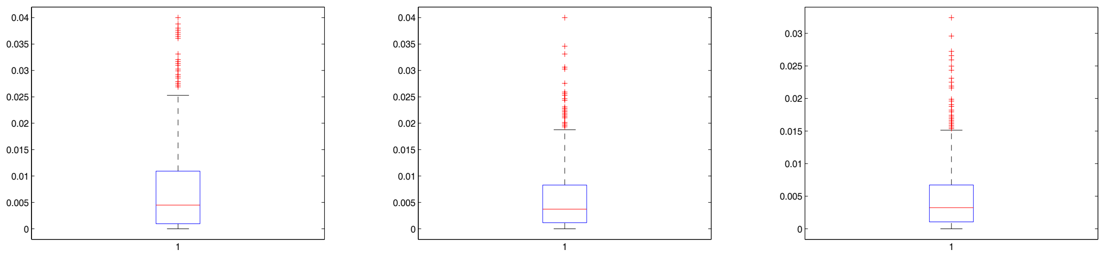

Figure 1 Boxplots of $\widehat{\xi}_{pn}(x)$ with $p = 0.5$ and $h_n = 0.3$ at $x = 0.5$ along $M = 1000$ Monte Carlo trials, for $\theta = 90\%$, PC=30%, n =200, 350 and 500, respectively.

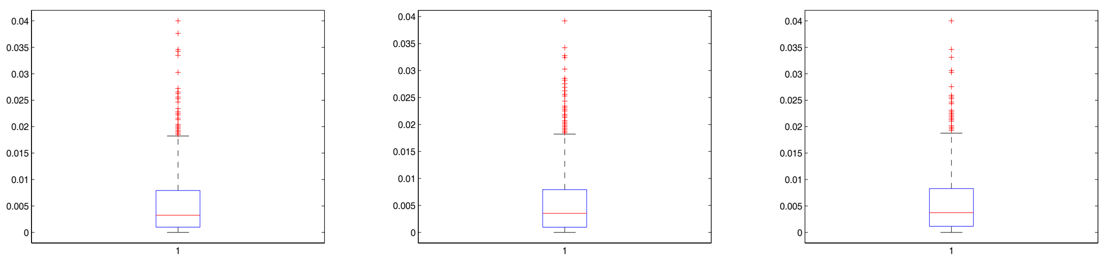

Figure 2 Boxplots of $\widehat{\xi}_{pn}(x)$ with $p = 0.5$ and $h_n = 0.3$ at $x = 0.5$ along $M = 1000$ Monte Carlo trials, for $\theta = 90\%$, $n = 350$, PC=10%, 15% and 30%, respectively.

From Figures 2–3, it can be seen that for the same sample size, the quality of the estimator does not seem to be affected by the percentage of truncated data $1 - \theta$ and the percentage of censoring.

In Figure 4, we plot the averages of the curves $\xi_p(x) = \sin(\pi x)$ and its estimator $\widehat{\xi}_{pn}(x)$ with $p = 0.5$ and $h_n = n^{-\frac{1}{5}}$ based on 100 replications for $\theta = 90\%$, PC=10%, n =150, 300 and 500, respectively. Figure 4 shows again that the quality of fit of the estimator increases as the sample size n increases.

3.2 Asymptotic normality

In this subsection, we examine how good is the asymptotic normality of the estimator $\widehat{\xi}_{pn}(\mathbf{x})$ with $p = 0.5$ at $x = 0.5$ by comparing the histograms and Normal-Probability-plots with the normal distribution. We draw M independent n-samples. In Figures 5–6, we plot the histograms and Normal-Probability-plots for $\theta = 90\%$, PC=10% and $h_n = n^{-\frac{1}{5}}$ based on

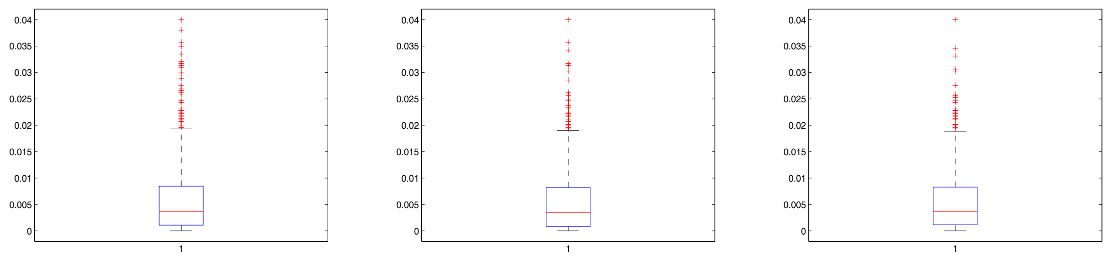

Figure 3 Boxplots of $\widehat{\xi}_{pn}(x)$ with $p = 0.5$ and $h_n = 0.3$ at $x = 0.5$ along $M = 1000$ Monte Carlo trials, for PC=30%, $n = 350$, θ =30%, 60% and 90%, respectively.

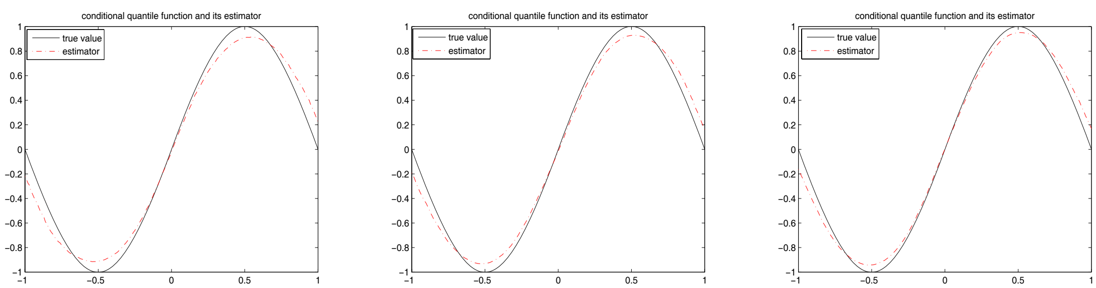

Figure 4 Function $\xi_p(x)$ and its estimator $\widehat{\xi}_{pn}(x)$ with $p = 0.5$ and $h_n = n^{-\frac{1}{5}}$ along $M = 100$ Monte Carlo trials, for $\theta = 90\%$, PC=10%, n =150, 300 and 500, respectively.

$M = 1000$ replications with sample sizes n =300 and 600, respectively. From Figures 5–6, it is seen that the sampling distribution of the estimator fits the normal distribution reasonably well; this fit being better when increasing the sample size.

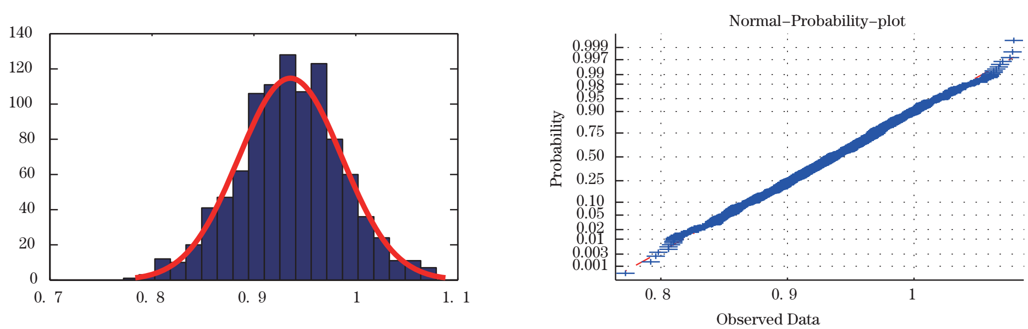

Figure 5 Histogram and Normal-Probability-plot of $\widehat{\xi}_{pn}(x)$ with $p = 0.5$ and $h_n = n^{-\frac{1}{5}}$ at $x = 0.5$ along $M = 1000$ Monte Carlo trials, for $\theta = 90\%$, PC=10%, n =300.

To study the influence of the dependence of the observations, we consider different degrees of dependence; specifically we choose in Figure 7, ρ =0.1, 0.3, 0.5, respectively, and plot the Normal-Probability-plots of $\widehat{\xi}_{pn}(x)$ with $p = 0.5$ and $h_n = n^{-\frac{1}{5}}$ at $x = 0.5$ based on $M = 1000$ replications with $\theta = 90\%$, PC=30%, and $n = 400$. Figure 7 shows that as the dependence of the observations increases, the quality of fit decreases.

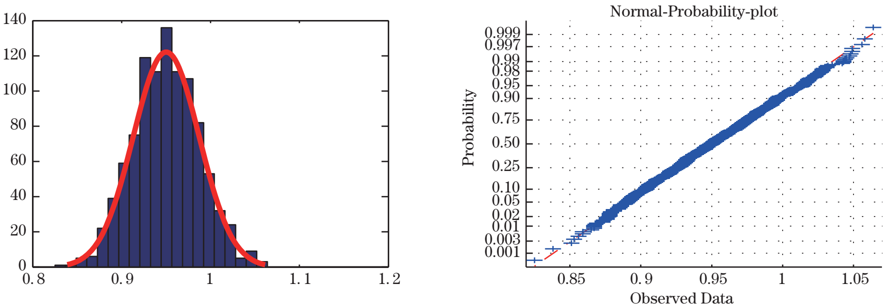

Figure 6 Histogram and Normal-Probability-plot of $\widehat{\xi}_{pn}(x)$ with $p = 0.5$ and $h_n = n^{-\frac{1}{5}}$ at $x = 0.5$ along $M = 1000$ Monte Carlo trials, for $\theta = 90\%$, PC=10%, n =600.

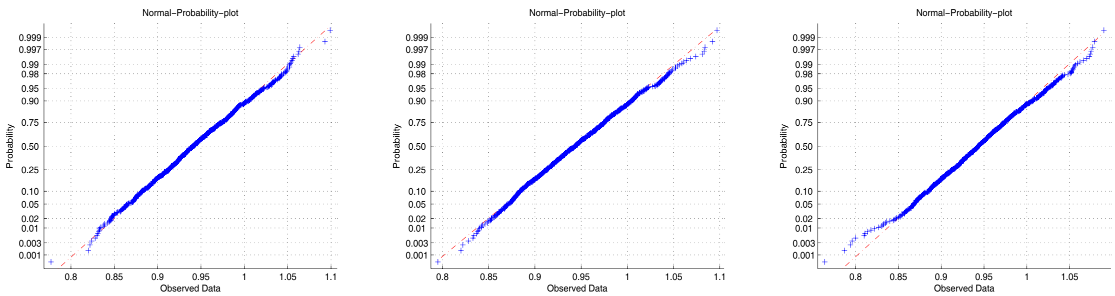

Figure 7 Normal-Probability-plots of $\widehat{\xi}_{pn}(x)$ with $p = 0.5$ and $h_n = n^{-\frac{1}{5}}$ at $x = 0.5$ along $M = 1000$ Monte Carlo trials, for $\theta = 90\%$, PC=30%, n =400, ρ =0.1, 0.3 and 0.5, respectively.

4 Proofs of the Main Results

Lemma 4.1 *Let* $\mathbf{x} \in D$ *and* $\alpha(n) = O(n^{-\lambda})$ *for some* $\lambda > 6$. *Suppose that conditions* (A1′)–(A4′) *and* (B1)–(B2) *hold, and that* $\tau_1 < a_{H(\cdot|\mathbf{x})}$ *and* $\frac{nh_n^{d+2r_0}}{\ln(n)} = O(1)$.

(a) *If* (B3)(i) *holds, then*

$$\sup_{s,t\in[\tau_1,\tau_2]:|s-t|\le c\left(\frac{\ln(n)}{nh_n^d}\right)^{\frac{1}{2}}} \left|[\widehat{F}_n(s|\mathbf{x}) - F(s|\mathbf{x})] - [\widehat{F}_n(t|\mathbf{x}) - F(t|\mathbf{x})]\right| = O\Big(\Big(\frac{\ln(n)}{nh_n^d}\Big)^{\frac{3}{4}}\Big) \text{ a.s.}$$

(b) *If* (B3)(ii) *holds, then*

$$\sup_{s,t\in[\tau_1,\tau_2]:|s-t|\le c\left(\frac{\ln(n)}{nh_n^d}\right)^{\frac{1}{2}}} \left|[\widehat{F}_n(s|\mathbf{x}) - F(s|\mathbf{x})] - [\widehat{F}_n(t|\mathbf{x}) - F(t|\mathbf{x})]\right| = O_p\Big(\Big(\frac{\ln(n)}{nh_n^d}\Big)^{\frac{3}{4}}\Big).$$

Proof of Lemma 4.1 We prove only (a); the proof of (b) is similar. From Lemma 5.2, we have

$$\begin{aligned}
&\left|[\widehat{F}_n(s|\mathbf{x}) - F(s|\mathbf{x})] - [\widehat{F}_n(t|\mathbf{x}) - F(t|\mathbf{x})]\right| \\
\le &\left|(1 - F(s|\mathbf{x}))\sum_{i=1}^n B_{ni}(\mathbf{x})\xi(Z_i, T_i, \delta_i, s, \mathbf{x}) - (1 - F(t|\mathbf{x}))\sum_{i=1}^n B_{ni}(\mathbf{x})\xi(Z_i, T_i, \delta_i, t, \mathbf{x})\right|
\end{aligned}$$

$$
\begin{aligned}
&+O\Big(\Big(\frac{\ln(n)}{nh_n^d}\Big)^{\frac34}\Big)\\
&\le \Big|(1-F(s|\mathbf{x}))\sum_{i=1}^n B_{ni}(\mathbf{x})[\xi(Z_i,T_i,\delta_i,s,\mathbf{x})-\xi(Z_i,T_i,\delta_i,t,\mathbf{x})]\Big|\\
&\quad+\Big|(F(s|\mathbf{x})-F(t|\mathbf{x}))\sum_{i=1}^n B_{ni}(\mathbf{x})\xi(Z_i,T_i,\delta_i,t,\mathbf{x})\Big|+O\Big(\Big(\frac{\ln(n)}{nh_n^d}\Big)^{\frac34}\Big)\\
&:=J_{1n}(s,t|\mathbf{x})+J_{2n}(s,t|\mathbf{x})+O\Big(\Big(\frac{\ln(n)}{nh_n^d}\Big)^{\frac34}\Big)\quad\text{a.s.}
\end{aligned}\tag{4.1}
$$

Note that

$$
\begin{aligned}
\sum_{i=1}^n B_{ni}(\mathbf{x})\xi(Z_i,T_i,\delta_i,s,\mathbf{x})&=\int_{a_{H(\cdot|\mathbf{x})}}^s\frac{\mathrm{d}\widehat H_{n1}^*(u|\mathbf{x})}{C(u|\mathbf{x})}-\int_{a_{H(\cdot|\mathbf{x})}}^s\frac{\widehat C_n(u|\mathbf{x})\mathrm{d}H_1^*(u|\mathbf{x})}{C^2(u|\mathbf{x})}\\
=\int_{a_{H(\cdot|\mathbf{x})}}^s\frac{\mathrm{d}(\widehat H_{n1}^*(u|\mathbf{x})-H_1^*(u|\mathbf{x}))}{C(u|\mathbf{x})}&-\int_{a_{H(\cdot|\mathbf{x})}}^s\frac{\widehat C_n(u|\mathbf{x})-C(u|\mathbf{x})}{C^2(u|\mathbf{x})}\mathrm{d}H_1^*(u|\mathbf{x}).
\end{aligned}\tag{4.2}
$$

Therefore

$$
\begin{aligned}
J_{1n}(s,t|\mathbf{x})&\le\Big|\Big[\frac{\widehat H_{n1}^*(y|\mathbf{x})-H_1^*(y|\mathbf{x})}{C(y|\mathbf{x})}\Big]_{y=t}^{y=s}\Big|+\Big|\int_t^s\frac{\widehat H_{n1}^*(u|\mathbf{x})-H_1^*(u|\mathbf{x})}{C^2(u|\mathbf{x})}\mathrm{d}C(u|\mathbf{x})\Big|\\
&\quad+\Big|\int_t^s\frac{\widehat C_n(u|\mathbf{x})-C(u|\mathbf{x})}{C^2(u|\mathbf{x})}\mathrm{d}H_1^*(u|\mathbf{x})\Big|\\
&:=J_{11n}(s,t|\mathbf{x})+J_{12n}(s,t|\mathbf{x})+J_{13n}(s,t|\mathbf{x}).
\end{aligned}\tag{4.3}
$$

Note that $C(y|\mathbf{x})=\theta^{-1}(\mathbf{x})L(y|\mathbf{x})(1-G(y|\mathbf{x}))(1-F(y|\mathbf{x}))$ and $H_1^*(y|\mathbf{x})=\theta^{-1}(\mathbf{x})\int_0^y L(t|\mathbf{x})(1-G(t|\mathbf{x}))f(t|\mathbf{x})\mathrm{d}t$. Then $C^{(0,1)}(y|\mathbf{x})$ and $h_1^*(y|\mathbf{x})$ are bounded for $y\in[\tau_1,\tau_2]$ from (A3'). Hence, using Lemmas 5.1–5.2, it follows that

$$
\begin{aligned}
&\sup_{s,t\in[\tau_1,\tau_2]:|s-t|\le c\left(\frac{\ln(n)}{nh_n^d}\right)^{\frac12}} J_{11n}(s,t|\mathbf{x})\\
&\le\sup_{s,t\in[\tau_1,\tau_2]:|s-t|\le c\left(\frac{\ln(n)}{nh_n^d}\right)^{\frac12}}\Big\{\Big|\frac{[\widehat H_{n1}^*(s|\mathbf{x})-H_1^*(s|\mathbf{x})-[\widehat H_{n1}^*(t|\mathbf{x})-H_1^*(t|\mathbf{x})]}{C(s|\mathbf{x})}\Big|\\
&\quad+\Big|\frac{[\widehat H_{n1}^*(t|\mathbf{x})-H_1^*(t|\mathbf{x})][C(t|\mathbf{x})-C(s|\mathbf{x})]}{C(s|\mathbf{x})C(t|\mathbf{x})}\Big|\Big\}\\
&\le O\Big(\Big(\frac{\ln(n)}{nh_n^d}\Big)^{\frac34}\Big)+C\sup_{s,t\in[\tau_1,\tau_2]:|s-t|\le c\left(\frac{\ln(n)}{nh_n^d}\right)^{\frac12}}|\widehat H_{n1}^*(t|\mathbf{x})-H_1^*(t|\mathbf{x})||s-t|\\
&=O\Big(\Big(\frac{\ln(n)}{nh_n^d}\Big)^{\frac34}\Big)+O\Big(\frac{\ln(n)}{nh_n^d}\Big)=O\Big(\Big(\frac{\ln(n)}{nh_n^d}\Big)^{\frac34}\Big)\quad\text{a.s.}
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
&\sup_{s,t\in[\tau_1,\tau_2]:|s-t|\le c\left(\frac{\ln(n)}{nh_n^d}\right)^{\frac12}} J_{12n}(s,t|\mathbf{x})\\
&\le C\sup_{|s-t|\le c\left(\frac{\ln(n)}{nh_n^d}\right)^{\frac12}}\ \sup_{\tau_1\le y\le\tau_2}|\widehat H_{n1}^*(y|\mathbf{x})-H_1^*(y|\mathbf{x})||s-t|=O\Big(\frac{\ln(n)}{nh_n^d}\Big)\quad\text{a.s.}
\end{aligned}
$$

and

$$\begin{aligned}&\sup_{s,t\in[\tau_1,\tau_2]:|s-t|\le c\left(\frac{\ln(n)}{nh_n^d}\right)^{\frac{1}{2}}} J_{13n}(s,t|\mathbf{x})\\
&\le C\sup_{|s-t|\le c\left(\frac{\ln(n)}{nh_n^d}\right)^{\frac{1}{2}}}\ \sup_{\tau_1\le y\le\tau_2}|\widehat{C}_n(y|\mathbf{x})-C(y|\mathbf{x})||s-t| = O\Big(\frac{\ln(n)}{nh_n^d}\Big)\ \text{ a.s.}\end{aligned}$$

Therefore, from (4.3) it follows that

$$\sup_{s,t\in[\tau_1,\tau_2]:|s-t|\le c\left(\frac{\ln(n)}{nh_n^d}\right)^{\frac{1}{2}}} J_{1n}(s,t|\mathbf{x}) = O\Big(\Big(\frac{\ln(n)}{nh_n^d}\Big)^{\frac{3}{4}}\Big)\ \text{ a.s.} \tag{4.4}$$

Using Lemma 5.1, from (4.2) one can verify that $\sum_{i=1}^{n} B_{ni}(\mathbf{x})\xi(Z_i,T_i,\delta_i,t,\mathbf{x}) = O\big(\big(\frac{\ln(n)}{nh_n^d}\big)^{\frac{1}{2}}\big)$ a.s. Therefore,

$$\begin{aligned}&\sup_{s,t\in[\tau_1,\tau_2]:|s-t|\le c\left(\frac{\ln(n)}{nh_n^d}\right)^{\frac{1}{2}}} J_{2n}(s,t|\mathbf{x})\\
&\le \sup_{s,t\in[\tau_1,\tau_2]:|s-t|\le c\left(\frac{\ln(n)}{nh_n^d}\right)^{\frac{1}{2}}} f(r|\mathbf{x})|s-t|\Big|\sum_{i=1}^{n} B_{ni}(\mathbf{x})\xi(Z_i,T_i,\delta_i,s,\mathbf{x})\Big|\\
&= O\Big(\frac{\ln(n)}{nh_n^d}\Big)\ \text{ a.s.,}\end{aligned} \tag{4.5}$$

where r is between s and t. Thus, the conclusion follows from (4.1) and (4.4)–(4.5).

Proof of Theorem 2.1 We prove only (a); the proof of (b) is similar. Observe that

$$\begin{aligned}|F(\widehat{\xi}_{pn}(\mathbf{x})|\mathbf{x})-F(\xi_p(\mathbf{x})|\mathbf{x})| &\le |\widehat{F}_n(\widehat{\xi}_{pn}(\mathbf{x})|\mathbf{x})-F(\widehat{\xi}_{pn}(\mathbf{x})|\mathbf{x})|\\
&\quad+|\widehat{F}_n(\widehat{\xi}_{pn}(\mathbf{x})|\mathbf{x})-F(\xi_p(\mathbf{x})|\mathbf{x})|.\end{aligned} \tag{4.6}$$

Since $F(\cdot|\mathbf{x})$ is continuous, $F(\xi_p(\mathbf{x})|\mathbf{x})=p$. Then from the definition of $\widehat{\xi}_{pn}(\mathbf{x})$, we have

$$\begin{aligned}&|\widehat{F}_n(\widehat{\xi}_{pn}(\mathbf{x})|\mathbf{x})-F(\xi_p(\mathbf{x})|\mathbf{x})|\\
&=\widehat{F}_n(\widehat{\xi}_{pn}(\mathbf{x})|\mathbf{x})-p\le \widehat{F}_n(\widehat{\xi}_{pn}(\mathbf{x})|\mathbf{x})-\widehat{F}_n(\widehat{\xi}_{pn}(\mathbf{x})^-|\mathbf{x})\\
&\le|\widehat{F}_n(\widehat{\xi}_{pn}(\mathbf{x})|\mathbf{x})-F(\widehat{\xi}_{pn}(\mathbf{x})|\mathbf{x})|+|F(\widehat{\xi}^-_{pn}(\mathbf{x})|\mathbf{x})-\widehat{F}_n(\widehat{\xi}^-_{pn}(\mathbf{x})|\mathbf{x})|,\end{aligned} \tag{4.7}$$

where $\widehat{F}_n(\widehat{\xi}_{pn}(\mathbf{x})^-|\mathbf{x})$ stands for the left-hand limit of $\widehat{F}_n(y|\mathbf{x})$ at $y=\widehat{\xi}_{pn}(\mathbf{x})^-$.

Since $0<p_0\le p_1<1$ with $\tau_1<\xi_{p_0}(\mathbf{x})\le\xi_{p_1}(\mathbf{x})<\tau_2$ for all $\mathbf{x}\in I$, $\tau_1<\xi_p(\mathbf{x})<\tau_2$ for $p\in[p_0,p_1]$. Hence, $\tau_1\le\widehat{\xi}_{pn}(\mathbf{x})\le\tau_2$ eventually from Lemma 5.1. Therefore, from (4.6)–(4.7) it follows that

$$|F(\widehat{\xi}_{pn}(\mathbf{x})|\mathbf{x})-F(\xi_p(\mathbf{x})|\mathbf{x})|\le 3\sup_{\mathbf{x}\in I}\ \sup_{\tau_1\le y\le\tau_2}|\widehat{F}_n(y|\mathbf{x})-F(y|\mathbf{x})|. \tag{4.8}$$

Then, the first part of the theorem follows from Lemma 5.1 and (D1). Note that

$$F(\widehat{\xi}_{pn}(\mathbf{x})|\mathbf{x})-F(\xi_p(\mathbf{x})|\mathbf{x})=(\widehat{\xi}_{pn}(\mathbf{x})-\xi_p(\mathbf{x}))f(\xi^*_{pn}(\mathbf{x})|\mathbf{x}),$$

where $\xi^*_{pn}(\mathbf{x})$ is between $\xi_p(\mathbf{x})$ and $\widehat{\xi}_{pn}(\mathbf{x})$. Then, by (4.8), we have

$$\sup_{\mathbf{x}\in I}\sup_{p_0\le p\le p_1}|\widehat{\xi}_{pn}(\mathbf{x})-\xi_p(\mathbf{x})|f(\xi^*_{pn}(\mathbf{x})|\mathbf{x})\le 3\sup_{\mathbf{x}\in I}\sup_{\tau_1\le y\le\tau_2}|\widehat{F}_n(y|\mathbf{x})-F(y|\mathbf{x})|.$$

Thus, the second part of the theorem follows from Lemma 5.1 and (D2).

Proof of Theorem 2.2 We prove only the conclusion in the case

$$\sup_{p_0\le p\le p_1}|R_{ni}(\xi_p(\mathbf{x})|\mathbf{x})|=O\Big(\frac{\ln(n)}{(nh_n^d)^{\frac{3}{4}}}\Big)\text{ a.s.}$$

for $i=1,2$.

Since $\frac{nh_n^{d+2r_0}}{\ln(n)}=O(1)$, (b) in Theorem 2.1 ensures that

$$\sup_{p_0\le p\le p_1}|\widehat{\xi}_{pn}(\mathbf{x})-\xi_p(\mathbf{x})|=O\Big(\Big(\frac{\ln(n)}{nh_n^d}\Big)^{\frac{1}{2}}\Big)\text{ a.s.}$$

Therefore, using a Taylor expansion, it follows that

$$\begin{aligned}
&\widehat{F}_n(\widehat{\xi}_{pn}(\mathbf{x})|\mathbf{x})-\widehat{F}_n(\xi_p(\mathbf{x})|\mathbf{x})\\
&=[F(\widehat{\xi}_{pn}(\mathbf{x})|\mathbf{x})-F(\xi_p(\mathbf{x})|\mathbf{x})]+\{[\widehat{F}_n(\widehat{\xi}_{pn}(\mathbf{x})|\mathbf{x})-F(\widehat{\xi}_{pn}(\mathbf{x})|\mathbf{x})]-[\widehat{F}_n(\xi_p(\mathbf{x})|\mathbf{x})-F(\xi_p(\mathbf{x})|\mathbf{x})]\}\\
&=f(\xi_p(\mathbf{x})|\mathbf{x})(\widehat{\xi}_{pn}(\mathbf{x})-\xi_p(\mathbf{x}))+\frac{f^{(0,1)}(\xi^*_{pn}(\mathbf{x})|\mathbf{x})}{2}(\widehat{\xi}_{pn}(\mathbf{x})-\xi_p(\mathbf{x}))^2+R_n^*(\xi_p(\mathbf{x})|\mathbf{x}),
\end{aligned}$$

where $\xi^*_{pn}(\mathbf{x})$ is between $\widehat{\xi}_{pn}(\mathbf{x})$ and $\xi_p(\mathbf{x})$, and $\sup_{p_0\le p\le p_1}|R_n^*(\xi_p(\mathbf{x})|\mathbf{x})|=O\big(\frac{\ln(n)}{(nh_n^d)^{\frac{3}{4}}}\big)$ a.s. by Lemma 4.1. Hence from $f(\xi_p(\mathbf{x})|\mathbf{x})>0$ and $\widehat{F}_n(\widehat{\xi}_{pn}(\mathbf{x})|\mathbf{x})=p$, we have

$$\begin{aligned}
\widehat{\xi}_{pn}(\mathbf{x})-\xi_p(\mathbf{x})&=\frac{p-\widehat{F}_n(\xi_p(\mathbf{x})|\mathbf{x})}{f(\xi_p(\mathbf{x})|\mathbf{x})}-\frac{2^{-1}f^{(0,1)}(\xi^*_{pn}(\mathbf{x})|\mathbf{x})(\widehat{\xi}_{pn}(\mathbf{x})-\xi_p(\mathbf{x}))^2+R_n^*(\xi_p(\mathbf{x})|\mathbf{x})}{f(\xi_p(\mathbf{x})|\mathbf{x})}\\
&=\frac{p-\widehat{F}_n(\xi_p(\mathbf{x})|\mathbf{x})}{f(\xi_p(\mathbf{x})|\mathbf{x})}+R_{n1}(\xi_p(\mathbf{x})|\mathbf{x}).
\end{aligned}\tag{4.9}$$

Note that (E1) implies that $f^{(0,1)}(\xi^*_{pn}(\mathbf{x})|\mathbf{x})$ is bounded. Then, according to $\sup_{p_0\le p\le p_1}|\widehat{\xi}_{pn}(\mathbf{x})-\xi_p(\mathbf{x})|^2=O\big(\frac{\ln(n)}{nh_n^d}\big)$ a.s. from (b) in Theorem 2.1, it follows that $\sup_{p_0\le p\le p_1}|R_{n1}(\xi_p(\mathbf{x})|\mathbf{x})|=O\big(\frac{\ln(n)}{(nh_n^d)^{\frac{3}{4}}}\big)$ a.s.

In addition, using Lemma 5.2 and $F(\xi_p(\mathbf{x})|\mathbf{x})=p$, we can write (4.9) as

$$\begin{aligned}
\widehat{\xi}_{pn}(\mathbf{x})-\xi_p(\mathbf{x})&=\frac{p-[\widehat{F}_n(\xi_p(\mathbf{x})|\mathbf{x})-F(\xi_p(\mathbf{x})|\mathbf{x})]-F(\xi_p(\mathbf{x})|\mathbf{x})}{f(\xi_p(\mathbf{x})|\mathbf{x})}+R_{n1}(\xi_p(\mathbf{x})|\mathbf{x})\\
&=-\frac{(1-F(\xi_p(\mathbf{x})|\mathbf{x}))\sum_{i=1}^n B_{ni}(\mathbf{x})\xi(Z_i,T_i,\delta_i,\xi_p(\mathbf{x}),\mathbf{x})+Q_n(\xi_p(\mathbf{x})|\mathbf{x})}{f(\xi_p(\mathbf{x})|\mathbf{x})}\\
&\quad+R_{n1}(\xi_p(\mathbf{x})|\mathbf{x})\\
&=-\frac{1-p}{f(\xi_p(\mathbf{x})|\mathbf{x})}\sum_{i=1}^n B_{ni}(\mathbf{x})\xi(Z_i,T_i,\delta_i,\xi_p(\mathbf{x}),\mathbf{x})+R_{n2}(\xi_p(\mathbf{x})|\mathbf{x})
\end{aligned}$$

and $\sup_{p_0\le p\le p_1}|R_{n2}(\xi_p(\mathbf{x})|\mathbf{x})| = O\big(\frac{\ln(n)}{(nh_n^d)^{\frac{3}{4}}}\big)$ a.s. from (D2').

Proof of Theorem 2.3 Note that $\frac{\ln^3(n)}{nh_n^d} \to 0$ implies that $(nh_n^d)^{\frac{1}{2}}\big(\frac{\ln(n)}{nh_n^d}\big)^{\frac{3}{4}} = \big[\frac{\ln^3(n)}{nh_n^d}\big]^{\frac{1}{4}} \to 0$. Then from Theorem 2.2 and $F(\xi_p(\mathbf{x})|\mathbf{x}) = p$, we have

$$(nh_n^d)^{\frac{1}{2}}(\widehat{\xi}_{pn}(\mathbf{x}) - \xi_p(\mathbf{x})) = -\frac{(nh_n^d)^{\frac{1}{2}}[\widehat{F}_n(\xi_p(\mathbf{x})|\mathbf{x}) - F(\xi_p(\mathbf{x})|\mathbf{x})]}{f(\xi_p(\mathbf{x})|\mathbf{x})} + o_p(1).$$

Therefore, from Lemma 5.3 it follows that $(nh_n^d)^{\frac{1}{2}}(\widehat{\xi}_{pn}(\mathbf{x}) - \xi_p(\mathbf{x})) \xrightarrow{\mathcal{D}} N(0, \Delta^2(p|\mathbf{x}))$.

Proof of Theorem 2.4 From Theorem 2.2 we write

$$\begin{aligned}
\frac{(nh_n^d)^{\frac{1}{2}}(\widehat{\xi}_{pn}(\mathbf{x}) - \xi_p(\mathbf{x}))}{\Delta(p|\mathbf{x})} &= -\frac{m^*(\mathbf{x})}{\widehat{m}_n^*(\mathbf{x})}\cdot\frac{1-p}{f(\xi_p(\mathbf{x})|\mathbf{x})\Delta(p|\mathbf{x})m^*(\mathbf{x})(nh_n^d)^{\frac{1}{2}}} \\
&\quad\times\sum_{i=1}^n\Big\{\Big[K\Big(\frac{\mathbf{x}-\mathbf{X}_i}{h_n}\Big)\xi(Z_i,T_i,\delta_i,\xi_p(\mathbf{x}),\mathbf{x}) \\
&\quad - E\Big(K\Big(\frac{\mathbf{x}-\mathbf{X}_i}{h_n}\Big)\xi(Z_i,T_i,\delta_i,\xi_p(\mathbf{x}),\mathbf{x})\Big)\Big] \\
&\quad + E\Big(K\Big(\frac{\mathbf{x}-\mathbf{X}_i}{h_n}\Big)\xi(Z_i,T_i,\delta_i,\xi_p(\mathbf{x}),\mathbf{x})\Big)\Big\} \\
&\quad + \frac{(nh_n^d)^{\frac{1}{2}}R_{n2}(\xi_p(\mathbf{x})|\mathbf{x})}{\Delta(p|\mathbf{x})} \\
&:= -\frac{m^*(\mathbf{x})}{\widehat{m}_n^*(\mathbf{x})}[I_{1n}(\mathbf{x}) + I_{2n}(\mathbf{x})] + I_{3n}(\mathbf{x}).
\end{aligned}$$

Let $\gamma_{1n}^* = \big(\frac{\ln(n)}{nh_n^d}\big)^{\frac{1}{3}}$. Then, using Lemma 5.4 we have

$$\begin{aligned}
&\sup_u\Big|P\Big(\frac{(nh_n^d)^{\frac{1}{2}}(\widehat{\xi}_{pn}(\mathbf{x}) - \xi_p(\mathbf{x}))}{\Delta(p|\mathbf{x})}\le u\Big) - \Phi(u)\Big| \\
&= \sup_u\Big|P\Big(\frac{m^*(\mathbf{x})}{\widehat{m}_n^*(\mathbf{x})}[I_{1n}(\mathbf{x}) + I_{2n}(\mathbf{x})] - I_{3n}(\mathbf{x}) < -u\Big) - \Phi(-u)\Big| \\
&\le \sup_u|P(I_{1n}(\mathbf{x}) < u) - \Phi(u)| + C(\gamma_{1n}^* + \gamma_{1n} + |I_{2n}(\mathbf{x})|) \\
&\quad + P\Big(\Big|\frac{\widehat{m}_n^*(\mathbf{x})}{m^*(\mathbf{x})} - 1\Big| > \gamma_{1n}^*\Big) + P\Big(\Big|\frac{\widehat{m}_n^*(\mathbf{x})}{m^*(\mathbf{x})}I_{3n}(\mathbf{x})\Big| > \gamma_{1n}\Big).
\end{aligned}\tag{4.10}$$

From Lemma 5.1, it follows that

$$P\Big(\Big|\frac{\widehat{m}_n^*(\mathbf{x})}{m^*(\mathbf{x})} - 1\Big| > \gamma_{1n}^*\Big) \le \frac{E(\widehat{m}_n^*(\mathbf{x}) - m^*(\mathbf{x}))^2}{(m^*(\mathbf{x}))^2(\gamma_{1n}^*)^2} \le C\frac{\ln(n)}{nh_n^d(\gamma_{1n}^*)^2} = C\gamma_{1n}^*.\tag{4.11}$$

Lemma 5.1 and Theorem 2.2 ensure that

$$P\Big(\Big|\frac{\widehat{m}_n^*(\mathbf{x})}{m^*(\mathbf{x})}I_{3n}(\mathbf{x})\Big| > \gamma_{1n}\Big) \le E\Big(\frac{\widehat{m}_n^*(\mathbf{x})}{m^*(\mathbf{x})}I_{3n}(\mathbf{x})\Big)^2\gamma_{1n}^{-2} \le C\frac{\ln^2(n)}{(nh_n^d)^{\frac{1}{2}}\gamma_{1n}^2} = C\gamma_{1n}.\tag{4.12}$$

Let $\Lambda(\mathbf{u}) = E(\xi(Z,T,\delta,\xi_p(\mathbf{x}),\mathbf{x})\mid \mathbf{X}=\mathbf{u}, T\le Z)$. Then

$$\Lambda(\mathbf{u}) = \int_0^{\xi_p(\mathbf{x})}\frac{\mathrm{d}H_1^*(t|\mathbf{u})}{C(t|\mathbf{x})} - \int_0^{\xi_p(\mathbf{x})}\frac{C(t|\mathbf{u})\mathrm{d}H_1^*(t|\mathbf{x})}{C^2(t|\mathbf{x})}.$$

Obviously, $\Lambda(\mathbf{x})=0$ and the function Λ has bounded the first r_0 partial derivatives in $U(\mathbf{x})$ from (A3'). Hence we have

$$\begin{aligned}|I_{2n}(\mathbf{x})| &= \frac{1-p}{f(\xi_p(\mathbf{x})|\mathbf{x})\Delta(p|\mathbf{x})m^*(\mathbf{x})}\Big(\frac{n}{h_n^d}\Big)^{\frac{1}{2}}\Big|E\Big(K\Big(\frac{\mathbf{x}-\mathbf{X}}{h_n}\Big)\xi(Z,T,\delta,\xi_p(\mathbf{x}),\mathbf{x})\Big)\Big| \\ &= \frac{1-p}{f(\xi_p(\mathbf{x})|\mathbf{x})\Delta(p|\mathbf{x})m^*(\mathbf{x})}\Big(\frac{n}{h_n^d}\Big)^{\frac{1}{2}}\Big|\int_{\mathbb{R}^d}K\Big(\frac{\mathbf{x}-\mathbf{u}}{h_n}\Big)m^*(\mathbf{u})\Lambda(\mathbf{u})\mathrm{d}\mathbf{u}\Big| \\ &= \frac{1-p}{f(\xi_p(\mathbf{x})|\mathbf{x})\Delta(p|\mathbf{x})m^*(\mathbf{x})}(nh_n^d)^{\frac{1}{2}}\Big|\int_{\mathbb{R}^d}K(\mathbf{u})m^*(\mathbf{x}-h_n\mathbf{u})\Lambda(\mathbf{x}-h_n\mathbf{u})\mathrm{d}\mathbf{u}\Big| \\ &= O\big((nh_n^{d+2r_0})^{\frac{1}{2}}\big)=O(\gamma_{2n}). \end{aligned} \tag{4.13}$$

Note that $\gamma_{1n}^*\le\gamma_{1n}$. Then from (4.10)–(4.13), it suffices to verify that

$$\begin{aligned}\sup_u|P(I_{1n}(\mathbf{x})<u)-\Phi(u)| &= O(h_n+(qp^{-1})^{\frac{1}{3}}+(pn^{-1})^{\frac{1}{3}} \\ &\quad + h_n^{\frac{d(1-\rho)}{3}}+\gamma_{3n}^{\frac{1}{3}}+\gamma_{4n}+\gamma_{5n}^{\frac{1}{4}}).\end{aligned} \tag{4.14}$$

In fact, let $w=[\frac{n}{p+q}]$, and $\eta_i(\mathbf{x})=\frac{1-p}{f(\xi_p(\mathbf{x})|\mathbf{x})\Delta(p|\mathbf{x})m^*(\mathbf{x})h_n^{\frac{d}{2}}}\big[K(\frac{\mathbf{x}-\mathbf{X}_i}{h_n})\xi(Z_i,T_i,\delta_i,\xi_p(\mathbf{x}),\mathbf{x})-E\big(K(\frac{\mathbf{x}-\mathbf{X}_i}{h_n})\xi(Z_i,T_i,\delta_i,\xi_p(\mathbf{x}),\mathbf{x})\big)\big]$. Define $y_{mn}(\mathbf{x})$, $y'_{mn}(\mathbf{x})$, $y''_{wn}(\mathbf{x})$ as follows:

$$y_{mn}(\mathbf{x})=\sum_{i=k_m}^{k_m+p-1}\eta_i(\mathbf{x}),\quad y'_{mn}(\mathbf{x})=\sum_{j=l_m}^{l_m+q-1}\eta_j(\mathbf{x}),\quad y''_{wn}(\mathbf{x})=\sum_{k=w(p+q)+1}^{n}\eta_k(\mathbf{x}),$$

where $k_m=(m-1)(p+q)+1$, $l_m=(m-1)(p+q)+p+1$. Then

$$\begin{aligned}I_{1n}(\mathbf{x}) &= \frac{1}{\sqrt{n}}\sum_{i=1}^n\eta_i(\mathbf{x})=\frac{1}{\sqrt{n}}\Big\{\sum_{m=1}^w y_{mn}(\mathbf{x})+\sum_{m=1}^w y'_{mn}(\mathbf{x})+y''_{wn}(\mathbf{x})\Big\} \\ &:= n^{-\frac{1}{2}}\{S'_n(\mathbf{x})+S''_n(\mathbf{x})+S'''_n(\mathbf{x})\}.\end{aligned}$$

Let $\tau_{1n}=qp^{-1}+h_n^{d(1-\rho)}+\gamma_{3n}$, $\tau_{2n}=pn^{-1}+h_n^{d(1-\rho)}$. By applying Lemma 5.4, it follows that

$$\begin{aligned}\sup_u|P(I_{1n}(\mathbf{x})<u)-\Phi(u)| &= \sup_u|P(n^{-\frac{1}{2}}\{S'_n(\mathbf{x})+S''_n(\mathbf{x})+S'''_n(\mathbf{x})\}\le u)-\Phi(u)| \\ &\le \sup_u|P(n^{-\frac{1}{2}}S'_n(\mathbf{x})\le u)-\Phi(u)|+P(n^{-\frac{1}{2}}|S''_n(\mathbf{x})|>\tau_{1n}^{\frac{1}{3}}) \\ &\quad +P(n^{-\frac{1}{2}}|S'''_n(\mathbf{x})|>\tau_{2n}^{\frac{1}{3}})+(2\pi)^{-\frac{1}{2}}(\tau_{1n}^{\frac{1}{3}}+\tau_{2n}^{\frac{1}{3}}).\end{aligned}$$

Then, to verify (4.14), we only need to prove that

$$n^{-1}E(S''_n(\mathbf{x}))^2=O(\tau_{1n}),\quad n^{-1}E(S'''_n(y|\mathbf{x}))^2=O(\tau_{2n}) \tag{4.15}$$

and

$$\sup_u|P(n^{-\frac{1}{2}}S'_n(\mathbf{x})\le u)-\Phi(u)|=O(h_n+qp^{-1}+pn^{-1}+h_n^{d(1-\rho)}+\gamma_{4n}+\gamma_{5n}^{\frac{1}{4}}). \tag{4.16}$$

(i) We verify (4.15). Note that

$$\frac{1}{n}E(S''_n(\mathbf{x}))^2=\frac{1}{n}\sum_{m=1}^w\sum_{i=l_m}^{l_m+q-1}E\eta_i^2(\mathbf{x})+\frac{2}{n}\sum_{m=1}^w\sum_{l_m\le i<j\le l_m+q-1}\mathrm{Cov}(\eta_i(\mathbf{x}),\eta_j(\mathbf{x}))$$

$$+\frac{2}{n}\sum_{1\le i<j\le w}\mathrm{Cov}(y'_{in}(\mathbf{x}),y'_{jn}(\mathbf{x})). \tag{4.17}$$

From (A1') and (A3'), we get

$$\begin{aligned}E\eta_i^2(\mathbf{x}) &\le \frac{C}{h_n^d}E\Big(K^2\Big(\frac{\mathbf{x}-\mathbf{X}_i}{h_n}\Big)\xi^2(Z_i,T_i,\delta_i,\xi_p(\mathbf{x}),\mathbf{x})\Big)\le\frac{C}{h_n^d}\int_{\mathbb{R}^d}K^2\Big(\frac{\mathbf{x}-\mathbf{s}}{h_n}\Big)m^*(\mathbf{s})\mathrm{d}\mathbf{s}\\ &\le C\int_{\mathbb{R}^d}K^2(\mathbf{s})m^*(\mathbf{x}-h_n\mathbf{s})\mathrm{d}\mathbf{s}\le C. \end{aligned}\tag{4.18}$$

Using (A1') and (A4'), from the proof in (4.13) for $i<j$, we have

$$\begin{aligned}&|\mathrm{Cov}(\eta_i(\mathbf{x}),\eta_j(\mathbf{x}))|\le E|\eta_i(\mathbf{x})\eta_j(\mathbf{x})|+(E\eta_1(\mathbf{x}))^2\\ &\le Ch_n^{-d}E\Big(K\Big(\frac{X_i-x}{h_n}\Big)K\Big(\frac{X_j-x}{h_n}\Big)\Big)+O(h_n^{d+2r_0})\\ &=Ch_n^d\int_{\mathbb{R}^2}K(\mathbf{s}_1)K(\mathbf{s}_2)v_{j-i}^*(\mathbf{x}-h_n\mathbf{s}_1,\mathbf{x}-h_n\sigma_2)\mathrm{d}\mathbf{s}_1\mathrm{d}\mathbf{s}_2+O(h_n^{d+2r_0})=O(h_n^d).\end{aligned}$$

On the other hand, from Lemma 5.5 (taking $p=q=20\lambda$), it follows that

$$|\mathrm{Cov}(\eta_i(\mathbf{x}),\eta_j(\mathbf{x}))|\le C[\alpha(j-i)]^{1-\frac{1}{10\lambda}}(E|\eta_i(\mathbf{x})|^{20\lambda})^{\frac{1}{10\lambda}}$$

and $E|\eta_i(\mathbf{x})|^{20\lambda}\le Ch_n^{-10\lambda d}EK^{20\lambda}\big(\frac{\mathbf{x}-\mathbf{X}_i}{h_n}\big)=h_n^{-d(10\lambda-1)}\int_{\mathbb{R}^d}K^{20\lambda}(\mathbf{s})m^*(\mathbf{x}-h_n\mathbf{s})\mathrm{d}\mathbf{s}=O(h_n^{-d(10\lambda-1)})$, which yield $|\mathrm{Cov}(\eta_i(\mathbf{x}),\eta_j(\mathbf{x}))|\le C[\alpha(j-i)]^{1-\frac{1}{10\lambda}}h_n^{-d(1-\frac{1}{10\lambda})}$. Let $c_n=h_n^{-d\rho}$ for $\frac{20\lambda-1}{\lambda(10\lambda-1)}\le\rho<1$. Then

$$\begin{aligned}&\frac{1}{n}\sum_{1\le i<j\le n}|\mathrm{Cov}(\eta_i(\mathbf{x}),\eta_j(\mathbf{x}))|\\ &\le\frac{C}{n}\Big(\sum_{1\le j-i\le c_n}+\sum_{c_n+1\le j-i\le n-1}\Big)\min\big\{h_n^d,[\alpha(j-i)]^{1-\frac{1}{10\lambda}}h_n^{-d(1-\frac{1}{10\lambda})}\big\}\\ &\le C\big\{c_nh_n^d+h_n^{-d(1-\frac{1}{10\lambda})}c_n^{-(\lambda-\frac{11}{10})}\big\}=O(h_n^{d(1-\rho)}).\end{aligned}\tag{4.19}$$

Using Lemma 5.5 again, we have

$$\begin{aligned}&\frac{1}{n}\Big|\sum_{1\le i<j\le w}\mathrm{Cov}(y'_{in}(\mathbf{x}),y'_{jn}(\mathbf{x}))\Big|\le\frac{(1-p)^2}{f^2(\xi_p(\mathbf{x})|\mathbf{x})\Delta^2(p|\mathbf{x})(m^*(\mathbf{x}))^2nh_n^d}\sum_{1\le i<j\le w}\sum_{s=l_i}^{l_i+q-1}\sum_{t=l_j}^{l_j+q-1}\\ &\Big|\mathrm{Cov}\Big(K\Big(\frac{\mathbf{x}-\mathbf{X}_s}{h_n}\Big)\xi(Z_s,T_s,\delta_s,\xi_p(\mathbf{x}),\mathbf{x}),K\Big(\frac{\mathbf{x}-\mathbf{X}_t}{h_n}\Big)\xi(Z_t,T_t,\delta_t,\xi_p(\mathbf{x}),\mathbf{x})\Big)\Big|\\ &\le\frac{C}{nh_n^d}\sum_{i=1}^{w-1}\sum_{s=l_i}^{l_i+q-1}\Big\|K\Big(\frac{\mathbf{x}-\mathbf{X}_s}{h_n}\Big)\xi(Z_s,T_s,\delta_s,\xi_p(\mathbf{x}),\mathbf{x})\Big\|_{2+\delta}^2\sum_{j=i+1}^{w}\sum_{t=l_j}^{l_j+q-1}\alpha^{\frac{\delta}{2+\delta}}(t-s)\\ &\le\frac{Cwq}{nh_n^d}\Big\{h_n^d\int_{\mathbb{R}^d}K(\mathbf{s})^{2+\delta}m^*(\mathbf{x}-h_n\mathbf{s})\mathrm{d}\mathbf{s}\Big\}^{\frac{2}{2+\delta}}\sum_{i=p}^{\infty}\alpha^{\frac{\delta}{2+\delta}}(i)\\ &\le C\frac{wq}{n}h_n^{-\frac{d\delta}{2+\delta}}u(p)\le Cqp^{-1}h_n^{-\frac{d\delta}{2+\delta}}u(p)=O(\gamma_{3n}).\end{aligned}\tag{4.20}$$

From (4.17)–(4.20), we obtain $n^{-1}E(S''_n(\mathbf{x}))^2=O(qp^{-1}+h_n^{d(1-\rho)}+\gamma_{3n})=O(\tau_{1n})$ and

$$\frac{1}{n}E(S'''_n(\mathbf{x}))^2=\frac{1}{n}\sum_{i=w(p+q)+1}^{n}E\eta_i^2(\mathbf{x})+\frac{2}{n}\sum_{w(p+q)+1\le i<j\le n}\mathrm{Cov}(\eta_i(\mathbf{x}),\eta_j(\mathbf{x}))$$

$$\leq C\cdot\frac{n-w(p+q)}{n}+\frac{2}{n}\sum_{1\leq i<j\leq n}|\mathrm{Cov}(Z_i,Z_j)|$$
$$=O(pn^{-1}+h_n^{d(1-\rho)})=O(\tau_{2n}).$$

(ii) We prove (4.16). Let $\pi_{mn}(\mathbf{x}), m=1,2,\cdots,w$ be independent random variables, where the distribution of π_{mn} is the same as that of $y_{mn}(\mathbf{x})$ for $m=1,2,\cdots,w$. Put $U_n=n^{-\frac{1}{2}}\sum_{m=1}^{w}\pi_{mn}(\mathbf{x})$ and $s_n^2=n^{-1}\sum_{m=1}^{w}Ey_{mn}^2(\mathbf{x})$. Then

$$\begin{aligned}&\sup_u|P(n^{-\frac{1}{2}}S_n'(\mathbf{x})\leq u)-\Phi(u)|\\ \leq&\sup_u|P(n^{-\frac{1}{2}}S_n'(\mathbf{x})\leq u)-P(n^{-\frac{1}{2}}U_n\leq u)|\\ &+\sup_u\Big|P(n^{-\frac{1}{2}}U_n\leq u)-\Phi\Big(\frac{u}{s_n}\Big)\Big|+\sup_u\Big|\Phi\Big(\frac{u}{s_n}\Big)-\Phi(u)\Big|.\end{aligned}\tag{4.21}$$

Note that

$$\begin{aligned}&E(\xi^2(Z_i,T_i,\delta_i,y,\mathbf{x})\mid\mathbf{X}=\mathbf{s},T\leq Z)\\ =&E\Big(\frac{I(Z\leq y,\delta=1)}{C^2(Z|\mathbf{x})}\;\Big|\;\mathbf{X}=\mathbf{s},T\leq Z\Big)+E\Big[\Big(\int_0^y\frac{I(T\leq t\leq Z)\mathrm{d}H_1^*(t|\mathbf{x})}{C^2(t|\mathbf{x})}\Big)^2\Big|\mathbf{X}=\mathbf{s},T\leq Z\Big]\\ &-2E\Big[\Big(\frac{I(Z\leq y,\delta=1)}{C(Z|\mathbf{x})}\Big)\Big(\int_0^y\frac{I(T\leq t\leq Z)\mathrm{d}H_1^*(t|\mathbf{x})}{C^2(t|\mathbf{x})}\Big)\Big|\mathbf{X}=\mathbf{s},T\leq Z\Big]\\ =&E\Big(\frac{I(Z\leq y,\delta=1)}{C^2(Z|\mathbf{x})}\;\Big|\;\mathbf{X}=\mathbf{s},T\leq Z\Big)=\int_0^y\frac{\mathrm{d}H_1^*(t|\mathbf{s})}{C^2(t|\mathbf{x})}.\end{aligned}$$

Then, in view of $m^*(\mathbf{s})=\theta^{-1}\theta(\mathbf{s})m(\mathbf{s})$, $h_1^*(y|\mathbf{s})=\theta^{-1}(\mathbf{s})L(y|\mathbf{s})(1-G(y|\mathbf{s}))f(y|\mathbf{s})$, and $C(y|\mathbf{s})=\theta^{-1}(\mathbf{s})L(y|\mathbf{s})(1-G(y|\mathbf{s}))(1-F(y|\mathbf{s}))$, from (A1'), (A3') and (4.13), we have

$$\begin{aligned}E\eta_i^2(\mathbf{x})=&\frac{(1-p)^2}{f^2(\xi_p(\mathbf{x})|\mathbf{x})\Delta^2(p|\mathbf{x})(m^*(\mathbf{x}))^2h_n^d}\Big\{E\Big(K^2\Big(\frac{\mathbf{x}-\mathbf{X}_i}{h_n}\Big)\xi^2(Z_i,T_i,\delta_i,\xi_p(\mathbf{x}),\mathbf{x})\Big)\\ &-\Big[E\Big(K\Big(\frac{\mathbf{x}-\mathbf{X}_i}{h_n}\Big)\xi(Z_i,T_i,\delta_i,\xi_p(\mathbf{x}),\mathbf{x})\Big)\Big]^2\Big\}\\ =&\frac{(1-p)^2}{f^2(\xi_p(\mathbf{x})|\mathbf{x})\Delta^2(p|\mathbf{x})(m^*(\mathbf{x}))^2}\int_{\mathbb{R}^d}K^2(\mathbf{s})m^*(\mathbf{x}-h_n\mathbf{s})\Big(\int_0^{\xi_p(\mathbf{x})}\frac{h_1^*(t|\mathbf{x}-h_n\mathbf{s})\mathrm{d}t}{C^2(t|\mathbf{x})}\Big)\mathrm{d}\mathbf{s}\\ &+O(h_n^{d+2r_0})=1+O(h_n).\end{aligned}\tag{4.22}$$

Then, from (4.18)–(4.19) and (4.22), it follows that

$$\begin{aligned}s_n^2=&\frac{1}{n}\sum_{m=1}^{w}\sum_{i=k_m}^{k_m+p-1}E\eta_i^2(\mathbf{x})+\frac{2}{n}\sum_{m=1}^{w}\sum_{k_m\leq i<j\leq k_m+p-1}\mathrm{Cov}(\eta_i(\mathbf{x}),\eta_j(\mathbf{x}))\\ =&\frac{1}{n}\Big\{\sum_{i=1}^{n}E\eta_i^2(\mathbf{x})-\sum_{m=1}^{w}\sum_{i=l_m}^{l_m+q-1}E\eta_i^2(\mathbf{x})-\sum_{i=w(p+q)+1}^{n}E\eta_i^2(\mathbf{x})\\ &+2\sum_{m=1}^{w}\sum_{k_m\leq i<j\leq k_m+p-1}\mathrm{Cov}(\eta_i(\mathbf{x}),\eta_j(\mathbf{x}))\Big\}\\ =&1+O(h_n+p^{-1}q+n^{-1}p+h_n^{d(1-\rho)}),\end{aligned}$$

which implies that $s_n^2 \to 1$ and

$$\sup_u \Big|\Phi\Big(\frac{u}{s_n}\Big) - \Phi(u)\Big| = O(|s_n^2 - 1|) = O(h_n + p^{-1}q + n^{-1}p + h_n^{d(1-\rho)}). \tag{4.23}$$

By the Berry-Esseen inequality (see [25, p. 154, Theorem 5.7]), for $l > 2$, there exists some constant $C > 0$ such that

$$\sup_u \Big|P(n^{-\frac{1}{2}}U_n \le u) - \Phi\Big(\frac{u}{s_n}\Big)\Big| \le \frac{C}{n^{l/2}s_n^l}\sum_{m=1}^{w} E|\pi_{mn}(\mathbf{x})|^l. \tag{4.24}$$

Taking $l = 2(1+\beta)$ and $\mu = \delta - 2\beta$, we have $l + \mu = 2 + \delta$. Note that $\beta \le \frac{\delta\lambda-(2+\delta)}{2\lambda+(2+\delta)}$ implies that $\lambda \ge \frac{(1+\beta)(2+\delta)}{\delta-2\beta} = \frac{l(l+\mu)}{2\mu}$. Then, using Lemma 5.6 (take $p = l$ and $q = l + \mu$) and $E|\eta_1(\mathbf{x})|^{2+\delta} \le Ch_n^{-\frac{d\delta}{2}}$, we have

$$\begin{aligned}\sum_{m=1}^{w} E|\pi_{mn}(\mathbf{x})|^l &= \sum_{m=1}^{w} E|y_{mn}(\mathbf{x})|^l = \sum_{m=1}^{w} E\Big|\sum_{i=k_m}^{k_m+p-1}\eta_i(\mathbf{x})\Big|^{2(1+\beta)} \\ &\le Cwp^{1+\beta}(E|\eta_1(\mathbf{x})|^{2+\delta})^{\frac{2(1+\beta)}{2+\delta}} \\ &= O\big(wp^{1+\beta}h_n^{-\frac{d\delta(1+\beta)}{2+\delta}}\big) = O\big(np^{\beta}h_n^{-\frac{d\delta(1+\beta)}{2+\delta}}\big),\end{aligned}$$

which, together with (4.24), yields

$$\sup_u \Big|P(n^{-\frac{1}{2}}U_n \le u) - \Phi\Big(\frac{u}{s_n}\Big)\Big| = O\big(n^{-(1+\beta)}np^{\beta}h_n^{-\frac{d\delta(1+\beta)}{2+\delta}}\big) = O(\gamma_{4n}). \tag{4.25}$$

Let $\varphi(t)$ and $\psi(t)$ be the characteristic functions of $n^{-\frac{1}{2}}S_n'(\mathbf{x})$ and $n^{-\frac{1}{2}}U_n$, respectively. By the Esseen inequality (see [25, p. 146, Theorem 5.3]), for any $\Gamma > 0$,

$$\begin{aligned}&\sup_u |P(n^{-\frac{1}{2}}S_n'(\mathbf{x}) \le u) - P(n^{-\frac{1}{2}}U_n \le u)| \\ &\le \int_{-\Gamma}^{\Gamma}\Big|\frac{\varphi(t)-\psi(t)}{t}\Big|\mathrm{d}t + \Gamma\sup_u\int_{|v|\le\frac{C}{\Gamma}}|P(n^{-\frac{1}{2}}U_n \le u+v) - P(n^{-\frac{1}{2}}U_n \le u)|\mathrm{d}v \\ &:= H_{1n} + H_{2n}.\end{aligned} \tag{4.26}$$

Using Lemma 5.7, we have

$$\begin{aligned}|\varphi(t)-\psi(t)| &= \Big|E\exp\Big(\mathrm{i}t\sum_{m=1}^{w} n^{-\frac{1}{2}}y_{mn}(\mathbf{x})\Big) - \prod_{m=1}^{w} E\exp(\mathrm{i}tn^{-\frac{1}{2}}y_{mn}(\mathbf{x}))\Big| \\ &\le C|t|\alpha^{\frac{1}{2}}(q)\sum_{m=1}^{w}\|n^{-\frac{1}{2}}y_{mn}(\mathbf{x})\|_2 \le C|t|\alpha^{\frac{1}{2}}(q)n^{-\frac{1}{2}}\sum_{m=1}^{w}\Big\{E\Big|\sum_{i=k_m}^{k_m+p-1}\eta_i(\mathbf{x})\Big|^2\Big\}^{\frac{1}{2}}.\end{aligned}$$

From (4.13) and $|\mathrm{Cov}(\eta_i(\mathbf{x}),\eta_j(\mathbf{x}))| \le C\min\big\{h_n^d, [\alpha(j-i)]^{1-\frac{1}{10\lambda}}h_n^{-d(1-\frac{1}{10\lambda})}\big\}$, we have

$$E\Big|\sum_{i=k_m}^{k_m+p-1}\eta_i(\mathbf{x})\Big|^2 = \sum_{i=k_m}^{k_m+p-1}E\eta_i^2(\mathbf{x}) + 2\sum_{k_m\le i<j\le k_m+p-1}\mathrm{Cov}(\eta_i(\mathbf{x}),\eta_j(\mathbf{x})) = O(p).$$

Thus $H_{1n} = O(\Gamma(w^2n^{-1}p\alpha(q))^{\frac{1}{2}}) = O(\Gamma(np^{-1}\alpha(q))^{\frac{1}{2}}) = O(\Gamma\gamma_{5n}^{\frac{1}{2}})$. From (4.25), we have

$$\sup_u |P(n^{-\frac{1}{2}}U_n \le u+v) - P(n^{-\frac{1}{2}}U_n \le u)|$$

$$\leq \sup_u \Big| P(n^{-\frac{1}{2}} U_n \leq u+v) - \Phi\Big(\frac{u+v}{s_n}\Big) \Big| + \sup_u \Big| P(n^{-\frac{1}{2}} U_n \leq u) - \Phi\Big(\frac{u}{s_n}\Big) \Big|$$
$$+ \sup_u \Big| \Phi\Big(\frac{u+v}{s_n}\Big) - \Phi\Big(\frac{u}{s_n}\Big) \Big| = O(\gamma_{4n}) + O\Big(\frac{|v|}{s_n}\Big),$$

which yields that $H_{2n} = O\big(\gamma_{4n} + \frac{1}{\Gamma}\big)$. Choose $\Gamma = \gamma_{5n}^{-\frac{1}{4}}$. Then from (4.26), we have

$$\sup_u |P(n^{-\frac{1}{2}} S_n'(\mathbf{x}) \leq u) - P(n^{-\frac{1}{2}} U_n \leq u)| = O(\gamma_{4n} + \gamma_{5n}^{\frac{1}{4}}). \tag{4.27}$$

Therefore, from (4.21), (4.23), (4.25) and (4.27), we have

$$\sup_u |P(n^{-\frac{1}{2}} S_n'(\mathbf{x}) \leq u) - \Phi(u)| = O(h_n + p^{-1}q + n^{-1}p + h_n^{d(1-\rho)} + \gamma_{4n} + \gamma_{5n}^{\frac{1}{4}}).$$

5 Appendix

In this section, we list some preliminary lemmas which have been used in the proofs of the main results in Section 4. Let $\{\chi_i, i \geq 1\}$ be a stationary α-mixing sequence of real random variables with mixing coefficients $\{\alpha(k)\}$.

Lemma 5.1 (see [20]) *Let $\alpha(n) = O(n^{-\lambda})$ for some $\lambda > 2$, and let τ be a finite positive constant. Set $\Gamma_{1n} = \max\big\{\big(\frac{\ln(n)}{nh_n^d}\big)^{\frac{1}{2}}, h_n^{r_0}\big\}$.*

(a) *Suppose that* (A1)–(A4) *are satisfied. If* (A5)(i) *holds, then* $\sup_{\mathbf{x}\in I} \sup_{\tau_1 \leq y \leq \tau_2} |\widehat{F}_n(y|\mathbf{x}) - F(y|\mathbf{x})| = O(\Gamma_{1n})$ a.s. *If* (A5)(ii) *holds, then* $\sup_{\mathbf{x}\in I} \sup_{\tau_1 \leq y \leq \tau_2} |\widehat{F}_n(y|\mathbf{x}) - F(y|\mathbf{x})| = O_p(\Gamma_{1n})$.

(b) *Let $\mathbf{x} \in D$. Suppose that* (A1′)–(A4′) *are satisfied. If* (A5′)(i) *holds, then*

$$\sup_{\tau_1 \leq y \leq \tau_2} |\widehat{F}_n(y|\mathbf{x}) - F(y|\mathbf{x})| = O(\Gamma_{1n}) \ \text{ a.s.}, \quad \sup_{0 \leq y \leq \tau} |\widehat{H}_{1n}^*(y|\mathbf{x}) - H_1^*(y|\mathbf{x})| = O(\Gamma_{1n}) \ \text{ a.s.},$$
$$\sup_{0 \leq y \leq \tau} |\widehat{C}_n(y|\mathbf{x}) - C(y|\mathbf{x})| = O(\Gamma_{1n}) \ \text{ a.s.}, \quad \textit{and} \quad |\widehat{m}_n^*(\mathbf{x}) - m^*(\mathbf{x})| = O(\Gamma_{1n}) \ \text{ a.s.}$$

If (A5′)(ii) *holds, then* $\sup_{\tau_1 \leq y \leq \tau_2} |\widehat{F}_n(y|\mathbf{x}) - F(y|\mathbf{x})| = O_p(\Gamma_{1n})$, $\sup_{0 \leq y \leq \tau} |\widehat{H}_{1n}^*(y|\mathbf{x}) - H_1^*(y|\mathbf{x})| = O_p(\Gamma_{1n})$, *and* $\sup_{0 \leq y \leq \tau} |\widehat{C}_n(y|\mathbf{x}) - C(y|\mathbf{x})| = O_p(\Gamma_{1n})$.

Lemma 5.2 (see [20]) *Set $\xi(Z, T, \delta, y, \mathbf{x}) = \frac{I(Z \leq y, \delta = 1)}{C(Z|\mathbf{x})} - \int_0^y \frac{I(T \leq t \leq Z)}{C^2(t|\mathbf{x})} \mathrm{d}H_1^*(t|\mathbf{x})$. Let $\mathbf{x} \in D$ and $\alpha(n) = O(n^{-\lambda})$ for some $\lambda > 0$. Suppose that conditions* (A1′)–(A4′) *and* (B1)–(B2) *hold, and that $\frac{nh_n^{d+2r_0}}{\ln(n)} = O(1)$. Set $\Gamma_{2n} = \big(\frac{\ln(n)}{nh_n^d}\big)^{\frac{3}{4}}$.*

(a) *Let $\lambda > 6$ and $\tau_1 < a_{H(\cdot|\mathbf{x})}$. Then for $y \in [\tau_1, \tau_2]$, we have*

$$\widehat{F}_n(y|\mathbf{x}) - F(y|\mathbf{x}) = (1 - F(y|\mathbf{x})) \sum_{i=1}^n B_{ni}(\mathbf{x}) \xi(Z_i, T_i, \delta_i, y, \mathbf{x}) + Q_n(y|\mathbf{x}),$$

where $\sup_{y \in [\tau_1, \tau_2]} |Q_n(y|\mathbf{x})| = O(\Gamma_{2n})$ a.s. *when* (B3)(i) *holds;* $\sup_{y \in [\tau_1, \tau_2]} |Q_n(y|\mathbf{x})| = O_p(\Gamma_{2n})$ *when* (B3)(ii) *holds.*

(b) *Let $\lambda > 4$. If* (B3)(i) *holds, then*

$$\sup_{s,t \in [0,\infty): |s-t| \leq c\left(\frac{\ln(n)}{nh_n^d}\right)^{\frac{1}{2}}} |[\widehat{H}_{1n}^*(s|\mathbf{x}) - H_1^*(s|\mathbf{x})] - [\widehat{H}_{1n}^*(t|\mathbf{x}) - H_1^*(t|\mathbf{x})]| = O(\Gamma_{2n}) \ \text{ a.s.}$$

If (B3)(ii) *holds, then*

$$\sup_{s,t\in[0,\infty):|s-t|\le c\left(\frac{\ln(n)}{nh_n^d}\right)^{\frac{1}{2}}} |[\widehat{H}_{1n}^*(s|\mathbf{x}) - H_1^*(s|\mathbf{x})] - [\widehat{H}_{1n}^*(t|\mathbf{x}) - H_1^*(t|\mathbf{x})]| = O_p(\Gamma_{2n}).$$

Lemma 5.3 (see [20]) *Let* $\mathbf{x} \in D$ *and* $\alpha(n) = O(n^{-\lambda})$ *for some* $\lambda > 6$*. Suppose that conditions* (A1′)–(A4′), (B1)–(B2), (B3)(ii) *and* (E2)–(E3) *hold. If* $\tau_1 < a_{H(\cdot|\mathbf{x})}$*, then for* $y \in [\tau_1, \tau_2]$*, we have* $(nh_n^d)^{\frac{1}{2}}[\widehat{F}_n(y|\mathbf{x}) - F(y|\mathbf{x})] \xrightarrow{d} N\big(0, \sigma^2(y|\mathbf{x})\big)$*, where*

$$\sigma^2(y|\mathbf{x}) = \frac{\theta(1-F(y|\mathbf{x}))^2}{m(\mathbf{x})}\int_0^y \frac{f(t|\mathbf{x})\mathrm{d}t}{L(t|\mathbf{x})(1-G(t|\mathbf{x}))(1-F(t|\mathbf{x}))^2}\int_{\mathbb{R}^d} K^2(\mathbf{s})\mathrm{d}\mathbf{s}.$$

Lemma 5.4 *Let* X*,* V *and* $Y_1, \cdots, Y_m$ *be random variables, and then for positive numbers* a*,* w_1*,* $\cdots$*,* w_m*, we have* $\sup_u |P(X \le uV) - \Phi(u)| \le \sup_u |P(X \le u) - \Phi(u)| + P(|V-1| > a) + a$ *and*

$$\sup_u \Big|P\Big(X + \sum_{i=1}^m Y_i \le u\Big) - \Phi(u)\Big| \le \sup_u |P(X \le u) - \Phi(u)| + \sum_{i=1}^m \frac{w_i}{\sqrt{2\pi}} + \sum_{i=1}^m P(|Y_i| > w_i).$$

Proof The first inequality is a consequence of Michel and Pfanzagl [22] and the second one follows from Lemma 3.1 of Liang and Fan [17].

Lemma 5.5 (see [11, Corollary A.2, p. 278]) *Suppose that* X *and* Y *are random variables such that* $E|X|^p < \infty$*,* $E|Y|^q < \infty$*, where* p*,* $q > 1$*,* $p^{-1} + q^{-1} < 1$*. Then*

$$|EXY - EXEY| \le 8\|X\|_p\|Y\|_q\Big\{\sup_{\substack{A\in\sigma(X)\\ B\in\sigma(Y)}} |P(A\cap B) - P(A)P(B)|\Big\}^{1-p^{-1}-q^{-1}}.$$

Lemma 5.6 (see [27, Theorem 4.1]) *Let* $2 < p < q \le \infty$*. Assume that* $E\chi_n = 0$ *and* $\alpha(n) = O(n^{-\gamma})$ *for* $\gamma > 0$*. Then there exists* $Q = Q(p, q, \gamma) < \infty$ *such that* $E\Big|\sum\limits_{i=1}^n \chi_i\Big|^p \le Qn^{\frac{p}{2}} \max\limits_{1\le i\le n} \|\chi_i\|_q^p$ *if* $\gamma \ge \frac{pq}{2(q-p)}$.

Lemma 5.7 (see [30]) *Let* p *and* q *be positive integers. Set* $\eta_r = \sum\limits_{j=(r-1)(p+q)+1}^{(r-1)(p+q)+p} \chi_j$ *for* $1 \le r \le w$*. If* $s > 0$*,* $r > 0$ *with* $\frac{1}{s} + \frac{1}{r} = 1$*, then there exists a constant* $C > 0$ *such that* $\Big|E\exp\big(\mathrm{i}t\sum\limits_{r=1}^w \eta_r\big) - \prod\limits_{r=1}^w E\exp(\mathrm{i}t\eta_r)\Big| \le C|t|\alpha^{\frac{1}{s}}(q)\sum\limits_{r=1}^w \|\eta_r\|_r$.

References

[1] Akritas, M. G. and LaValley, M. P., A generalized product-limit estimator for truncated data, *J. Nonparametric Statist.*, **17**, 2005, 643–663.

[2] Beran, R., Nonparametric regression with randomly censored survival data, Technical Report, Department of Statistics, University of California, Berkeley, 1981.

[3] Cai, T., Wei, L. J. and Wilcox, M., Semiparametric regression analysis for clustered survival data, *Biometrika*, **87**, 2000, 867–878.

[4] Cai, Z. W., Regression quantiles for time series, *Econometric Theory*, **18**, 2002, 169–192.

[5] Cai, J. and Kim, J., Nonparametric quantile estimation with correlated failure time data, *Lifetime Data Analysis*, **9**, 2003, 357–371.

[6] Dabrowska, D., Nonparametric quantile regression with censored data, *Sankhyā, Ser. A*, **54**, 1992, 252–259.

[7] Doukhan, P., Mixing: Properties and Examples, Lecture Notes in Statistics, Vol. 85, Springer-Verlag, Berlin, 1994.

[8] Ferraty, F., Rabhi, A. and Vieu, P., Conditional quantiles for dependent functional data with application to the climatic *El Niño* phenomenon, *Sankhyā*, **67**, 2005, 378–398.

[9] González-Manteiga, W. and Cadarso-Suárez, Z., Asymptotic properties of a generalized Kaplan-Meier estimator with some applications, *J. Nonparametric Statist.*, **4**, 1994, 65–78.

[10] Gürler, Ü., Stute, W. and Wang, J. L., Weak and strong quantile representations for randomly truncated data with applications, *Statist. Probab. Lett.*, **17**, 1993, 139–148.

[11] Hall, P. and Heyde, C. C., Martingale Limit Theory and Its Application, Academic Press, New York, 1980.

[12] Honda, T., Nonparametric estimation of a conditional quantile for α-mixing processes, *Ann. Inst. Statist. Math.*, **52**, 2000, 459–470.

[13] Iglesias-Pérez, C. and González-Manteiga, W., Strong representation of a generalized product-limit estimator for truncated and censored data with some applications, *J. Nonparametric Statist.*, **10**, 1999, 213–244.

[14] Iglesias-Pérez, C., Strong representation of a conditional quantile function estimator with truncated and censored data, *Statist. Probab. Lett.*, **65(2)**, 2003, 79–91.

[15] Kang, S. S. and Koehler, K. J., Modification of the Greenwood formula for correlated failure times, *Biometrics*, **53**, 1997, 885–899.

[16] Lemdani, M., Ould-Saïd, E. and Poulin, N., Asymptotic properties of a conditional quantile estimator with randomly truncated data, *J. Multivar. Analysis*, **100**, 2009, 546–559.

[17] Liang, H. Y. and Fan, G. L., Berry-Esseen-type bounds of estimators in a semiparametric model with linear process errors, *J. Multivar. Analysis*, **100**, 2009, 1–15.

[18] Liang, H. Y. and de Uña-Álvarez, J., Asymptotic properties of conditional quantile estimator for censored dependent observations, *Ann. Inst. Statist. Math.*, **63**, 2011, 267–289.

[19] Liang, H. Y. and de Uña-Álvarez, J., Conditional quantile estimation with auxiliary information for left-truncated and dependent data, *J. Statist. Plan. Inference*, **141**, 2011, 3475–3488.

[20] Liang, H. Y., de Uña-Álvarez, J. and Iglesias-Pérez, C., Asymptotic properties of conditional distribution estimator with truncated, censored and dependent data, *Test*, **21**(4), 2012, 790–810.

[21] Lo, S. and Singh, K., The product-limit estimator and the bootstrap: Some asymptotic representations, *Probab. Theory Related Fields*, **71**, 1985, 455–465.

[22] Michel, R. and Pfanzagl, J., The accuracy of the normal approximation for minimum constrast estimates, *Z. Wahrsch. Verw. Gebiete*, **18**, 1971, 73–84.

[23] Ould-Saïd, E., A strong uniform convergence rate of kernel conditional quantile estimator under random censorship, *Statist. Probab. Lett.*, **76**, 2006, 579–586.

[24] Ould-Saïd, E., Yahia, D. and Necir, A., A strong uniform convergence rate of a kernel conditional quantile estimator under random left-truncation and dependent data, *Electronic J. Statist.*, **3**, 2009, 426–445.

[25] Petrov, V. V., Limit Theorems of Probability Theory, Oxford Univ. Press Inc., New York, 1995.

[26] Polonik, W. and Yao, Q., Set-indexed conditional empirical and quantile processes based on dependent data, *J. Multivar. Analysis*, **80**, 2002, 234–255.

[27] Shao, Q. and Yu, H., Weak convergence for weighted empirical processes of dependent sequences, *Ann. Probab.*, **24**, 1996, 2098–2127.

[28] Van Keilegom, I. and Veraverbeke, N., Bootstrapping quantiles in a fixed design regression model with censored data, *J. Statist. Plan. Inference*, **69**, 1998, 115–131.

[29] Withers, C. S., Conditions for linear processes to be strong mixing, *Z. Wahrsch. Verw. Gebiete*, **57**, 1981, 477–480.

[30] Yang, S. C. and Li, Y. M., Uniformly asymptotic normality of the regression weighted estimator for strong mixing samples, *Acta Math. Sinica*, **49**(5), 2006, 1163–1170.

[31] Zhou, X., Sun, L. Q. and Ren, H., Quantile estimation for left truncted and right censored data, *Statist. Sinica*, **10**, 2000, 1217–1229.