Available online at www.sciencedirect.com
ScienceDirect
C. R. Acad. Sci. Paris, Ser. I 344 (2007) 205–210
COMPTES RENDUS MATHEMATIQUE
http://france.elsevier.com/direct/CRASS1/

Statistics

Mean square convergence for estimators of additive regression under random censorship

Mohammed Debbarh, Vivian Viallon

L.S.T.A. université de Paris 6, 175, rue du Chevaleret, 75013 Paris, France

Received 19 June 2006; accepted after revision 1 December 2006

Available online 10 January 2007

Presented by Paul Deheuvels

Abstract

In this Note, we establish the mean square convergence rate for estimators of an additive regression function under random censorship. To build our estimator, the marginal integration method is coupled with some *Inverse Probability of Censoring Weighted* [$I.P.C.W.$] estimates of the multivariate regression function. ***To cite this article: M. Debbarh, V. Viallon, C. R. Acad. Sci. Paris, Ser. I 344 (2007).***

Résumé

Convergence en moyenne quadratique de l'estimateur de la fonction de régression additive en données censurées. Dans cette Note, nous proposons d'établir la vitesse de convergence en moyenne quadratique de l'estimateur d'une fonction de régression additive en données censurées. Pour construire nos estimateurs, nous combinons la méthode d'intégration marginale à des estimateurs de la fonction de régression multivariée de type *Inverse Probability of Censoring Weighted* [$I.P.C.W.$]. ***Pour citer cet article : M. Debbarh, V. Viallon, C. R. Acad. Sci. Paris, Ser. I 344 (2007).***

Version française abrégée

Soit $(Y_1, C_1, \mathbf{X}_1), \ldots, (Y_n, C_n, \mathbf{X}_n)$, $n \geqslant 1$, une suite de répliques indépendantes et identiquement distribuées du triplet $(Y, C, \mathbf{X})$ à valeurs dans $\mathbb{R} \times \mathbb{R} \times \mathbb{R}^d$, $d \geqslant 2$, où Y est la variable d'intérêt, C une variable de censure et $\mathbf{X} = (X_1, \ldots, X_d)$ une variable de conditionnement. Soit ψ une fonction donnée, mesurable à valeurs dans $\mathbb{R}$, telle que

$$\text{(A)} \quad \psi(y) = 0 \quad \text{si } y \in [\tau, +\infty[, \quad \text{avec } \tau < T_F := \sup\{t\colon\ P(Y > t) > 0\}.$$

Soit m_ψ la fonction de régression de $\psi(Y)$ sachant $\mathbf{X}$. Nous supposons ici que m_ψ est additive i.e.,

E-mail addresses: debbarh@ccr.jussieu.fr (M. Debbarh), viallon@ccr.jussieu.fr (V. Viallon).

1631-073X/$ – see front matter
doi:10.1016/j.crma.2006.12.002

$$m_\psi(\mathbf{x}) = E\big(\psi(Y) \mid \mathbf{X}=\mathbf{x}\big), \quad \forall \mathbf{x} = (x_1, \ldots, x_d) \in \mathbb{R}^d,$$
$$= \mu + \sum_{\ell=1}^{d} m_\ell(x_\ell).$$

Soit $\widehat{m}^\star_{\psi,\mathrm{add}}$ l'estimateur de la fonction de régression additive m_ψ défini en (9) en combinant la méthode d'intégration marginale, les idées de Jones et al. [8], et l'estimateur de la régression multivariée censurée proposé par Carbonez et al. [4] et Kohler et al. [11]. Sous les hypothèses (C.1)–(C.4), (Q.1), (Q.2), (K.1), (K.2), (F.1), (F.2) et (H.1), (H.2) présentées ci-dessous, nous obtenons la vitesse de convergence en moyenne quadratique de l'estimateur de $\widehat{m}^\star_{\psi,\mathrm{add}}$ vers m_ψ, pour tout $\mathbf{x}$ appartenant à un certain compact,

$$E\big(\widehat{m}^\star_{\psi,\mathrm{add}}(\mathbf{x}) - m_\psi(\mathbf{x})\big)^2 = \mathcal{O}\big(n^{-2k/(2k+1)}\big).$$

1. Introduction

In many statistical application domains, the variable of interest is only partially observed, because of right censoring. In medical studies, for instance, considering the survival time of a patient, this variable is *censored* for individuals who are still alive at the end of the study or individuals who dropped-out before the termination of the study. Moreover, in this setting, the variable of interest is often related to numerous covariates, making the use of nonparametric estimates unsuitable because of the well-known *curse of dimensionality* [14]. Working under the additive model assumption enables us to get round this issue. Namely, combining the marginal integration method (see Newey [13], Linton and Nielsen [12]) with an initial $I.P.C.W.$ estimator of the multivariate censored regression function (see Carbonez et al. [4] and Kohler et al. [11]), we derive a new estimate of this quantity, for which, under the additive model assumption, the rate of mean square convergence is shown to be independent of the dimension of the covariates (see Theorem 3.1). A similar result can be found in Camlong et al. [3] for uncensored data.

Let $(Y, C, \mathbf{X}), (Y_1, C_1, \mathbf{X}_1), (Y_2, C_2, \mathbf{X}_2), \ldots$ be independent and identically distributed $\mathbb{R} \times \mathbb{R} \times \mathbb{R}^d$-valued random variables, $d \geqslant 2$. Here Y is the variable of interest, C a censoring variable and $\mathbf{X} = (X_1, \ldots, X_d)$ a vector of concomitant variables. Set for all $t \in \mathbb{R}$ $F(t) = P(Y > t)$ and $G(t) = P(C > t)$ the right-continuous survival functions pertaining to Y and C respectively. Denote by ψ a given real measurable function fulfilling the following assumption.

(A) $\psi(y) = 0$ for all $y \in [\tau, +\infty[$, with $\tau < T_F := \sup\{t\colon F(t) > 0\}$.

Consider the additive regression function of $\psi(Y)$ evaluated at $\mathbf{X} = \mathbf{x}$, defined by,

$$m_\psi(\mathbf{x}) = E\big(\psi(Y) \mid \mathbf{X}=\mathbf{x}\big), \quad \forall\, \mathbf{x} = (x_1, \ldots, x_d) \in \mathbb{R}^d, \tag{1}$$
$$= \mu + \sum_{\ell=1}^{d} m_\ell(x_\ell). \tag{2}$$

In this Note, we work under the identifiability assumption $Em_\ell(X_\ell) = 0$, $\ell = 1, \ldots, d$, which ensures that $\mu = E\psi(Y)$.

Remark 1. (i) Let $\mathbb{I}_E$ be the indicator function of a given set E. With the particular choice $\psi(\cdot) = \psi_{t_0}(\cdot) = \mathbb{I}_{(-\infty;t_0]}(\cdot)$, for any given $0 \leqslant t_0 \leqslant \tau$, our work allows to treat the case of the censored conditional survival function (see for instance Beran [2], Dabrowska [5], Deheuvels and Derzko [6]). Moreover, other choices of the function ψ may lead to some properties for the conditional density function. These properties would enable to derive some results for the conditional hazard rate and the conditional quantiles.

(ii) The choice $\psi(y) = y$ violates our condition (A). Thus, our work does not allow to treat the 'classical' regression function. However, under some additional – and quite strong – assumptions (especially the existence of a constant T such that $P(Y > T) = 0$, $P(Y = T) > 0$ and $P(C > T) > 0$), some modifications in our proofs would allow to treat this special and interesting case (see for instance Kohler et al. [10]).

In the right censorship model, the variables Y_i and C_i are not directly observed, so that only $\mathbf{X}_i$, $Z_i = \min\{Y_i, C_i\}$ and $\delta_i = \mathbb{I}_{\{Y_i \leqslant C_i\}}$ are at our disposal. Therefore, to estimate m_ψ, we work with the observed sample $(\mathbf{X}_i, Z_i, \delta_i)_{1\leqslant i\leqslant n}$.

2. Additive regression estimate under random censorship

Let K_1, K_2, K_3 and K be some kernels respectively defined in $\mathbb{R}$, $\mathbb{R}^{d-1}$, $\mathbb{R}^d$ and $\mathbb{R}^d$. Introduce the kernel estimator $\hat{f}_n$ of the marginal density function f of $\mathbf{X}$,

$$\hat{f}_n(\mathbf{x}) = \frac{1}{nh_n^d}\sum_{j=1}^n K\left(\frac{\mathbf{x}-\mathbf{X}_j}{h_n}\right),$$

where $(h_n)_{n\geqslant 1}$ is a given sequence of positive numbers. In the sequel, we denote by G_n the Kaplan–Meier [9] estimator of G. Namely, for all $y \in \mathbb{R}$,

$$G_n(y) = 1 - \prod_{1\leqslant i\leqslant n}\left(\frac{N_n(Z_i)-1}{N_n(Z_i)}\right)^{\beta_i}, \quad \text{with } \beta_i = \mathbb{I}_{\{Z_i\leqslant y\}}(1-\delta_i) \text{ and } N_n(y) = \sum_{j=1}^n \mathbb{I}_{\{Z_j\geqslant y\}}. \tag{3}$$

To estimate the regression function defined in (1), the two following $I.P.C.W.$ estimators will be used (see Carbonez et al. [4], Kohler et al. [11] and Jones et al. [8]):

$$\widetilde{m}^{\star}_{\psi,n}(\mathbf{x}) = \sum_{i=1}^n W_{n,i}(\mathbf{x})\frac{\delta_i\psi(Z_i)}{G_n(Z_i)} \quad \text{with } W_{n,i}(\mathbf{x}) = \frac{K_3(\frac{\mathbf{x}-\mathbf{X}_i}{h_{1,n}})}{nh_{1,n}^d\hat{f}_n(\mathbf{X}_i)}, \tag{4}$$

and

$$\widetilde{m}^{\star}_{\psi,n,\ell}(\mathbf{x}) = \sum_{i=1}^n W^{\ell}_{n,i}(\mathbf{x})\frac{\delta_i\psi(Z_i)}{G_n(Z_i)} \quad \text{with } W^{\ell}_{n,i}(\mathbf{x}) = \frac{K_1\big(\frac{x_\ell - X_{i,\ell}}{h_{1,n}}\big)K_2\big(\frac{\mathbf{x}_{-\ell}-\mathbf{X}_{i,-\ell}}{h_{2,n}}\big)}{nh_{1,n}h_{2,n}^{d-1}\hat{f}_n(\mathbf{X}_i)}, \ \ell = 1,\ldots,d, \tag{5}$$

where $(h_{j,n})_{n\geqslant 1}$, $j = 1, 2$, are two sequences of positive numbers. For all $\mathbf{x} = (x_1,\ldots,x_d) \in \mathbb{R}^d$ and every $\ell = 1,\ldots,d$, set $\mathbf{x}_{-\ell} = (x_1,\ldots, x_{\ell-1}, x_{\ell+1},\ldots,x_d)$. To estimate the additive components, we use the marginal integration method. To do so, let $q_1,\ldots,q_d$ be d given integration density functions defined in $\mathbb{R}$. Then, setting $q(\mathbf{x}) = \prod_{\ell=1}^d q_\ell(x_\ell)$ and $q_{-\ell}(\mathbf{x}_{-\ell}) = \prod_{j\neq\ell} q_j(x_j)$, we define

$$\eta_\ell(x_\ell) = \int_{\mathbb{R}^{d-1}} m_\psi(\mathbf{x})q_{-\ell}(\mathbf{x}_{-\ell})\,\mathrm{d}\mathbf{x}_{-\ell} - \int_{\mathbb{R}^d} m_\psi(\mathbf{x})q(\mathbf{x})\,\mathrm{d}\mathbf{x}, \quad \ell = 1,\ldots,d, \tag{6}$$

in such a way that the two following equalities hold:

$$\eta_\ell(x_\ell) = m_\ell(x_\ell) - \int_{\mathbb{R}} m_\ell(z)q_\ell(z)\,\mathrm{d}z, \quad \ell = 1,\ldots,d, \quad \text{and} \quad m_\psi(\mathbf{x}) = \sum_{\ell=1}^d \eta_\ell(x_\ell) + \int_{\mathbb{R}^d} m_\psi(\mathbf{z})q(\mathbf{z})\,\mathrm{d}\mathbf{z}. \tag{7}$$

In view of (5) and (6), a natural estimator of the ℓ-th component η_ℓ is given by

$$\hat{\eta}^{\star}_\ell(x_\ell) = \int_{\mathbb{R}^{d-1}} \widetilde{m}^{\star}_{\psi,n,\ell}(\mathbf{x})q_{-\ell}(\mathbf{x}_{-\ell})\,\mathrm{d}\mathbf{x}_{-\ell} - \int_{\mathbb{R}^d} \widetilde{m}^{\star}_{\psi,n,\ell}(\mathbf{x})q(\mathbf{x})\,\mathrm{d}\mathbf{x}, \quad \ell = 1,\ldots,d, \tag{8}$$

from which we deduce an estimate $\widehat{m}^{\star}_{\psi,\mathrm{add}}$ of the additive regression function,

$$\widehat{m}^{\star}_{\psi,\mathrm{add}} = \sum_{\ell=1}^d \hat{\eta}^{\star}_\ell(x_\ell) + \int_{\mathbb{R}^d} \widetilde{m}^{\star}_{\psi,n}(\mathbf{x})q(\mathbf{x})\,\mathrm{d}\mathbf{x}. \tag{9}$$

3. Hypotheses and results

Before stating our results, we introduce some assumptions and additional notations. Consider the following hypotheses pertaining to $(Y, C, \mathbf{X})$. Let $\mathcal{C}_1,\ldots,\mathcal{C}_d$, be d compact intervals on $\mathbb{R}$ and set $\mathcal{C} = \mathcal{C}_1 \times\cdots\times \mathcal{C}_d$. For every subset $\mathcal{E}$ of $\mathbb{R}^q$, $q \geqslant 1$, and any $\alpha > 0$, introduce the α-neighborhood $\mathcal{E}^\alpha$ of $\mathcal{E}$, defined by $\mathcal{E}^\alpha = \{x\colon \inf_{y\in\mathcal{E}} \|x-y\|_{\mathbb{R}^q} \leqslant \alpha\}$, with $\|\cdot\|_{\mathbb{R}^q}$ standing for the Euclidian norm on $\mathbb{R}^q$.

(C.1): C and $(\mathbf{X}, Y)$ are independent.
(C.2): G is continuous and $G(\tau) > 0$.
(C.3): There exists a positive constant M such that $\sup_{y \leqslant \tau} |\psi(y)| \leqslant M < \infty$.
(C.4): m_ψ is a k-times continuously differentiable function, $k \geqslant 1$, and

$$\sup_{\mathbf{x}} \left| \frac{\partial^k m_\psi}{\partial x_\ell^k}(\mathbf{x}) \right| < \infty; \quad \ell = 1, \ldots, d.$$

Denote by f_ℓ, $\ell = 1, \ldots, d$, the density functions of X_ℓ, $\ell = 1, \ldots, d$. The functions f and f_ℓ, $\ell = 1, \ldots, d$, will be supposed to be continuous, and we will assume the existence of a constant $\alpha > 0$ such that the following conditions hold.

(F.1): $\forall x_\ell \in \mathcal{C}_\ell^\alpha$, $f_\ell(x_\ell) > 0$, $\ell = 1, \ldots, d$, and $\forall \mathbf{x} \in \mathcal{C}^\alpha$ with $f(\mathbf{x}) > 0$.
(F.2): f is k'-times continuously differentiable on $\mathcal{C}$, $k' > kd$.

The kernels K_1, K_2, K_3 and K defined in $\mathbb{R}$, $\mathbb{R}^{d-1}$, $\mathbb{R}^d$ and $\mathbb{R}^d$ respectively, are assumed to be continuous, with compact support and integrating to 1. Moreover, they are assumed to fulfill (K.1), (K.2) introduced below.

(K.1): The kernels K_1, K_2 and K_3 are of order k.
(K.2): The kernel K is of order k'.

The known integration density functions $q_{-\ell}$ and q_ℓ, $\ell = 1, \ldots, d$, satisfy the following assumptions:

(Q.1): $q_{-\ell}$ is bounded and continuous, $\ell = 1, \ldots, d$.
(Q.2): q_ℓ has k continuous and bounded derivatives, with compact support included in $\mathcal{C}_\ell$, $\ell = 1, \ldots, d$.

Finally, we will work under the following conditions on the smoothing parameters h_n and $h_{j,n}$, $j = 1, 2$.

(H.1): $h_n = c'(\frac{\log n}{n})^{1/(2k'+d)}$, for a fixed $0 < c' < \infty$.
(H.2): $h_{1,n} = c_1 n^{-1/(2k+1)}$ and $h_{2,n} = c_2 n^{-1/(2k+1)}$, for fixed $0 < c_1, c_2 < \infty$.

Theorem 3.1. *Under the conditions* (C.1-2-3-4), (F.1-2), (K.1-2), (Q.1-2) *and* (H.1-2)*, we have, for all* $\mathbf{x} \in \mathcal{C}$,

$$E\big(\widehat{m}_{\psi,\mathrm{add}}^\star(\mathbf{x}) - m_\psi(\mathbf{x})\big)^2 = \mathcal{O}\big(n^{-2k/(2k+1)}\big).$$

Proof. Denote by $\widehat{\widehat{m}}_{\psi,\mathrm{add}}$, $\hat{\hat{\eta}}_\ell$, $\widetilde{\widetilde{m}}_{\psi,n}$ and $\widetilde{\widetilde{m}}_{\psi,n,\ell}$, $\ell = 1, \ldots, d$, [resp. $\widehat{m}_{\psi,\mathrm{add}}$, $\hat{\eta}_\ell$, $\widetilde{m}_{\psi,n}$ and $\widetilde{m}_{\psi,n,\ell}$, $\ell = 1, \ldots, d$] the versions of $\widehat{m}_{\psi,\mathrm{add}}^\star$, $\hat{\eta}_\ell^\star$, $\widetilde{m}_{\psi,n}^\star$ and $\widetilde{m}_{\psi,n,\ell}^\star$, $\ell = 1, \ldots, d$, when f and G are known (G_n and $\hat{f}_n$ are formally replaced by f and G in (4) and (5)) [resp. G is known and f is unknown]. Using the classical inequality $(a+b)^2 \leqslant 2(a^2+b^2)$, if follows that, for all $\mathbf{x} \in \mathcal{C}$,

$$E\big(\widehat{m}_{\psi,\mathrm{add}}^\star(\mathbf{x}) - m_\psi(\mathbf{x})\big)^2 \leqslant 4E\big(\widehat{\widehat{m}}_{\psi,\mathrm{add}}(\mathbf{x}) - m_\psi(\mathbf{x})\big)^2 + 4E\big(\widehat{m}_{\psi,\mathrm{add}}(\mathbf{x}) - \widehat{\widehat{m}}_{\psi,\mathrm{add}}(\mathbf{x})\big)^2 + 2E\big(\widehat{m}_{\psi,\mathrm{add}}^\star(\mathbf{x}) - \widehat{m}_{\psi,\mathrm{add}}(\mathbf{x})\big)^2 =: I_1(\mathbf{x}) + I_2(\mathbf{x}) + I_3(\mathbf{x}). \tag{10}$$

First consider the term $I_1(\mathbf{x})$. Set

$$\widehat{C}_n = \int_{\mathbb{R}^d} \widetilde{\widetilde{m}}_{\psi,n}(\mathbf{x}) q(\mathbf{x})\,\mathrm{d}\mathbf{x}, \qquad C_n = \mu + \int_{\mathbb{R}^{d-1}} \sum_{j=2}^d m_j(z_j) \mathcal{G}(\mathbf{z}_{-1})\,\mathrm{d}\mathbf{z}_{-1},$$

$$C = \int_{\mathbb{R}} m_1(x_1) q_1(x_1)\,\mathrm{d}x_1 \quad \text{and} \quad \mathcal{G}(\mathbf{u}_{-1}) = \int_{\mathbb{R}^{d-1}} \frac{1}{h_{2,n}^{d-1}} K_2\left(\frac{\mathbf{x}_{-1} - \mathbf{u}_{-1}}{h_{2,n}}\right) q_{-1}(\mathbf{x}_{-1})\,\mathrm{d}\mathbf{x}_{-1}.$$

By using the inequality $(\sum_{i=1}^d a_i)^2 \leqslant d\sum_{i=1}^d a_i^2$, we obtain, for all $\mathbf{x}\in\mathcal{C}$,

$$I_1(\mathbf{x}) \leqslant 8d\sum_{\ell=1}^{d} E\big(\hat{\hat{\eta}}_\ell(x_\ell)-\eta_\ell(x_\ell)\big)^2 + 8E\big(\widehat{C}_n - C_n - C\big)^2 + \mathcal{O}\big(n^{-2k/(2k+1)}\big). \tag{11}$$

Under (F.1-2), (K.1), (Q.1-2) and (H.2), following the same lines as in Camlong et al. [3], it can be shown that, for all $\mathbf{x}\in\mathcal{C}$,

$$I_1(\mathbf{x}) = \mathcal{O}\big(n^{-2k/(2k+1)}\big) \quad \text{as } n\to\infty. \tag{12}$$

Next, consider the function I_2. The Jensen's inequality yields, for all $\mathbf{x}\in\mathcal{C}$,

$$\begin{aligned} I_2(\mathbf{x}) \leqslant{}& 16d\sum_{\ell=1}^{d}\int_{\mathbb{R}^{d-1}} E\big(\widetilde{m}_{\psi,n,\ell}(\mathbf{x}) - \widetilde{\widetilde{m}}_{\psi,n,\ell}(\mathbf{x})\big)^2 q_{-\ell}^2(\mathbf{x}_{-\ell})\,\mathrm{d}\mathbf{x}_{-\ell} \\ &+ 16d\sum_{\ell=1}^{d}\int_{\mathbb{R}^{d}} E\big(\widetilde{m}_{\psi,n,\ell}(\mathbf{x}) - \widetilde{\widetilde{m}}_{\psi,n,\ell}(\mathbf{x})\big)^2 q^2(\mathbf{x})\,\mathrm{d}\mathbf{x} + 8\int_{\mathbb{R}^{d}} E\big(\widetilde{\widetilde{m}}_{\psi,n}(\mathbf{x}) - \widetilde{m}_{\psi,n}(\mathbf{x})\big)^2 q^2(\mathbf{x})\,\mathrm{d}\mathbf{x}. \end{aligned} \tag{13}$$

Using the decomposition $1/f = 1/\hat{f}_n + (\hat{f}_n - f)/(\hat{f}_n f)$, it is easily derived that, for all $\ell = 1,\dots,d$, there exists a constant $0 < M_1 < \infty$, such that, for all $\mathbf{x}\in\mathcal{C}$ and for n large enough, we have, under (Q.2),

$$\begin{aligned} &E\big(\widetilde{m}_{\psi,n,\ell}(\mathbf{x}) - \widetilde{\widetilde{m}}_{\psi,n,\ell}(\mathbf{x})\big)^2 \\ &\quad\leqslant M_1 E\left(\frac{1}{h_{1,n}h_{2,n}^{d-1}}\left|K_1\left(\frac{x_\ell - X_{i,\ell}}{h_{1,n}}\right)K_2\left(\frac{\mathbf{x}_{-\ell} - \mathbf{X}_{i,-\ell}}{h_{2,n}}\right)\right| \times \sup_{\mathbf{x}\in\mathcal{C}^\alpha}\big|\hat{f}_n(\mathbf{x}) - f(\mathbf{x})\big|\right)^2. \end{aligned} \tag{14}$$

Under our assumptions, we can use, for example, the result of Ango-Nze and Rios [1], which ensures that,

$$\sup_{\mathbf{x}\in\mathcal{C}^\alpha}\big|\hat{f}_n(\mathbf{x}) - f(\mathbf{x})\big| = \mathcal{O}\left(\left(\frac{\log n}{n}\right)^{k'/(2k'+d)}\right) \quad \text{a.s.}$$

By combining this result with (13) and (14), it is readily shown that, for all $\mathbf{x}\in\mathcal{C}$,

$$I_2(\mathbf{x}) = \mathcal{O}\left(\left(\frac{\log n}{n}\right)^{2k'/(2k'+d)}\right) = \mathcal{O}\big(n^{-2k/(2k+1)}\big) \quad \text{as } n\to\infty. \tag{15}$$

Finally, to evaluate $I_3(\mathbf{x})$, first observe that, for all $\mathbf{x}\in\mathcal{C}$,

$$\begin{aligned} I_3(\mathbf{x}) \leqslant{}& 8d\sum_{\ell=1}^{d}\int_{\mathbb{R}^{d-1}} E\big(\widetilde{m}^\star_{\psi,n,\ell}(\mathbf{x}) - \widetilde{\widetilde{m}}_{\psi,n,\ell}(\mathbf{x})\big)^2 q_{-\ell}^2(\mathbf{x}_{-\ell})\,\mathrm{d}\mathbf{x}_{-\ell} \\ &+ 8d\sum_{\ell=1}^{d}\int_{\mathbb{R}^{d}} E\big(\widetilde{m}^\star_{\psi,n,\ell}(\mathbf{x}) - \widetilde{\widetilde{m}}_{\psi,n,\ell}(\mathbf{x})\big)^2 q^2(\mathbf{x})\,\mathrm{d}\mathbf{x} + 4\int_{\mathbb{R}^{d}} E\big(\widetilde{\widetilde{m}}^\star_{\psi,n}(\mathbf{x}) - \widetilde{m}_{\psi,n}(\mathbf{x})\big)^2 q^2(\mathbf{x})\,\mathrm{d}\mathbf{x}. \end{aligned} \tag{16}$$

But, under (A), we have

$$\big|\widetilde{m}^\star_{\psi,n}(\mathbf{x}) - \widetilde{m}_{\psi,n}(\mathbf{x})\big| \leqslant \sup_{y\leqslant\tau}\big|\psi(y)\big|\sup_{y\leqslant\tau}\big|G_n(y) - G(y)\big|\sup_{y\leqslant\tau}\frac{1}{G_n(y)G(y)}\sum_{i=1}^{n}\big|W_{n,i}(\mathbf{x})\big|. \tag{17}$$

Obviously, a similar upper bound can be obtained for $|\widetilde{m}^\star_{\psi,n,\ell}(\mathbf{x}) - \widetilde{m}_{\psi,n,\ell}(\mathbf{x})|$, $\ell = 1,\dots,d$. Moreover, under the assumptions (A) and (C.2), we have $G(\tau) > 0$ and $F(\tau) > 0$. Thus, we can apply the iterated law of the logarithm of Földes and Rejtő [7], which ensures that

$$\sup_{y\leqslant\tau}\big|G_n(y) - G(y)\big| = \mathcal{O}\left(\frac{\sqrt{\log\log n}}{n}\right) \quad \text{a.s.}$$

Besides, under the conditions imposed on K and f, the term $\sum_{i=1}^{n} |W_{n,i}(\mathbf{x})|$ is almost surely uniformly bounded on $\mathcal{C}$, for n large enough. Combining these two last results with (16) and (17), it follows, under (C.2) and (C.3) that, for all $\mathbf{x} \in \mathcal{C}$,

$$I_3(\mathbf{x}) = \mathcal{O}\left(\frac{\log\log n}{n}\right) \quad \text{as } n \to \infty. \tag{18}$$

By combining (10), (12), (15) and (18), we achieve the proof of Theorem 3.1. □

Remark 2. The result of Földes and Rejtő [7] does not apply if (A) is not satisfied, and the conclusion of Theorem 3.1 is generally false when this assumption does not hold, unless some additional hypotheses are imposed (see Remark 1(ii)).

References

[1] P. Ango-Nze, R. Rios, Density estimation in L^∞ norm for mixing processes, J. Statist. Plann. Inference 83 (1) (2000) 75–90.
[2] R. Beran, Nonparametric regression with randomly censored data, Technical report, Univ. California Press, Berkeley, 1981.
[3] C. Camlong-Viot, P. Sarda, P. Vieu, Additive time series: the kernel integration method, Math. Methods Statist. 9 (4) (2000) 358–375.
[4] A. Carbonez, L. Györfi, E.C. van der Meulen, Partitioning-estimates of a regression function under random censoring, Statist. Decisions 13 (1) (1995) 21–37.
[5] D.M. Dabrowska, Nonparametric regression with censored covariates, J. Multivariate Anal. 54 (2) (1995) 253–283.
[6] P. Deheuvels, G. Derzko, Nonparametric estimation of conditional lifetime distributions under random censorship, in: Advances in Statistical Methods for the Health Sciences: Applications to Cancer and AIDS Studies, Genome Sequence Analysis and Survival Analysis, Springer, New York, 2006, in press.
[7] A. Földes, L. Rejtő, A L.I.L. type result for the product-limit estimator, Z. Wahrsch. Verw. Gebiete 56 (1981) 75–86.
[8] M.C. Jones, S.J. Davies, B.U. Park, Versions of kernel-type regression estimators, J. Amer. Statist. Assoc. 89 (1994) 825–832.
[9] E.L. Kaplan, P. Meier, Non parametric estimation for incomplete observations, J. Amer. Statist. Assoc. 53 (1958) 457–481.
[10] M. Kohler, S. Kul, K. Máthé, Least squares estimates for censored regression, Preprint, Available at http://www.mathematik.uni-stuttgart.de/mathA/lst3/kohler/hfm-pub-en.html, 2006.
[11] M. Kohler, K. Máthé, M. Pintér, Prediction from randomly right censored data, J. Multivariate Anal. 80 (1) (2002) 73–100.
[12] O.B. Linton, J.P. Nielsen, A kernel method of estimating structured nonparametric regression based on marginal integration, Biometrika 82 (1995) 93–100.
[13] W.K. Newey, Kernel estimation of partial means and a general variance estimator, Econometric Theory 10 (2) (1994) 233–253.
[14] C.J. Stone, Additive regression and other nonparametric models, Ann. Statist. 13 (2) (1985) 689–705.