Ann. Inst. Statist. Math.
Vol. 49, No. 3, 467–491 (1997)

ESTIMATION AND BOOTSTRAP WITH CENSORED DATA IN FIXED DESIGN NONPARAMETRIC REGRESSION

INGRID VAN KEILEGOM AND NOËL VERAVERBEKE

Department of Mathematics, Limburgs Universitair Centrum, Universitaire Campus, B-3590 Diepenbeek, Belgium

(Received March 1, 1996; revised June 18, 1996)

Abstract. We study Beran's extension of the Kaplan-Meier estimator for the situation of right censored observations at fixed covariate values. This estimator for the conditional distribution function at a given value of the covariate involves smoothing with Gasser Müller weights. We establish an almost sure asymptotic representation which provides a key tool for obtaining central limit results. To avoid complicated estimation of asymptotic bias and variance parameters, we propose a resampling method which takes the covariate information into account. An asymptotic representation for the bootstrapped estimator is proved and the strong consistency of the bootstrap approximation to the conditional distribution function is obtained.

Key words and phrases: Asymptotic normality, asymptotic representation, bootstrap approximation, fixed design, kernel estimator, nonparametric regression, right censoring.

1. Introduction

At fixed design points $0 \leq x_1 \leq \cdots \leq x_n \leq 1$ we have nonnegative responses $Y_1, \ldots, Y_n$ such as survival times or failure times. These responses are independent random variables and the distribution function of the response Y_i at x_i will be denoted by $F_{x_i}(t) = P(Y_i \leq t)$.

As often occurs in clinical trials or industrial life tests, the responses $Y_1, \ldots, Y_n$ are subject to random right censoring, i.e. the observed random variables at design point x_i are in fact T_i and δ_i $(i = 1, \ldots, n)$, with

$$T_i = \min(Y_i, C_i) \quad \text{and} \quad \delta_i = I(Y_i \leq C_i)$$

where $C_1, \ldots, C_n$ are nonnegative independent censoring variables with distribution functions $G_{x_i}(t) = P(C_i \leq t)$. We will assume independence of the Y_i and the C_i for each i. Consequently we have that the distribution function $H_{x_i}(t) = P(T_i \leq t)$ satisfies the relation

$$1 - H_{x_i}(t) = (1 - F_{x_i}(t))(1 - G_{x_i}(t)). \tag{1.1}$$

At a given fixed design value $x \in [0,1]$, we write F_x, G_x, H_x for the distribution function of respectively the response Y_x at x, the censoring variable C_x at x and $T_x = \min(Y_x, C_x)$. Also we will write $\delta_x = I(Y_x \le C_x)$. (Note that for the design variables x_i we write Y_i, C_i, T_i, δ_i instead of Y_{x_i}, C_{x_i}, T_{x_i}, δ_{x_i}.)

Here we consider a general dependence of F_x on x. As a particular example we could consider a general heteroscedastic regression model, where $F_x(t) = F((t - \mu(x))/\sigma(x))$ for some distribution function F with mean 0 and variance 1 and for some unknown smooth functions μ and σ on $[0,1]$.

This paper concerns nonparametric estimation of $F_x(t)$ and is organized as follows. Below, in Section 2, we define the distribution function estimator F_{xh} for F_x. It is Beran's generalization of the usual Kaplan-Meier estimator, taking regression into account and depending on a bandwidth sequence $\{h_n\}$. In Theorem 2.1 we establish a basic almost sure representation for F_{xh}. This representation then leads in Section 3 to the basic asymptotic properties of the estimator like asymptotic normality and weak convergence. In Section 4 we propose a bootstrap version F^*_{xhg} of the estimator F_{xh}. This estimator depends on the bandwidth sequence $\{h_n\}$ and on a preliminary bandwidth sequence $\{g_n\}$. The latter is used to generate resampled data (x_i, T_i^*, δ_i^*) from the original data (x_i, T_i, δ_i). The main theorem in this Section 4 then gives an almost sure asymptotic representation for the bootstrap estimator F^*_{xhg}. In Section 5 we show the validity of the proposed bootstrap procedure in the sense that the bootstrap distribution of $(nh_n)^{1/2}(F^*_{xhg}(t) - F_{xg}(t))$ is strongly consistent for the distribution function of $(nh_n)^{1/2}(F_{xh}(t) - F_x(t))$. The Appendix (Section 6) contains a series of basic Lemmas A.1–A.5 on empirical distribution functions of the kernel type which are used very frequently in the paper.

2. The Kaplan-Meier type estimator and its almost sure asymptotic representation

In the case of no censoring, a natural nonparametric estimator for $F_x(t)$ is the kernel estimator due to Stone (1977) with Gasser-Müller type weights. It is given by

$$\sum_{i=1}^{n} w_{ni}(x; h_n) I(Y_i \le t)$$

where

$$w_{ni}(x; h_n) = \frac{1}{c_n(x; h_n)} \int_{x_{i-1}}^{x_i} \frac{1}{h_n} K\left(\frac{x-z}{h_n}\right) dz \qquad i = 1, \ldots, n$$

$$c_n(x; h_n) = \int_0^{x_n} \frac{1}{h_n} K\left(\frac{x-z}{h_n}\right) dz.$$

Here $x_0 = 0$, K is a known probability density function (kernel) and $\{h_n\}$ is a sequence of positive constants (bandwidth), tending to 0 as $n \to \infty$.

In the present case of censoring, Beran (1981) was the first who studied regression problems in a fully nonparametric way. His estimator is a generalization of the product-limit estimator of Kaplan and Meier (1958) and some of its asymptotic properties have been studied by Dabrowska (1987, 1989, 1992) (for the random design case). In absence of ties, the estimator is given by (we refer to Section 4 for a definition allowing for ties)

$$F_{xh}(t) = 1 - \left\{ \prod_{T_{(i)} \leq t} \left(1 - \frac{w_{n(i)}(x; h_n)}{1 - \sum_{j=1}^{i-1} w_{n(j)}(x; h_n)} \right)^{\delta_{(i)}} \right\} I(t < T_{(n)}). \tag{2.1}$$

Here $T_{(1)} < \cdots < T_{(n)}$ are the ordered $T_1, \ldots, T_n$ and the $\delta_{(i)}$ and $w_{n(i)}(x; h_n)$ are the corresponding δ_i and $w_{ni}(x; h_n)$. Clearly, $F_{xh}(t)$ is a step function with jumps only at the uncensored observations. Note that if we think the weights $w_{ni}(x; h_n)$ all equal to n^{-1}, then $F_{xh}(t)$ becomes the classical Kaplan-Meier estimator. On the other hand, in the case of no censoring ($T_i = Y_i$ and $\delta_i = 1$ for all i) the estimator equals the kernel estimator of Stone (1977), which has been studied in Aerts *et al.* (1994*a*) in the fixed design regression problem with complete observations.

We will use the notation $H_x^u(t) = P(T_x \leq t, \delta_x = 1) = \int_0^t (1 - G_x(s-))dF_x(s)$ for the subdistribution function of the uncensored observations and the cumulative hazard function Λ_x is defined by

$$\Lambda_x(t) = \int_0^t \frac{dF_x(s)}{1 - F_x(s-)} = \int_0^t \frac{dH_x^u(s)}{1 - H_x(s-)}. \tag{2.2}$$

We now replace H_x and H_x^u by the following kernel type estimators

$$H_{xh}(t) = \sum_{i=1}^n w_{ni}(x; h_n) I(T_i \leq t) \tag{2.3}$$

$$H_{xh}^u(t) = \sum_{i=1}^n w_{ni}(x; h_n) I(T_i \leq t, \delta_i = 1) \tag{2.4}$$

which leads to the following Nelson-Aalen type estimator for $\Lambda_x(t)$:

$$\Lambda_{xh}(t) = \int_0^t \frac{dH_{xh}^u(s)}{1 - H_{xh}(s-)}.$$

Some notation to be used is the following. For the design points $x_1, \ldots, x_n$ we denote $\underline{\Delta}_n = \min_{1 \leq i \leq n}(x_i - x_{i-1})$ and $\overline{\Delta}_n = \max_{1 \leq i \leq n}(x_i - x_{i-1})$. For the kernel K we will use $\|K\|_\infty = \sup_{u \in \mathbb{R}} K(u)$, $\|K\|_2^2 = \int_{-\infty}^{\infty} K^2(u)du$, $\mu_1^K = \int_{-\infty}^{\infty} uK(u)du$, $\mu_2^K = \int_{-\infty}^{\infty} u^2 K(u)du$. We will constantly use the following assumptions on the design and on the kernel:

(C1) $x_n \to 1$, $\overline{\Delta}_n = O(n^{-1})$, $\overline{\Delta}_n - \underline{\Delta}_n = o(n^{-1})$

(C2) K is a probability density function with finite support $[-L, L]$ for some $L > 0$, $\mu_1^K = 0$, and K is Lipschitz of order 1.

Note that $c_n(x; h_n) = 1$ for n sufficiently large (depending on x) since $x_n \to 1$ and K has finite support. This makes that in all proofs of asymptotic results, we will take $c_n(x; h_n) = 1$.

Concerning the support of distribution functions we will use the following notation: if L is any (sub)distribution function, then T_L denotes the right endpoint of its support, i.e. $T_L = \inf\{t : L(t) = L(\infty)\}$. Clearly, $T_{H_x} = \min(T_{F_x}, T_{G_x})$.

In the formulation of our results, we will need typical types of smoothness conditions on functions like $H_x(t)$ and $H_x^u(t)$. We formulate them here for a general (sub)distribution function $L_x(t)$, $0 \le x \le 1$, $t \in \mathbb{R}$, and for a fixed $T > 0$.

(C3) $\dot{L}_x(t) = \frac{\partial}{\partial x} L_x(t)$ exists and is continuous in $(x,t) \in [0,1] \times [0,T]$

(C4) $L'_x(t) = \frac{\partial}{\partial t} L_x(t)$ exists and is continuous in $(x,t) \in [0,1] \times [0,T]$

(C5) $\ddot{L}_x(t) = \frac{\partial^2}{\partial x^2} L_x(t)$ exists and is continuous in $(x,t) \in [0,1] \times [0,T]$

(C6) $L''_x(t) = \frac{\partial^2}{\partial t^2} L_x(t)$ exists and is continuous in $(x,t) \in [0,1] \times [0,T]$

(C7) $\dot{L}'_x(t) = \frac{\partial^2}{\partial x \partial t} L_x(t)$ exists and is continuous in $(x,t) \in [0,1] \times [0,T]$.

Note that (C5) and (C7) imply (C3) and that (C6) and (C7) imply (C4). Also, (C3) implies that $L_x(t)$ is Lipschitz in the following sense: there is a constant C_L such that for $0 \le x,\ x' \le 1$:

$$\sup_{0<t<T} |L_x(t) - L_{x'}(t)| \le C_L |x - x'|.$$

Similarly, (C4) implies: there is a constant $\tilde{C}_L$ such that for $0 \le t,\ t' \le T$:

$$\sup_{0\le x\le 1} |L_x(t) - L_x(t')| \le \tilde{C}_L |t - t'|.$$

Also note that imposing conditions (C3) and (C4) on $H_x(t)$ and $H_x^u(t)$ implies that $F_x(t)$ and $G_x(t)$ are continuous in $(x,t) \in [0,1] \times [0,T]$.

Finally, we note that, since most of the results in this paper are at a fixed design point x, we could relax the above conditions (C3)–(C7) by requiring the continuity only in $U_x \times [0,T]$ instead of $[0,1] \times [0,T]$, with U_x some open neighborhood of the fixed point x.

The proposed estimator F_{xh} in (2.1) has a complicated structure which does not allow immediate study of its properties. As Lo and Singh (1986) did for the ordinary Kaplan-Meier estimator, we now prove an a.s. asymptotic representation. It represents $F_{xh}(t)$ as a weighted sum plus a remainder term, which under certain conditions, is of the a.s. order $O((nh_n)^{-3/4}(\log n)^{3/4})$ as $n \to \infty$. It should be noted here that, for (a slightly different version) of F_{xh}, Gonzalez-Manteiga and Cadarso-Suarez (1994) recently obtained an a.s. asymptotic representation with remainder term $O((nh_n)^{-3/4}(\log n)^{3/4} + h_n^2)$, for bandwidth sequences h_n satisfying $\frac{\log n}{nh_n} \to 0$ and $nh_n^3 \to \infty$ (where the last condition, however, could be weakened to $nh_n^2 \to \infty$). For situations where the remainder term should be $o((nh_n)^{-1/2})$, the extra conditions $\frac{\log n}{(nh_n)^3} \to 0$ and $nh_n^5 \to 0$ are therefore required. This last condition, however, is not satisfied for the optimal bandwidth sequence $h_n = Cn^{-1/5}$ (which minimizes the approximate mean squared error). Since our bootstrap investigation in Sections 4 and 5 will be carried out for this optimal bandwidth sequence, we have to reconsider the main steps in the proof and use for instance our Lemma A.5(b) in the estimation of their term II(3) (Gonzalez-Manteiga and Cadarso-Suarez (1994), p. 75).

THEOREM 2.1. *Assume* (C1), (C2), $H_x(t)$ *and* $H_x^u(t)$ *satisfy* (C5), (C6) *and* (C7) *in* $[0,T]$ *with* $T < T_{H_x}$, $h_n \to 0$, $\frac{\log n}{nh_n} \to 0$, $\frac{nh_n^5}{\log n} = O(1)$. *Then, for* $t < T_{H_x}$:

$$F_{xh}(t) - F_x(t) = \sum_{i=1}^{n} w_{ni}(x;h_n) g_{tx}(T_i,\delta_i) + r_n(t,x)$$

where

$$\begin{aligned} g_{tx}(T_i,\delta_i) = (1-F_x(t))\bigg\{ \int_0^t \frac{I(T_i \le s) - H_x(s)}{(1-H_x(s))^2} dH_x^u(s) \\ + \frac{I(T_i \le t, \delta_i = 1) - H_x^u(t)}{1-H_x(t)} \\ - \int_0^t \frac{I(T_i \le s, \delta_i = 1) - H_x^u(s)}{(1-H_x(s))^2} dH_x(s) \bigg\} \end{aligned}$$

and where

$$\sup_{0\le t\le T} |r_n(t,x)| = O((nh_n)^{-3/4}(\log n)^{3/4}) \qquad a.s.$$

as $n \to \infty$.

PROOF. Because of continuity, $1 - F_x(t) = \exp(-\Lambda_x(t))$. Introducing for $t < T_{H_{xh}}$

$$\tilde{\Lambda}_{xh}(t) = \int_0^t \frac{dH_{xh}^u(s)}{1-H_{xh}(s)}$$

we have the following identity

$$F_{xh}(t) - F_x(t) = [e^{-\Lambda_x(t)} - e^{-\tilde{\Lambda}_{xh}(t)}] - [1 - F_{xh}(t) - e^{-\tilde{\Lambda}_{xh}(t)}].$$

By one term Taylor expansion of the second and two term Taylor expansion of the first part, we obtain

$$F_{xh}(t) - F_x(t) = (1-F_x(t))(\tilde{\Lambda}_{xh}(t) - \Lambda_x(t)) + R_{n1}(t) + R_{n2}(t) \tag{2.5}$$

where

$$\begin{aligned} R_{n1}(t) &= -\frac{1}{2}\exp(-\Lambda_{xh}^{\circ}(t))[\tilde{\Lambda}_{xh}(t) - \Lambda_x(t)]^2 \\ R_{n2}(t) &= \exp(-\Lambda_{xh}^{\circ\circ}(t))[-\log(1-F_{xh}(t)) - \tilde{\Lambda}_{xh}(t)] \end{aligned}$$

with $\Lambda_{xh}^{\circ}(t)$ between $\tilde{\Lambda}_{xh}(t)$ and $\Lambda_x(t)$ and $\Lambda_{xh}^{\circ\circ}(t)$ is between $-\log(1-F_{xh}(t))$ and $\tilde{\Lambda}_{xh}(t)$.

Furthermore, for $t < T_{H_{xh}}$, (and omitting the integration variables in the notation):

$$\begin{aligned} \tilde{\Lambda}_{xh}(t) - \Lambda_x(t) &= \int_0^t \frac{dH_{xh}^u}{1-H_{xh}} - \int_0^t \frac{dH_x^u}{1-H_x} \\ &= \int_0^t \left[\frac{1}{1-H_{xh}} - \frac{1}{1-H_x}\right] dH_x^u + \int_0^t \frac{1}{1-H_x} d(H_{xh}^u - H_x^u) \\ &\quad + \int_0^t \left[\frac{1}{1-H_{xh}} - \frac{1}{1-H_x}\right] d(H_{xh}^u - H_x^u). \end{aligned}$$

Writing for the integrand in the first term

$$\frac{H_{xh}-H_x}{(1-H_{xh})(1-H_x)}=\frac{H_{xh}-H_x}{(1-H_x)^2}+\frac{(H_{xh}-H_x)^2}{(1-H_x)^2(1-H_{xh})}$$

and integrating by parts in the second term, we arrive at

$$\begin{aligned}(2.6)\qquad \tilde{\Lambda}_{xh}(t)-\Lambda_x(t)=&\int_0^t\frac{H_{xh}-H_x}{(1-H_x)^2}dH_x^u+\frac{H_{xh}^u(t)-H_x^u(t)}{1-H_x(t)}\\&-\int_0^t\frac{H_{xh}^u-H_x^u}{(1-H_x)^2}dH_x+R_{n3}(t)+R_{n4}(t)\end{aligned}$$

where

$$\begin{aligned}R_{n3}(t)&=\int_0^t\frac{(H_{xh}-H_x)^2}{(1-H_x)^2(1-H_{xh})}dH_x^u\\R_{n4}(t)&=\int_0^t\left[\frac{1}{1-H_{xh}}-\frac{1}{1-H_x}\right]d[H_{xh}^u-H_x^u].\end{aligned}$$

Because $H_x(T)<1$ and $H_{xh}(T)\to H_x(T)$ a.s. (by Lemma A.2), we may suppose that $T<T_{H_{xh}}$. For $R_{n3}(t)$ we have

$$\begin{aligned}\sup_{0\le t\le T}|R_{n3}(t)|&\le\left(\sup_{0\le t\le T}|H_{xh}(t)-H_x(t)|\right)^2\frac{1}{(1-H_{xh}(T))(1-H_x(T))^2}\\&=O((nh_n)^{-1}\log n)\qquad\text{a.s.}\end{aligned}$$

by application of Lemma A.2 and Lemma A.4(b). By Lemma 2.1 below

$$\sup_{0\le t\le T}|R_{n4}(t)|=O((nh_n)^{-3/4}(\log n)^{3/4})\qquad\text{a.s.}$$

Also from (2.6), Lemma A.4 and the bounds for $R_{n3}(t)$ and $R_{n4}(t)$:

$$(2.7)\qquad \sup_{0\le t\le T}|\tilde{\Lambda}_{xh}(t)-\Lambda_x(t)|=O((nh_n)^{-1/2}(\log n)^{1/2})\qquad\text{a.s.}$$

This gives that

$$\sup_{0\le t\le T}|R_{n1}(t)|=O((nh_n)^{-1}\log n)\qquad\text{a.s.}$$

In Lemma 2.2 below, we will show

$$\sup_{0\le t\le T}|R_{n2}(t)|=O((nh_n)^{-1})\qquad\text{a.s.}$$

This, together with (2.5) and (2.6), shows that the theorem is proved.

We now prove the two lemmas used above.

LEMMA 2.1. *Under the conditions of Theorem 2.1, as $n \to \infty$,*

$$\sup_{0\le t\le T}\left|\int_0^t \left(\frac{1}{1-H_{xh}}-\frac{1}{1-H_x}\right) d(H^u_{xh}-H^u_x)\right| = O((nh_n)^{-3/4}(\log n)^{3/4}) \qquad a.s. \tag{2.8}$$

PROOF. Partitioning the interval $[0,T]$ into $k_n = O((nh_n)^{1/2}(\log n)^{-1/2})$ subintervals $[t_i,t_{i+1}]$ of length $O((nh_n)^{-1/2}(\log n)^{1/2})$, we have, as in the proof of Lemma 2 of Lo and Singh (1986), that the left hand side in (2.8) is bounded above by

$$\begin{aligned}&2\max_{1\le i\le k_n}\sup_{t_i\le y\le t_{i+1}}\left|\frac{1}{1-H_{xh}(y)}-\frac{1}{1-H_{xh}(t_i)}-\frac{1}{1-H_x(y)}+\frac{1}{1-H_x(t_i)}\right|\\ &\quad + k_n\frac{\sup_{0\le t\le T}|H_{xh}(t)-H_x(t)|}{(1-H_{xh}(T))(1-H_x(T))}\\ &\quad \cdot\max_{1\le i\le k_n}|H^u_{xh}(t_{i+1})-H^u_{xh}(t_i)-H^u_x(t_{i+1})+H^u_x(t_i)|.\end{aligned} \tag{2.9}$$

To estimate the first term in (2.9) we further subdivide each $[t_i,t_{i+1}]$ into $a_n = O((nh_n)^{1/4}(\log n)^{-1/4})$ subintervals $[t_{ij},t_{i,j+1}]$ of length $O((nh_n)^{-3/4}(\log n)^{3/4})$. Using that $\sup_{0\le t\le T}|H_{xh}(t)-H_x(t)| = O((nh_n)^{-1/2}(\log n)^{1/2})$ a.s. (which follows by application of Lemma A.4(b) to H_{xh}), we obtain, as in Lo and Singh (1986) that the first term in (2.9) is a.s. bounded by $C\max_{1\le i\le k_n}\max_{1\le j\le a_n}|H_{xh}(t_{ij})-H_{xh}(t_i)-H_x(t_{ij})+H_x(t_i)|+O((nh_n)^{-3/4}(\log n)^{3/4})$, for some constant $C>0$. Applying Lemma A.5 and Corollary A.1 to the functions H_{xh} and H_x gives that this term is $O((nh_n)^{-3/4}(\log n)^{3/4})$ a.s. The second term in (2.9) is treated similarly and leads to the same order.

LEMMA 2.2. *Assume* (C1), (C2), $H_x(t)$ *satisfies* (C3) *in* $[0,T]$ *with* $T < T_{H_x}$, $h_n \to 0$, $\frac{\log n}{nh_n}\to 0$. *Then, as* $n\to\infty$,

$$\sup_{0<t<T}|-\log(1-F_{xh}(t))-\tilde\Lambda_{xh}(t)| = O((nh_n)^{-1}) \qquad a.s.$$

PROOF. Because $H_x(T)<1$ and $H_{xh}(T)\to H_x(T)$ a.s., we may suppose that $T<T_{H_{xh}}$. If $t\le T$, then

$$\begin{aligned}\tilde\Lambda_{xh}(t) &= \int_0^t \frac{1-G_{xh}(s-)}{1-G_{xh}(s)}\frac{dF_{xh}(s)}{1-F_{xh}(s)}\\ &= \int_0^t \frac{G_{xh}(s)-G_{xh}(s-)}{1-G_{xh}(s)}\frac{dF_{xh}(s)}{1-F_{xh}(s)} - \log(1-F_{xh}(t)).\end{aligned}$$

Since $\sup_{0\le t\le T}|G_{xh}(t)-G_{xh}(t-)| = O((nh_n)^{-1})$ a.s., the result follows.

3. Central limit results

In this section we consider some major consequences of the a.s. representation in Theorem 2.1 concerning the limiting distribution of $F_{xh}(t)$. It should be noted that Theorem 2.1 also leads to other properties of the estimator. For instance, under the conditions of Theorem 2.1, we have the following rate of uniform consistency result: as $n \to \infty$,

$$\sup_{0\le t\le T} |F_{xh}(t) - F_x(t)| = O((nh_n)^{-1/2}(\log n)^{1/2}) \quad \text{a.s.} \tag{3.1}$$

This follows from (2.5) and (2.7). We will not discuss this type of properties of the estimator here. They can be found in Van Keilegom and Veraverbeke (1996) and include inequalities for $P(\sup_{0\le t\le T} |F_{xh}(t) - F_x(t)| > \varepsilon)$ and for $P(\sup_{0\le s,t\le T;|t-s|\le a_n} |F_{xh}(t) - F_{xh}(s) - F_x(t) + F_x(s)| > \varepsilon)$ for some $\varepsilon > 0$ and $a_n \to 0$.

The first major consequence of Theorem 2.1 is of course the asymptotic normality result for $(nh_n)^{1/2}(F_{xh}(t)-F_x(t))$. Looking at the main term in the asymptotic representation of Theorem 2.1 gives

$$\begin{aligned}
(3.2)\quad & \sum_{i=1}^n w_{ni}(x;h_n)g_{tx}(T_i,\delta_i) \\
&= \sum_{i=1}^n w_{ni}(x;h_n)\xi_{tx}(T_i,\delta_i,x_i) \\
&\quad + (1-F_x(t))\left\{\int_0^t \frac{EH_{xh}-H_x}{(1-H_x)^2}dH_x^u + \int_0^t \frac{d(EH_{xh}^u - H_x^u)}{1-H_x}\right\}
\end{aligned}$$

with

$$\begin{aligned}
\xi_{tx}(T_i,\delta_i,x_i) &= g_{tx}(T_i,\delta_i) - Eg_{tx}(T_i,\delta_i) \\
&= (1-F_x(t))\left\{\int_0^t \frac{I(T_i\le s) - H_{x_i}(s)}{(1-H_x(s))^2}dH_x^u(s)\right. \\
&\qquad + \frac{I(T_i\le t,\delta_i=1) - H_{x_i}^u(t)}{1-H_x(t)} \\
&\qquad \left. - \int_0^t \frac{I(T_i\le s,\delta_i=1) - H_{x_i}^u(s)}{(1-H_x(s))^2}dH_x(s)\right\}.
\end{aligned}$$

If (C1) and (C2) hold, and if H_x and H_x^u satisfy (C3) and (C5), then we can apply Lemma A.1(b) to $EH_{xh} - H_x$ and to $EH_{xh}^u - H_x^u$. This gives that the second term in (3.2) equals, uniformly in t,

$$\frac{1}{2}(1-F_x(t))\int_0^t \left\{\frac{\ddot{H}_x(s)dH_x^u(s)}{(1-H_x(s))^2} + \frac{d\ddot{H}_x^u(s)}{1-H_x(s)}\right\}\mu_2^K h_n^2 + o(h_n^2) + O(n^{-1}). \tag{3.3}$$

We have $E\xi_{tx}(T_i,\delta_i,x_i) = 0$. In order to deal with $\sum_{i=1}^n w_{ni}^2(x;h_n)\operatorname{Var}\xi_{tx}(T_i,\delta_i,x_i)$ we recall the following lemma.

LEMMA 3.1. *Assume* (C1), (C2), $h_n \to 0$, $nh_n \to \infty$. *If* $\gamma : [0,1] \to \mathbb{R}$ *is a Lipschitz function, then, as* $n \to \infty$,

$$\sum_{i=1}^{n} w_{ni}^2(x;h_n)\gamma(x_i) = \frac{1}{nh_n}\gamma(x)\|K\|_2^2 + o((nh_n)^{-1}).$$

We omit the proof of this lemma since it is standard in all variance calculations with Gasser-Müller weights. This enables us to obtain the following result for the variance.

LEMMA 3.2. *Assume* (C1), (C2), $H_x(t)$ *and* $H_x^u(t)$ *satisfy* (C3) *in* $[0,T]$ *with* $T < T_{H_x}$, $h_n \to 0$, $nh_n \to \infty$. *Then, for* $t \le T$, *as* $n \to \infty$,

$$\operatorname{Var}\left(\sum_{i=1}^{n} w_{ni}(x;h_n)\xi_{tx}(T_i,\delta_i,x_i)\right) = \frac{\|K\|_2^2}{nh_n}(1-F_x(t))^2\int_0^t \frac{dH_x^u(s)}{(1-H_x(s))^2} + o((nh_n)^{-1}).$$

PROOF. Using integration by parts, we can write

$$g_{tx}(T_i,\delta_i) = (1-F_x(t))\left\{-\int_0^{\min(T_i,t)} \frac{dH_x^u(s)}{(1-H_x(s))^2} + \frac{I(T_i \le t, \delta_i = 1)}{1-H_x(T_i)}\right\}.$$

Hence, some straightforward calculations show that

$$\begin{aligned}(3.4)\quad &\operatorname{Var}(g_{tx}(T_i,\delta_i))\\ &= (1-F_x(t))^2\Bigg\{2\int_0^t \frac{1}{(1-H_x(s))^2}\int_s^t \frac{H_x(y)-H_{x_i}(y)}{(1-H_x(y))^2}dH_x^u(y)dH_x^u(s)\\ &\quad + 2\int_0^t \frac{1}{(1-H_x(s))^2}\int_s^t \frac{d(H_x^u(y)-H_{x_i}^u(y))}{1-H_x(y)}dH_x^u(s)\\ &\quad + \int_0^t \frac{dH_{x_i}^u(y)}{(1-H_x(y))^2}\\ &\quad - \left[-\int_0^t \frac{H_x(s)-H_{x_i}(s)}{(1-H_x(s))^2}dH_x^u(s)\right.\\ &\qquad \left. + \int_0^t \frac{d(H_{x_i}^u(s)-H_x^u(s))}{1-H_x(s)}\right]^2\Bigg\}\end{aligned}$$

from which the result follows via Lemma 3.1.

We are now ready to state the asymptotic normality result for the estimator $F_{xh}(t)$. The (a) part requires the condition $nh_n^5 \to 0$. Since the optimal bandwidth $h_n = Cn^{-1/5}$ for some $C > 0$ (i.e. the bandwidth which minimizes the approximate mean squared error) is not covered by (a), we state this case in part (b).

THEOREM 3.1. *Assume* (C1), (C2), $H_x(t)$ *and* $H_x^u(t)$ *satisfy* (C5), (C6) *and* (C7) *in* $[0,T]$ *with* $T < T_{H_x}$.

(a) *If* $nh_n^5 \to 0$ *and* $\frac{(\log n)^3}{nh_n} \to 0$, *then for* $t \le T$, *as* $n \to \infty$,

$$(nh_n)^{1/2}(F_{xh}(t) - F_x(t)) \xrightarrow{d} N(0; s_x^2(t)).$$

(b) *If* $h_n = Cn^{-1/5}$ *for some* $C > 0$, *then for* $t \le T$, *as* $n \to \infty$,

$$(nh_n)^{1/2}(F_{xh}(t) - F_x(t)) \xrightarrow{d} N(b_x(t); s_x^2(t))$$

where

$$b_x(t) = \frac{1}{2}(1 - F_x(t)) \int_0^t \left\{ \frac{\ddot{H}_x(s)dH_x^u(s)}{(1 - H_x(s))^2} + \frac{d\ddot{H}_x^u(s)}{1 - H_x(s)} \right\} \mu_2^K C^{5/2} \tag{3.5}$$

$$s_x^2(t) = \|K\|_2^2(1 - F_x(t))^2 \int_0^t \frac{dH_x^u(s)}{(1 - H_x(s))^2}. \tag{3.6}$$

PROOF. (a) The condition $nh_n^5 \to 0$ implies $\frac{nh_n^5}{\log n} = O(1)$. Hence by Theorem 2.1 and the fact that the bias term is $O(h_n^2 + n^{-1})$ (applying Lemma A.1), it follows that the limiting distribution of $(nh_n)^{1/2}(F_{xh}(t) - F_x(t))$ is the same as that of $(nh_n)^{1/2} \sum_{i=1}^n w_{ni}(x; h_n)\xi_{tx}(T_i, \delta_i, x_i)$. The result follows by checking Liapunov's condition. The Liapunov ratio is easily seen to be $O((nh_n)^{-1/2}) = o(1)$.

(b) Similar to (a), but taking into account the precise form of the bias term in (3.3).

Under the conditions of Theorem 3.1, it is also possible to show that the stochastic process $W_{nx}(t) = (nh_n)^{1/2}(F_{xh}(t) - F_x(t))$, $0 \le t \le T$ (where $T < T_{H_x}$) converges weakly in $D[0,T]$ (the space of right continuous functions with left hand limits endowed with the Skorohod topology) to a Gaussian process $W_x(t)$ with covariance function ($0 \le s \le t \le T$):

$$\mathrm{Cov}(W_x(t), W_x(s)) = \|K\|_2^2(1 - F_x(t))(1 - F_x(s)) \int_0^s \frac{dH_x^u(y)}{(1 - H_x(y))^2}.$$

This is the analogue of the result in Theorem 5 of Breslow and Crowley (1974) for the usual Kaplan-Meier estimator. In Van Keilegom and Veraverbeke (1997) we establish this result by showing the asymptotic normality of the finite dimensional distributions together with a tightness argument. The paper also covers the weak convergence of the corresponding bootstrapped process (where the bootstrap procedure is defined as in the next section) as well as the analogous results for the quantile process and the bootstrapped quantile process. As an application, confidence bands for both the distribution and quantile function are obtained.

4. The bootstrap procedure

In the rest of the paper we introduce a bootstrap procedure for approximating the distribution of $(nh_n)^{1/2}(F_{xh}(t) - F_x(t))$. This then provides us with an alternative to the normal approximation in Theorem 3.1 and avoids estimation of the complicated mean and variance parameters of the latter.

Our procedure combines both the bootstrap ideas of Efron (1981) for censored data and of Aerts *et al.* (1994*b*) for fixed design regression. Given the design points x_i, the responses Y_i and censoring times C_i $(i = 1, \ldots, n)$ we define the random variables Y_i^* and C_i^* (independently) as follows:

$$\begin{array}{lll} Y_1^*, \ldots, Y_n^* & \text{are independent;} & Y_i^* \sim F_{x_i g} \\ C_1^*, \ldots, C_n^* & \text{are independent;} & C_i^* \sim G_{x_i g}. \end{array}$$

Here $F_{x_i g}$ is the estimator for F_{x_i} as defined in (2.1), but with a bandwidth sequence $\{g_n\}$, which is different from $\{h_n\}$. The distribution $G_{x_i g}$ is the analogous estimator for G_{x_i}. Then define, for $i = 1, \ldots, n$,

$$T_i^* = \min(Y_i^*, C_i^*) \quad \text{and} \quad \delta_i^* = I(Y_i^* \leq C_i^*).$$

It is readily verified that the above procedure is equivalent to one where the pairs (T_i^*, δ_i^*) are drawn (with replacement) from $(T_1, \delta_1), \ldots, (T_n, \delta_n)$, giving probability $w_{nj}(x_i; g_n)$ to (T_j, δ_j) for $j = 1, \ldots, n$.

Based on the bootstrap sample $(T_1^*, \delta_1^*), \ldots, (T_n^*, \delta_n^*)$, the bootstrap analogue of the Kaplan-Meier type estimator in (2.1) is given by

$$F_{xhg}^*(t) = 1 - \left\{ \prod_{T_{(i)}^* \leq t} \left(1 - \frac{w_{n(i)}(x; h_n)}{1 - \sum_{j=1}^{i-1} w_{n(j)}(x; h_n)} \right)^{\delta_{(i)}^*} \right\} I(t < T_{(n)}^*) \tag{4.1}$$

where $T_{(1)}^* \leq \cdots \leq T_{(n)}^*$ and $\delta_{(i)}^*$ and $w_{n(i)}(x; h_n)$ correspond to $T_{(i)}^*$. In case of ties, we make the usual convention that uncensored observations are considered to occur just before censored observations. It is easy to see that F_{xhg}^* in (4.1) is well defined in the case that two or more observations occur at the same time and that formula (4.1) can also be written as

$$F_{xhg}^*(t) = 1 - \left\{ \prod_{T_{(i)} \leq t} \left(1 - \frac{\bar{\bar{w}}_{n(i)}(x; h_n)}{1 - \sum_{j=1}^{i-1} \bar{w}_{n(j)}(x; h_n)} \right) \right\} I(t < T_{(n)}^*)$$

where $\bar{w}_{n(i)}(x; h_n) = \sum_{k=1}^n w_{nk}(x; h_n) I(T_k^* = T_{(i)})$ and $\bar{\bar{w}}_{n(i)}(x; h_n) = \sum_{k=1}^n w_{nk}(x; h_n) I(T_k^* = T_{(i)}, \delta_k^* = 1)$. In a similar way, replacing the (T_i, δ_i) by (T_i^*, δ_i^*), the bootstrap analogues of $H_{xh}(t)$ and $H_{xh}^u(t)$ in (2.3) and (2.4) are denoted as $H_{xhg}^*(t)$ and $H_{xhg}^{*u}(t)$ respectively.

In the sequel we will use notations P^*, E^*, $\text{Var}^*, \ldots$ for probability, expectation, variance, ... conditionally on the original observations.

The parameter g_n which is used to construct the resampled values is an appropriate pilot bandwidth sequence which is typically asymptotically larger than h_n, i.e. $g_n/h_n \to \infty$ in a certain way. This technique of oversmoothing with the initial bandwidth has been successfully used in other resampling schemes in regression (e.g. Härdle and Mammen (1991), Aerts *et al.* (1994*b*)). It entails that the bootstrap bias and bootstrap variance are asymptotically appropriate estimators for the bias and variance terms.

It should be noted that the important problem of bandwidth selection is not dealt with here, merely because this problem is a research topic on its own and is beyond the scope of this paper. To the best of our knowledge, no research concerning bandwidth selection in the present situation has been done yet. We therefore mention below some ideas for future research on this interesting problem. We propose two methods, both of which consist of minimizing a specific estimate of the mean squared error defined as

$$MSE(h_n) = E((F_{xh}(t) - F_x(t))^2).$$

Primarily, one could prove an a.s. asymptotic representation for $MSE(h_n)$ of the form

$$MSE(h_n) = AMSE(h_n) + \text{remainder term}$$

and minimize the asymptotic $MSE(AMSE)$ with respect to h_n. This method, however, will require further estimation of certain unknown quantities in the expression of the $AMSE$ (see e.g. Sánchez-Sellero *et al.* (1995) where this "plug-in" method is used in the context of density estimation with censored and truncated data). To overcome this additional problem (it might require the selection of a second bandwidth), one could use a bootstrap bandwidth selection procedure, as done in e.g. González-Manteiga *et al.* (1996) in the context of hazard rate estimation. Instead of minimizing the MSE, h_n is determined here such that

$$MSE^*(h_n) = E^*((F^*_{xhg}(t) - F_{xg}(t))^2)$$

is minimal for a given sample (here, the pilot bandwidth g_n should first be estimated in an optimal way). Again, one could prove an a.s. asymptotic representation for $MSE^*(h_n)$ and minimize the dominant term in this representation. The main advantage of this method is that there are no unknown quantities to be estimated.

The results require a slightly stronger version of condition (C2). We will denote it by (C2′):

(C2′) K is a twice differentiable probability density function with finite support $[-L, L]$ for some $L > 0$, $\mu_1^K = 0$, K'' is continuous and $K(-L) = K'(-L) = K(L) = K'(L) = 0$.

From now on we state our results for the fixed bandwidth sequence h_n of optimal rate, i.e. $h_n = Cn^{-1/5}$ for some constant $C > 0$.

We begin with two lemmas which collect some properties of $H^*_{xhg}(t)$. Analogous results hold for $H^{*u}_{xhg}(t)$.

LEMMA 4.1. *Assume* (C1), (C2′), $H_x(t)$ *satisfies* (C5), (C6) *and* (C7) *in* $[0,T]$ *with* $T < T_{H_x}$, $h_n = Cn^{-1/5}$ *for some* $C > 0$, $g_n \to 0$, $\frac{ng_n^5}{\log n} \to \infty$. *Then, as* $n \to \infty$,

(a) $\sup_{0\le t\le T} |\operatorname{bias} H^*_{xhg}(t) - \operatorname{bias} H_{xh}(t)| = o((nh_n)^{-1/2})$ *a.s., where*

$$\operatorname{bias} H^*_{xhg}(t) = E^* H^*_{xhg}(t) - H_{xg}(t)$$
$$\operatorname{bias} H_{xh}(t) = EH_{xh}(t) - H_x(t).$$

(b) $\sup_{0\le t\le T} |H^*_{xhg}(t) - H_{xg}(t)| = O_{P^*}((nh_n)^{-1/2}(\log n)^{1/2})$ *a.s.*

PROOF. (a) Partitioning $[0,T]$ into $k_n = O(ng_n(\log n)^{-1})$ subintervals $[t_i, t_{i+1}]$ of length $O((ng_n)^{-1}\log n)$, we have that

$$\begin{aligned}
&\sup_{0\le t\le T} |\operatorname{bias} H^*_{xhg}(t) - \mathit{bias} H_{xh}(t)| \\
&\quad\le \max_{1\le i\le k_n} |\operatorname{bias} H^*_{xhg}(t_i) - \operatorname{bias} H_{xh}(t_i)| \\
&\qquad + \max_{1\le i\le k_n} |H_{xg}(t_i) - H_{xg}(t_{i-1}) + EH_{xh}(t_i) - EH_{xh}(t_{i-1})| \\
&\quad\le \max_{1\le i\le k_n} |\operatorname{bias} H^*_{xhg}(t_i) - \operatorname{bias} H_{xh}(t_i)| \\
&\qquad + \max_{1\le i\le k_n} |H_{xg}(t_i) - H_{xg}(t_{i-1}) - H_x(t_i) + H_x(t_{i-1})| \\
&\qquad + \max_{1\le i\le k_n} |EH_{xh}(t_i) - EH_{xh}(t_{i-1}) - H_x(t_i) + H_x(t_{i-1})| \\
&\qquad + 2\max_{1\le i\le k_n} |H_x(t_i) - H_x(t_{i-1})| \\
&\quad= T_1 + T_2 + T_3 + T_4.
\end{aligned}$$

For T_1 we have, as in the proof of Lemma 10 in Aerts *et al.* (1994*b*) that

$$(nh_n)^{1/2} T_1 = \frac{1}{2}\mu_2^K C^{5/2} \max_{1\le i\le k_n} |\ddot{H}_{xg}(t_i) - \ddot{H}_x(t_i)| + o(1)$$

where $\ddot{H}_{xg}(y) = \sum_{j=1}^n w_{nj}^{(2)}(x;g_n) I(T_j \le y)$ and $w_{nj}^{(2)}(x;g_n) = \frac{1}{g_n^3}\int_{x_{j-1}}^{x_j} K''(\frac{x-z}{g_n})dz$. Now,

$$\begin{aligned}
\max_{1\le i\le k_n} |\ddot{H}_{xg}(t_i) - \ddot{H}_x(t_i)| \le{}& \max_{1\le i\le k_n} |\ddot{H}_{xg}(t_i) - E\ddot{H}_{xg}(t_i)| \\
&+ \max_{1\le i\le k_n} |E\ddot{H}_{xg}(t_i) - \ddot{H}_x(t_i)|.
\end{aligned}$$

The first term is $o(1)$, using Bernstein's theorem. As to the second term, we write

$$\begin{aligned}
&\max_{1\le i\le k_n} |E\ddot{H}_{xg}(t_i) - \ddot{H}_x(t_i)| \\
&\quad\le \max_{1\le i\le k_n} \left| \sum_{j=1}^n w_{nj}^{(2)}(x;g_n) H_{x_j}(t_i) - \frac{1}{g_n^3}\int_0^{x_n} K''\left(\frac{x-z}{g_n}\right) H_z(t_i)dz \right|
\end{aligned}$$

$$+ \max_{1\le i\le k_n}\left|\frac{1}{g_n^3}\int_0^{x_n} K''\left(\frac{x-z}{g_n}\right)H_z(t_i)dz - \ddot{H}_x(t_i)\right|$$

$$\le \max_{1\le i\le k_n}\left\{\sum_{j=1}^n \frac{1}{g_n^3}\int_{x_{j-1}}^{x_j}\left|K''\left(\frac{x-z}{g_n}\right)\right| |H_{x_j}(t_i) - H_z(t_i)|dz\right\}$$

$$+ \max_{1\le i\le k_n}\left|\int_{-L}^{L} K(u)[\ddot{H}_{x-h_nu}(t_i) - \ddot{H}_x(t_i)]du\right| \qquad \text{(for } n \text{ large)}$$

which is $o(1)$. Using Lemma A.5 with $a_n = c(ng_n)^{-1}\log n$ for some $c > 0$, it follows that T_2 and T_3 are $o((nh_n)^{-1})$ a.s. Also, $T_4 = o((nh_n)^{-1/2})$, using the Lipschitz continuity of H_x.

(b) We write: $|H^*_{xhg}(t) - H_{xg}(t)| \le |H^*_{xhg}(t) - E^*H^*_{xhg}(t)| + |\,\text{bias}\, H^*_{xhg}(t) - \text{bias}\, H_{xh}(t)| + |\,\text{bias}\, H_{xh}(t)|$.

The second term is $o((nh_n)^{-1/2})$ a.s. by the (a) part of the lemma. The last term is $O(h_n^2 + n^{-1})$ using Lemma A.1(b). To the first term we apply Bernstein's inequality and the usual argument for replacing the supremum by a maximum: partitioning the interval $[0,T]$ in $O((nh_n)^{1/2}(\log n)^{-1/2})$ subintervals $[t_i, t_{i+1}]$ such that $H_{xg}(t_{i+1}) - H_{xg}(t_i) = O((nh_n)^{-1/2}(\log n)^{1/2})$ a.s. This is possible since the jump sizes of H_{xg} are of order $O((ng_n)^{-1})$ and since $\frac{ng_n^5}{\log n} \to \infty$.

LEMMA 4.2. *Assume* (C1), (C2), $h_n = Cn^{-1/5}$ *for some* $C > 0$, $g_n \to 0$, $\frac{ng_n^5}{\log n} \to \infty$ *and* $\frac{ng_n^5}{\log n}\frac{h_n}{g_n} = O(1)$.

(a) *If* $H_x(t)$ *satisfies* (C3), (C6) *and* (C7) *in* $[0,T]$ *with* $T < T_{H_x}$, *then as* $n \to \infty$,

$$\sup_{0\le s,t\le T}\ \sup_{|t-s|\le c(nh_n)^{-1/2}(\log n)^{1/2}} |E^*H^*_{xhg}(t) - E^*H^*_{xhg}(s) - H_{xg}(t) + H_{xg}(s)|$$
$$= O((nh_n)^{-3/4}(\log n)^{3/4}) \qquad a.s.$$

(b) *If* $H_x(t)$ *satisfies* (C3) *and* (C5) *in* $[0,T]$ *with* $T < T_{H_x}$, *then as* $n \to \infty$,

$$\sup_{0\le t\le T} |H_{xg}(t) - H_x(t)| = O((nh_n)^{-1/2}(\log n)^{1/2}) \qquad a.s.$$

(c) *If* (C2′) *holds and* $H_x(t)$ *and* $H^u_x(t)$ *satisfy* (C5), (C6) *and* (C7) *in* $[0,T]$ *with* $T < T_{H_x}$, *then as* $n \to \infty$,

$$\sup_{0\le t\le T}\left|\int_0^t\left(\frac{1}{1-H^*_{xhg}} - \frac{1}{1-H_{xg}}\right)d(H^{*u}_{xhg} - H^u_{xg})\right|$$
$$= O_{P^*}((nh_n)^{-3/4}(\log n)^{3/4}) \qquad a.s.$$

PROOF. (a) We write

$$\begin{aligned}
&E^* H^*_{xhg}(t) - E^* H^*_{xhg}(s) - H_{xg}(t) + H_{xg}(s) \\
&\quad = \sum_{i=1}^{n} w_{ni}(x; h_n)[H_{x_i g}(t) - H_{x_i g}(s) - H_{x_i}(t) + H_{x_i}(s)] \\
&\qquad + [EH_{xh}(t) - EH_{xh}(s) - H_x(t) + H_x(s)] \\
&\qquad + [H_x(t) - H_x(s) - H_{xg}(t) + H_{xg}(s)].
\end{aligned}$$

The second and the third term are of the required order by Lemma A.5. For the first term the proof is completely analogous to that of Lemma A.5.

(b) By using Lemma A.1(b) instead of Lemma A.1(a) in the proof of Lemma A.3(b), it follows that

$$\begin{aligned}
&P\left(\sup_{0 \le t \le T} |H_{xg}(t) - H_x(t)| > c(nh_n)^{-1/2}(\log n)^{1/2}\right) \\
&\quad \le d_0 c(nh_n)^{-1/2}(\log n)^{1/2} n g_n \exp\left(-\frac{1}{4} d_1 n g_n c^2 (nh_n)^{-1} \log n\right) \\
&\quad \le K n g_n n^{-d_1 c^2/4} \qquad \text{(for some } K > 0 \text{ and for } n \text{ large)}
\end{aligned}$$

provided $\frac{ng_n^5}{\log n}\frac{h_n}{g_n} = O(1)$. Now apply Borel-Cantelli after proper choice of c.

(c) The proof parallels completely that of Lemma 2.1 above: the same partitionings and the same inequalities. Also, use is made of part (b) of Lemma 4.1 and parts (a) and (b) of Lemma 4.2.

THEOREM 4.1. *Assume* (C1), (C2′), $H_x(t)$ *and* $H_x^u(t)$ *satisfy* (C5), (C6) *and* (C7) *in* $[0, T]$ *with* $T < T_{H_x}$, $h_n = Cn^{-1/5}$ *for some* $C > 0$, $g_n \to 0$, $\frac{ng_n^5}{\log n} \to \infty$ *and* $\frac{ng_n^5}{\log n}\frac{h_n}{g_n} = O(1)$. *Then, for* $t < T_{H_x}$,

$$F^*_{xhg}(t) - F_{xg}(t) = \sum_{i=1}^{n} w_{ni}(x; h_n) g_{tx}(T_i^*, \delta_i^*) - \sum_{i=1}^{n} w_{ni}(x; g_n) g_{tx}(T_i, \delta_i) + r_n^*(t, x)$$

where g_{tx} *is as in Theorem* 2.1 *and where, as* $n \to \infty$,

$$\sup_{0<t<T} |r_n^*(t, x)| = O_{P^*}((nh_n)^{-3/4}(\log n)^{3/4}) \qquad a.s.$$

PROOF. We have

$$\begin{aligned}
(4.2)\quad &F^*_{xhg}(t) - F_{xg}(t) \\
&\quad = (1 - F_{xg}(t))[1 - e^{\log(1 - F^*_{xhg}(t)) - \log(1 - F_{xg}(t))}] \\
&\quad = (1 - F_{xg}(t))\Big\{-[\log(1 - F^*_{xhg}(t)) - \log(1 - F_{xg}(t))] \\
&\qquad\qquad - \frac{1}{2}[\log(1 - F^*_{xhg}(t)) - \log(1 - F_{xg}(t))]^2 e^{\theta_n}\Big\} \\
&\quad = (1 - F_{xg}(t))\left\{A - \frac{1}{2}B\right\}
\end{aligned}$$

where θ_n is between 0 and $\log(1 - F^*_{xhg}(t)) - \log(1 - F_{xg}(t))$. It is easy to show that

$$A = \int_0^t \frac{H^*_{xhg} - H_{xg}}{(1 - H_{xg})^2} dH^u_{xg} + \frac{H^{*u}_{xhg}(t) - H^u_{xg}(t)}{1 - H_{xg}(t)} - \int_0^t \frac{H^{*u}_{xhg} - H^u_{xg}}{(1 - H_{xg})^2} dH_{xg}$$
$$+ R_{n1}(t) + R_{n2}(t) + R_{n3}(t) + R_{n4}(t)$$

where

$$R_{n1}(t) = -\log(1 - F^*_{xhg}(t)) - \int_0^t \frac{dH^{*u}_{xhg}}{1 - H^*_{xhg}}$$
$$R_{n2}(t) = \int_0^t \frac{(H^*_{xhg} - H_{xg})^2}{(1 - H_{xg})^2(1 - H^*_{xhg})} dH^u_{xg}$$
$$R_{n3}(t) = \int_0^t \left(\frac{1}{1 - H^*_{xhg}} - \frac{1}{1 - H_{xg}} \right) d(H^{*u}_{xhg} - H^u_{xg})$$
$$R_{n4}(t) = \int_0^t \frac{dH^u_{xg}}{1 - H_{xg}} + \log(1 - F_{xg}(t)).$$

Direct application of Lemma 2.2 gives

$$\sup_{0 \le t \le T} |R_{n4}(t)| = O((ng_n)^{-1}) = O((nh_n)^{-1}) \qquad \text{a.s.}$$

Following the lines of the proof of Lemma 2.2, we obtain

$$\sup_{0 \le t \le T} |R_{n1}(t)| \le 2 \max_{1 \le i \le n} w_{ni}(x; h_n) \frac{1}{(1 - H^*_{xhg}(T))^2} = O_{P^*}((nh_n)^{-1}) \qquad \text{a.s.}$$

since $\max_{1 \le i \le n} w_{ni}(x; h_n) = O((nh_n)^{-1})$ and $H^*_{xhg}(T) \stackrel{P^*}{\to} H_x(T)$ by application of Lemma A.2 and Lemma 4.1(b). Clearly, $\sup_{0 \le t \le T} |R_{n2}(t)| = O_{P*}((nh_n)^{-1} \log n)$ a.s. by Lemma 4.1(b). For $R_{n3}(t)$, we use Lemma 4.2(c). The first term in A equals

$$\int_0^t \frac{H^*_{xhg} - H_{xg}}{(1 - H_x)^2} dH^u_x + R_{n5}(t) + R_{n6}(t)$$

where

$$R_{n5}(t) = \int_0^t (H^*_{xhg} - H_{xg}) \left(\frac{1}{(1 - H_{xg})^2} - \frac{1}{(1 - H_x)^2} \right) dH^u_{xg}$$
$$R_{n6}(t) = \int_0^t \frac{H^*_{xhg} - H_{xg}}{(1 - H_x)^2} d(H^u_{xg} - H^u_x).$$

$R_{n5}(t)$ is uniformly bounded by

$$\frac{2}{(1 - H_{xg}(T))^2 (1 - H_x(T))^2} \sup_{0 \le t \le T} |H^*_{xhg}(t) - H_{xg}(t)| \sup_{0 \le t \le T} |H_{xg}(t) - H_x(t)|$$

and this is $O_{P^*}((nh_n)^{-1}\log n)$ using Lemmas 4.1(b) and 4.2(b). Using some analogues of Lemma 2.1, it is easy to show that $\sup_{0\le t\le T}|R_{n6}(t)| = O_{P^*}((nh_n)^{-3/4}(\log n)^{3/4})$ a.s. The second and third term in A can be worked out in a similar way. Hence,

$$A - \int_0^t \frac{H^*_{xhg} - H_{xg}}{(1-H_x)^2}dH^u_x + \frac{H^{*u}_{xhg}(t) - H^u_{xg}(t)}{1-H_x(t)} - \int_0^t \frac{H^{*u}_{xhg} - H^u_{xg}}{(1-H_x)^2}dH_x + \rho^*_n(t)$$
$$= (1-F_x(t))^{-1}\left[\sum_{i=1}^n w_{ni}(x;h_n)g_{tx}(T^*_i,\delta^*_i) - \sum_{i=1}^n w_{ni}(x;g_n)g_{tx}(T_i,\delta_i)\right] + \rho^*_n(t)$$

where $\sup_{0\le t\le T}|\rho^*_n(t)| = O_{P^*}((nh_n)^{-3/4}(\log n)^{3/4})$ a.s.

To deal with the term $-\frac{1}{2}B$ in (4.2), we first note that $B = A^2e^{\theta_n}$ and that $e^{\theta_n} \le \frac{1}{1-F_{xg}(T)}$. From Lemma 4.1(b) it follows that $A = O_{P^*}((nh_n)^{-1/2}(\log n)^{1/2})$ a.s. Hence, by Lemma 4.3 below, $B = O_{P^*}((nh_n)^{-1}\log n)$ a.s. To complete the proof of Theorem 4.1, we still have to replace the factor $1 - F_{xg}(t)$ in (4.2) by $1 - F_x(t)$. This is allowed by Lemma 4.3 below.

LEMMA 4.3. *Assume* (C1), (C2), $H_x(t)$ *and* $H^u_x(t)$ *satisfy* (C3) *and* (C5) *in* $[0,T]$ *with* $T < T_{H_x}$, $h_n = Cn^{-1/5}$ *for some* $C > 0$, $g_n \to 0$, $\frac{ng_n^5}{\log n} \to \infty$ *and* $\frac{ng_n^5}{\log n}\frac{h_n}{g_n} = O(1)$. *Then, as* $n \to \infty$,

$$\sup_{0\le t\le T}|F_{xg}(t) - F_x(t)| = O((nh_n)^{-1/2}(\log n)^{1/2}) \qquad a.s.$$

PROOF. Applying partial integration on $\int_0^t \frac{d(1-F_{xg}(s))}{1-F_x(s)}$ and using that $\Lambda_{xg}(t) = \int_0^t \frac{dH^u_{xg}(s)}{1-H_{xg}(s-)} = \int_0^t \frac{dF_{xg}(s)}{1-F_{xg}(s-)}$, it is easy to show that

$$F_{xg}(t) - F_x(t) = (1-F_x(t))\int_0^t \frac{1-F_{xg}(s-)}{1-F_x(s)}d(\Lambda_{xg}(s) - \Lambda_x(s))$$

and hence that $\sup_{0\le t\le T}|F_{xg}(t) - F_x(t)| \le 3\sup_{0\le t\le T}|\Lambda_{xg}(t) - \Lambda_x(t)|$. For $1 - H_x(T) > \delta > 0$, some easy calculations show that

$$P\left(\sup_{0\le t\le T}|F_{xg}(t) - F_x(t)| > \varepsilon\right)$$
$$\le 2P\left(\sup_{0\le t\le T}|H_{xg}(t) - H_x(t)| > \frac{\varepsilon\delta^2}{12}\right) + P\left(\sup_{0\le t\le T}|H^u_{xg}(t) - H^u_x(t)| > \frac{\varepsilon\delta^2}{12}\right)$$

and hence, using Lemma 4.2(b),

$$\sup_{0\le t\le T}|F_{xg}(t) - F_x(t)| = O((nh_n)^{-1/2}(\log n)^{1/2}) \qquad \text{a.s.}$$

5. Strong consistency of the bootstrap approximation

In this section we show the validity of the proposed bootstrap procedure by proving that $P^*((nh_n)^{1/2}(F^*_{xhg}(t) - F_{xg}(t)) \le y)$ is a strongly consistent estimator for $P((nh_n)^{1/2}(F_{xh}(t) - F_x(t)) \le y)$. We prove this result below for a bandwidth h_n with optimal rate: $h_n = Cn^{-1/5}$. In this way the bootstrap distribution is an alternative for the normal approximation in Theorem 3.1(b) which is $\Phi((y - b_x(t))/s_x(t))$ where $b_x(t)$ and $s^2_x(t)$ are the bias and variance parameters as given in (3.5) and (3.6).

THEOREM 5.1. *Assume* (C1), (C2′), $H_x(t)$ *and* $H^u_x(t)$ *satisfy* (C5), (C6) *and* (C7) *in* $[0, T]$ *with* $T < T_{H_x}$, $h_n = Cn^{-1/5}$ *for some* $C > 0$, $g_n \to 0$, $\frac{ng_n^5}{\log n} \to \infty$ *and* $\frac{ng_n^5}{\log n}\frac{h_n}{g_n} = O(1)$. *Then, for* $t \le T$, *as* $n \to \infty$,

$$\sup_{y\in\mathbb{R}} |P^*((nh_n)^{1/2}(F^*_{xhg}(t) - F_{xg}(t)) \le y) - P((nh_n)^{1/2}(F_{xh}(t) - F_x(t)) \le y)| = o(1) \quad a.s.$$

PROOF. Since from Theorem 3.1,

$$\sup_{y\in\mathbb{R}} \left| P((nh_n)^{1/2}(F_{xh}(t) - F_x(t)) \le y) - \Phi\left(\frac{y - b_x(t)}{s_x(t)}\right)\right| = o(1)$$

we only have to show that

$$\sup_{y\in\mathbb{R}} \left| P^*((nh_n)^{1/2}(F^*_{xhg}(t) - F_{xg}(t)) \le y) - \Phi\left(\frac{y - b_x(t)}{s_x(t)}\right)\right| = o(1) \quad \text{a.s.}$$

Since from Theorem 4.1, $P^*((nh_n)^{1/2}|r^*_n(t, x)| > \varepsilon) \to 0$ a.s., it suffices to show that

$$\sup_{y\in\mathbb{R}} \left| P^*\left((nh_n)^{1/2}\left(\sum_{i=1}^n w_{ni}(x; h_n)g_{tx}(T^*_i, \delta^*_i) - \sum_{i=1}^n w_{ni}(x; g_n)g_{tx}(T_i, \delta_i)\right) \le y\right) - \Phi\left(\frac{y - b_x(t)}{s_x(t)}\right)\right| = o(1) \quad \text{a.s.}$$

Now, with the shorthand notation G^*_i for $g_{tx}(T^*_i, \delta^*_i)$ and G_i for $g_{tx}(T_i, \delta_i)$, and using the inequality $\sup_{x\in\mathbb{R}} |\Phi(a+bx) - \Phi(x)| \le |a| + \max(b, b^{-1}) - 1$, this expression is bounded above by

$$(5.1) \quad \sup_{y\in\mathbb{R}} \left| P^*\left(\frac{\sum_{i=1}^n w_{ni}(x; h_n)G^*_i - E^*(\sum_{i=1}^n w_{ni}(x; h_n)G^*_i)}{(\mathrm{Var}^*(\sum_{i=1}^n w_{ni}(x; h_n)G^*_i))^{1/2}} \le y\right) - \Phi(y)\right|$$
$$+ \frac{(nh_n)^{1/2}|\hat{b}_{xn}(t) - b_{xn}(t)| + |(nh_n)^{1/2}b_{xn}(t) - b_x(t)|}{((nh_n)\,\mathrm{Var}^*(\sum_{i=1}^n w_{ni}(x; h_n)G^*_i))^{1/2}}$$
$$+ \max\left\{\frac{s_x(t)}{((nh_n)\,\mathrm{Var}^*(\sum_{i=1}^n w_{ni}(x; h_n)G^*_i))^{1/2}}, \frac{((nh_n)\,\mathrm{Var}^*(\sum_{i=1}^n w_{ni}(x; h_n)G^*_i))^{1/2}}{s_x(t)}\right\} - 1$$

where

$$\hat{b}_{xn}(t) = \sum_{i=1}^{n} w_{ni}(x; h_n) E^*(G_i^*) - \sum_{i=1}^{n} w_{ni}(x; g_n) G_i \tag{5.2}$$

$$b_{xn}(t) = \sum_{i=1}^{n} w_{ni}(x; h_n) E(G_i). \tag{5.3}$$

Since the G_i^* are conditionally independent we can show that the first term in (5.1) is $o(1)$ a.s. by checking Liapunov's condition

$$\frac{\sum_{i=1}^{n} w_{ni}^3(x; h_n) E^*|G_i^* - E^*(G_i^*)|^3}{(\mathrm{Var}^*(\sum_{i=1}^{n} w_{ni}(x; h_n) G_i^*))^{3/2}} = o(1) \qquad \text{a.s.} \tag{5.4}$$

Since $E^*|G_i^* - E^*(G_i^*)|^3 \leq 4\{E^*|G_i^*|^3 + (E^*|G_i^*|)^3\}$ and since the G_i^* are uniformly bounded (by $3(1 - H_x(T))^{-2}$), the numerator in (5.4) is $O(\sum_{i=1}^{n} w_{ni}^3(x; h_n)) = O((nh_n)^{-2})$ a.s. The denominator in (5.4) is $O((nh_n)^{-3/2})$ a.s. by Lemma 5.1 below, so that the Liapunov ratio in (5.4) is $O((nh_n)^{-1/2})$ a.s. For the second and third term in (5.1), we use Lemma 5.1 below and (3.3) to see that these are $o(1)$ a.s. as $n \to \infty$. This proves the theorem.

It remains to prove a lemma on the bootstrap bias and variance which was used in the above theorem.

LEMMA 5.1. *Assume* (C1), $h_n = Cn^{-1/5}$ *for some* $C > 0$, $g_n \to 0$, $\frac{ng_n^5}{\log n} \to \infty$, $t \leq T < T_{H_x}$.

(a) *If* (C2′) *holds*, $H_x(t)$ *and* $H_x^u(t)$ *satisfy* (C5), (C6) *and* (C7) *in* $[0, T]$, *then, as* $n \to \infty$,

$$(nh_n)^{1/2}(\hat{b}_{xn}(t) - b_{xn}(t)) \to 0 \qquad a.s.$$

where $\hat{b}_{xn}(t)$ *and* $b_{xn}(t)$ *are given by* (5.2) *and* (5.3).

(b) *If* (C2) *holds*, $H_x(t)$ *and* $H_x^u(t)$ *satisfy* (C3) *in* $[0, T]$, *then, as* $n \to \infty$,

$$(nh_n)\,\mathrm{Var}^*\left(\sum_{i=1}^{n} w_{ni}(x; h_n) g_{tx}(T_i^*, \delta_i^*)\right) \to s_x^2(t) \qquad a.s.$$

PROOF. (a) Because

$$\begin{aligned}
&(nh_n)^{1/2}|\hat{b}_{xn}(t) - b_{xn}(t)| \\
&\quad = (nh_n)^{1/2}(1 - F_x(t)) \\
&\qquad \cdot \left| \int_0^t \frac{1}{(1 - H_x(y))^2}[E^* H_{xhg}^*(y) - H_{xg}(y) - EH_{xh}(y) + H_x(y)] dH_x^u(y) \right. \\
&\qquad + \frac{1}{1 - H_x(t)}[E^* H_{xhg}^{*u}(t) - H_{xg}^u(t) - EH_{xh}^u(t) + H_x^u(t)] \\
&\qquad \left. - \int_0^t \frac{1}{(1 - H_x(y))^2}[E^* H_{xhg}^{*u}(y) - H_{xg}^u(y) - EH_{xh}^u(y) + H_x^u(y)] dH_x(y) \right|
\end{aligned}$$

the result follows from Lemma 4.1(a).

(b) The left hand side equals $(nh_n)\sum_{i=1}^n w_{ni}^2(x;h_n)h(x_i)$ where $h(x_i)$ is obtained from the expression (3.4) by replacing every H_{x_i} by H_{x_ig} and every $H_{x_i}^u$ by $H_{x_ig}^u$. Since $\mathrm{Var}(\sum_{i=1}^n w_{ni}(x;h_n)g_{tx}(T_i,\delta_i)) \to s_x^2(t)$ as $n \to \infty$ (see Lemma 3.2), it suffices to prove that

$$\max_{1\leq i\leq n}\sup_{0\leq t\leq T}|H_{x_ig}(t)-H_{x_i}(t)|\left(\sum_{i=1}^n w_{ni}^2(x;h_n)\right) = o((nh_n)^{-1}) \qquad \text{a.s.}$$

which is obviously satisfied.

Appendix

In this Appendix we prove some basic results for empirical distribution functions of the kernel type which play a major role in fixed design regression models. In the paper these results are frequently applied to either H_{xh} or H_{xh}^u in (2.3) and (2.4). We state them for a general empirical

$$L_{xh}(t) = \sum_{i=1}^n w_{ni}(x;h_n)I(Z_i \leq t)$$

which is an estimator for the (sub)distribution function

$$L_x(t) = P(Z_x \leq t)$$

and where $Z_1,\ldots,Z_n$ are independent random variables with (sub)distributions $L_{x_1},\ldots,L_{x_n}$ and Z_x is the response at an $x \in [0,1]$. (The proofs go through in exactly the same way for empiricals of the type $\sum_{i=1}^n w_{ni}(x;h_n)I(Z_i \leq t, \delta_i = 1)$.)

We start with a result on bias and variance which is well known and can be found in e.g. Aerts *et al.* (1994*a*).

LEMMA A.1. (Bias and variance)

(a) *Assume* (C1), (C2), $L_x(t)$ *satisfies* (C3), $h_n \to 0$, $nh_n \to \infty$. *Then, as* $n \to \infty$,

$$\sup_{0\leq t\leq T}|EL_{xh}(t) - L_x(t)| = o(h_n).$$

(b) *Assume* (C1), (C2), $L_x(t)$ *satisfies* (C3) *and* (C5), $h_n \to 0$. *Then, as* $n \to \infty$,

$$\sup_{0\leq t\leq T}|EL_{xh}(t) - L_x(t)| = O(h_n^2 + n^{-1}).$$

More in particular:

$$\sup_{0\leq t\leq T}\left|EL_{xh}(t) - L_x(t) - \frac{1}{2}\mu_2^K\ddot{L}_x(t)h_n^2\right| = o(h_n^2) + O(n^{-1}).$$

(c) *Assume* (C1), (C2), $L_x(t)$ *satisfies* (C3), $h_n \to 0$, $nh_n \to \infty$. *Then, as* $n \to \infty$,

$$\operatorname{Var} L_{xh}(t) = \frac{1}{nh_n} L_x(t)(1 - L_x(t))\|K\|_2^2 + o((nh_n)^{-1}).$$

LEMMA A.2. (Pointwise strong consistency)
Assume (C1), (C2), $L_x(t)$ *satisfies* (C3), $h_n \to 0$, $\frac{\log n}{nh_n} \to 0$. *Then, as* $n \to \infty$, *for* $t \leq T$:

$$L_{xh}(t) - L_x(t) \to 0 \qquad a.s.$$

PROOF. With $X_{in} = w_{ni}(x;h_n)[I(Z_i \leq t) - L_{x_i}(t)]$ we have $L_{xh}(t) - L_x(t) = \sum_{i=1}^n X_{in} + EL_{xh}(t) - L_x(t)$, and by Lemma A.1(a) it suffices to prove strong consistency of $\sum_{i=1}^n X_{in}$. We have: $|X_{in}| \leq w_{ni}(x;h_n) \leq \|K\|_\infty \overline{\Delta}_n/h_n$. Also $EX_{in} = 0$ and $\sum_{i=1}^n \operatorname{Var}(X_{in}) = \sum_{i=1}^n w_{ni}^2(x;h_n)L_{x_i}(t)(1 - L_{x_i}(t)) \leq \sum_{i=1}^n w_{ni}^2(x;h_n) \leq \|K\|_\infty \overline{\Delta}_n/h_n$. Hence, by Bernstein's inequality (see e.g. Serfling (1980)), for all $\varepsilon > 0$,

$$P\left(\left|\sum_{i=1}^n X_{in}\right| > \varepsilon\right) \leq 2\exp(-c\varepsilon^2 h_n/\overline{\Delta}_n)$$

for some constant $c > 0$. By (C1) and the condition $\frac{\log n}{nh_n} \to 0$, the right hand side can be made integrable.

LEMMA A.3. (Dvoretzky-Kiefer-Wolfowitz type exponential bounds)
Assume (C1), (C2), $h_n \to 0$, $nh_n \to \infty$.
(a) *For* $\varepsilon > 0$ *and* n *sufficiently large such that*

$$\varepsilon^2 \geq \frac{3}{2}\|K\|_2^2 \frac{1}{nh_n} \tag{A.1}$$

we have for any $T > 0$

$$P\left(\sup_{0 \leq t \leq T} |L_{xh}(t) - EL_{xh}(t)| > \varepsilon\right) \leq d_0 \varepsilon nh_n e^{-d_1 nh_n \varepsilon^2}. \tag{A.2}$$

(b) *If moreover* $L_x(t)$ *satisfies* (C3), *then for* $\varepsilon > 0$ *and* n *sufficiently large such that*

$$\varepsilon^2 \geq \max\left(6\|K\|_2^2 \frac{1}{nh_n}, 16C_L^2 \left(\int |u|K(u)du\right)^2 h_n^2\right) \tag{A.3}$$

we have

$$P\left(\sup_{0 \leq t \leq T} |L_{xh}(t) - L_x(t)| > \varepsilon\right) \leq \frac{1}{2} d_0 \varepsilon nh_n e^{-(d_1 nh_n \varepsilon^2)/4}. \tag{A.4}$$

Here d_0 *and* d_1 *are absolute constants* ($d_0 = 8e^2/\|K\|_2^2, d_1 = 4/(3\|K\|_2^2)$).

PROOF. (a) Applying a general exponential bound result of Singh (1975) gives that the left hand side of (A.2) is bounded by

$$\frac{4e^2\varepsilon}{\sum_{i=1}^n w_{ni}^2(x;h_n)} \exp\left\{-2\frac{\varepsilon^2}{\sum_{i=1}^n w_{ni}^2(x;h_n)}\right\}$$

provided $\varepsilon^2 \geq \sum_{i=1}^n w_{ni}^2(x;h_n)$. Now,

$$\begin{aligned}
&\sum_{i=1}^n w_{ni}^2(x;h_n) - \frac{1}{nh_n}\|K\|_2^2 \\
&= \sum_{i=1}^n \left(\int_{x_{i-1}}^{x_i} \frac{1}{h_n} K\left(\frac{x-z}{h_n}\right) dz\right)^2 - \sum_{i=1}^n \frac{x_i - x_{i-1}}{h_n} \int_{x_{i-1}}^{x_i} \frac{1}{h_n} K^2\left(\frac{x-z}{h_n}\right) dz \\
&\quad + \sum_{i=1}^n \frac{x_i - x_{i-1}}{h_n} \int_{x_{i-1}}^{x_i} \frac{1}{h_n} K^2\left(\frac{x-z}{h_n}\right) dz - \frac{1}{nh_n}\|K\|_2^2 \\
&= O((nh_n)^{-2}) + o((nh_n)^{-1}) = o((nh_n)^{-1})
\end{aligned}$$

(see also Aerts and Geertsema (1990)).

Hence, for n sufficiently large, $\frac{1}{2}\|K\|_2^2 \frac{1}{nh_n} \leq \sum_{i=1}^n w_{ni}^2(x;h_n) \leq \frac{3}{2}\|K\|_2^2 \frac{1}{nh_n}$. This, together with condition (A.1) on ε gives the desired bound.

(b) For $\varepsilon > 0$, the left hand side in (A.4) is bounded above by

$$P\left(\sup_{0\leq t\leq T} |L_{xh}(t) - EL_{xh}(t)| > \varepsilon - \sup_{0\leq t\leq T} |EL_{xh}(t) - L_x(t)|\right). \tag{A.5}$$

Now, from the proof of Lemma A.1(a) and the condition (A.3) on ε,

$$\sup_{0\leq t\leq T} |EL_{xh}(t) - L_x(t)| \leq 2C_L\left(\int |u|K(u)du\right) h_n \leq \frac{\varepsilon}{2}.$$

Again by the condition on ε,

$$\left(\varepsilon - \sup_{0\leq t\leq T} |EL_{xh}(t) - L_x(t)|\right)^2 \geq \frac{\varepsilon^2}{4} \geq \frac{3}{2}\|K\|_2^2 \frac{1}{nh_n}.$$

This allows to apply the (a)-part to (A.5) which leads to the bound in (A.4).

LEMMA A.4. (Rates of uniform strong consistency)

(a) *Assume* (C1), (C2), $L_x(t)$ *satisfies* (C3), $h_n \to 0$, $nh_n \to \infty$, $\frac{nh_n^3}{\log n} = O(1)$. *Then, as* $n \to \infty$,

$$\sup_{0\leq t\leq T} |L_{xh}(t) - L_x(t)| = O((nh_n)^{-1/2}(\log n)^{1/2}) \quad a.s.$$

(b) *Assume* (C1), (C2), $L_x(t)$ *satisfies* (C3) *and* (C5), $h_n \to 0$, $\frac{nh_n^5}{\log n} = O(1)$. *Then, as* $n \to \infty$,

$$\sup_{0\leq t\leq T} |L_{xh}(t) - L_x(t)| = O((nh_n)^{-1/2}(\log n)^{1/2}) \quad a.s.$$

PROOF. (a) Apply Lemma A.3(b) with the choice $\varepsilon = \varepsilon_n = c(nh_n)^{-1/2}(\log n)^{1/2}$ for some appropriately chosen constant $c > 0$. Apply Borel-Cantelli.

(b) This follows in a similar way if we use Lemma A.1(b) in the proof of Lemma A.3(b).

LEMMA A.5. (Almost sure behaviour of the modulus of continuity)

(a) *Assume* (C1), (C2), $L_x(t)$ *satisfies* (C4), $h_n \to 0$. *Let* $\{a_n\}$ *be a sequence of positive constants, tending to* 0 *as* $n \to \infty$, *with*

$$a_n(nh_n)(\log n)^{-1} > \Delta > 0$$

for all n *sufficiently large. Then, as* $n \to \infty$,

$$\text{(A.6)} \qquad \sup_{0 \le s,t \le T} \sup_{|t-s| \le a_n} |L_{xh}(t) - L_{xh}(s) - EL_{xh}(t) + EL_{xh}(s)| \\ = O(a_n^{1/2}(nh_n)^{-1/2}(\log n)^{1/2}) \qquad a.s.$$

(b) *Assume* (C1), (C2), $L_x(t)$ *satisfies* (C3), (C6) *and* (C7), $h_n \to 0$. *Let* $\{a_n\}$ *be any sequence of positive constants, tending to* 0 *as* $n \to \infty$. *Then, as* $n \to \infty$,

$$\sup_{0 \le s,t \le T} \sup_{|t-s| \le a_n} |EL_{xh}(t) - EL_{xh}(s) - L_x(t) + L_x(s)| = O(n^{-1} + a_nh_n + a_n^2).$$

PROOF. (a) Partition the interval $[0,T]$ into $m = [\frac{T}{a_n}]$ subintervals of length $\bar{a}_n = \frac{T}{m} : 0 = t_0 < t_1 < \cdots < t_m = T$ with $t_i = i\bar{a}_n$ for $i = 0, \ldots, m$. Let $I_{ni} = [t_i - \bar{a}_n, t_i + \bar{a}_n]$, $i = 1, \ldots, m-1$. We have: $a_n \le \bar{a}_n < 2a_n$ for n large. Hence, for $s, t \in [0,T]$ with $|t - s| \le a_n$, there exists an interval I_{ni} such that $t, s \in I_{ni}$. Partition each interval I_{ni} by a grid $t_{ij} = t_i + j\frac{\bar{a}_n}{b_n}$, $j = -b_n, \ldots, b_n$, where $\{b_n\}$ is a sequence of positive integers such that $b_n \sim a_n^{1/2}(nh_n)^{1/2}(\log n)^{-1/2}$. Using the monotonicity of $L_{xh}(t)$ and $EL_{xh}(t)$, we have that the left hand side in (A.6) is majorized by

$$\text{(A.7)} \qquad \max_{1 < i < m-1} \max_{-b_n < j,k < b_n} |L_{xh}(t_{ik}) - L_{xh}(t_{ij}) - EL_{xh}(t_{ik}) + EL_{xh}(t_{ij})| \\ + 2 \max_{1 \le i \le m-1} \max_{-b_n \le j \le b_n - 1} |EL_{xh}(t_{i,j+1}) - EL_{xh}(t_{ij})|.$$

From the Lipschitz continuity of L_x (implied by condition (C4)), it follows that the second term in (A.7) is $O(\frac{\bar{a}_n}{b_n}) = O(\frac{a_n}{b_n}) = O(a_n^{1/2}(nh_n)^{-1/2}(\log n)^{1/2})$. As to the first term in (A.7), we have that $L_{xh}(t_{ik}) - L_{xh}(t_{ij}) - EL_{xh}(t_{ik}) + EL_{xh}(t_{ij}) = \sum_{r=1}^{n} X_{rijk}$, where $X_{rijk} = w_{nr}(x; h_n)\{[I(Z_r \le t_{ik}) - I(Z_r \le t_{ij})] - [L_{x_r}(t_{ik}) - L_{x_r}(t_{ij})]\}$. We have: $|X_{rijk}| \le w_{nr}(x; h_n) \le \|K\|_\infty \overline{\Delta}_n / h_n$, $E(X_{rijk}) = 0$ and

$$\begin{aligned} \operatorname{Var}(X_{rijk}) &= w_{nr}^2(x; h_n)\{L_{x_r}(t_{ik})(1 - L_{x_r}(t_{ik})) + L_{x_r}(t_{ij})(1 - L_{x_r}(t_{ij})) \\ &\qquad - 2(L_{x_r}(\min(t_{ik}, t_{ij})) - L_{x_r}(t_{ik})L_{x_r}(t_{ij}))\} \\ &= w_{nr}^2(x; h_n)\{-(L_{x_r}(t_{ik}) - L_{x_r}(t_{ij}))^2 \\ &\qquad + [L_{x_r}(t_{ik}) - L_{x_r}(\min(t_{ik}, t_{ij}))] \\ &\qquad + [L_{x_r}(t_{ij}) - L_{x_r}(\min(t_{ik}, t_{ij}))]\} \\ &\le Cw_{nr}^2(x; h_n)a_n \end{aligned}$$

for some constant $C > 0$, using the Lipschitz continuity of L_x. It follows that $\sum_{r=1}^n \mathrm{Var}(X_{rijk}) \leq C\|K\|_\infty \overline{\Delta}_n a_n / h_n$. Let $\lambda_n = c_1 a_n^{1/2}(nh_n)^{-1/2}(\log n)^{1/2}$, where c_1 is a positive constant to be specified further on. By Bernstein's inequality,

$$P\left(\max_{1\leq i\leq m-1} \max_{-b_n\leq j,k\leq b_n} |L_{xh}(t_{ik}) - L_{xh}(t_{ij}) - EL_{xh}(t_{ik}) + EL_{xh}(t_{ij})| > \lambda_n\right)$$
$$\leq 2(m-1)(2b_n+1)^2 \exp\left\{-\lambda_n^2 \Big/ \left(2C\frac{\|K\|_\infty\overline{\Delta}_n}{h_n}a_n + \frac{2}{3}\frac{\|K\|_\infty\overline{\Delta}_n}{h_n}\lambda_n\right)\right\}.$$

From the condition in the lemma it follows that $\lambda_n \leq c_1 \frac{a_n}{\Delta^{1/2}}$ for n large, so that the bound becomes $2(m-1)(2b_n+1)^2 \exp\{-\frac{c'}{c''+c_1} c_1^2 \log n\} = O(\frac{nh_n}{\log n}) n^{-(c'/(c''+c_1))c_1^2}$ for some $c', c'' > 0$. Since, by proper choice of c_1, this can be made summable, we arrive at the conclusion via the Borel-Cantelli lemma.

(b) $$|EL_{xh}(t) - EL_{xh}(s) - L_x(t) + L_x(s)|$$
$$\leq \sum_{i=1}^n \int_{x_{i-1}}^{x_i} \frac{1}{h_n} K\left(\frac{x-z}{h_n}\right)\{|L_{x_i}(t) - L_z(t)| + |L_{x_i}(s) - L_z(s)|\}dz$$
$$+ \int_0^{x_n} \frac{1}{h_n}K\left(\frac{x-z}{h_n}\right)|[L_z(t) - L_x(t)] - [L_z(s) - L_x(s)]|dz.$$

The first term is $O(n^{-1})$ uniformly, using (C1) and (C3). The second term can be written as

$$\int_{(x-x_n)/h_n}^{x/h_n} K(u)|[L_{x-h_nu}(t) - L_x(t)] - [L_{x-h_nu}(s) - L_x(s)]|du$$
$$= \int_{(x-x_n)/h_n}^{x/h_n} K(u)\Big|(t-s)(L'_{x-h_nu}(s) - L'_x(s)) + \frac{1}{2}(t-s)^2(L''_{x-h_nu}(\theta_0) - L''_x(\theta_0))\Big|du$$

where θ_0 is an intermediate point between t and s. In view of (C6) and (C7), this is uniformly bounded as $O(a_nh_n + a_n^2)$.

COROLLARY A.1.

(a) *If we assume* $\frac{\log n}{nh_n} \to 0$, *then we can apply Lemma* A.5(a) *with* $a_n = c_0(nh_n)^{-1/2}(\log n)^{1/2}$, c_0 *some constant. The order is* $O((nh_n)^{-3/4}(\log n)^{3/4})$ *a.s.*

(b) *If we take* $a_n = c_0(nh_n)^{-1/2}(\log n)^{1/2}$ *in Lemma* A.5(b), *then the order is* $O((nh_n)^{-3/4}(\log n)^{3/4})$, *provided* $\frac{\log n}{nh_n} \to 0$ *and* $\frac{nh_n^5}{\log n} = O(1)$.

REFERENCES

Aerts, M. and Geertsema, J. (1990). Bounded length confidence intervals in nonparametric regression, *Sequential Anal.*, **9**, 171–192.

Aerts, M., Janssen, P. and Veraverbeke, N. (1994*a*). Asymptotic theory for regression quantile estimators in the heteroscedastic regression model, *Asymptotic Statistics* (eds. P. Mandl and M. Hušková), 151–161, Physica, Heidelberg.

Aerts, M., Janssen, P. and Veraverbeke, N. (1994*b*). Bootstrapping regression quantiles, *J. Nonparametric Statistics*, **4**, 1–20.

Beran, R. (1981). Nonparametric regression with randomly censored survival data, Tech. Report, University of California, Berkeley.

Breslow, N. and Crowley, J. (1974). A large sample study of the life table and product limit estimates under random censorship, *Ann. Statist.*, **2**, 437–453.

Dabrowska, D. M. (1987). Non-parametric regression with censored survival time data, *Scand. J. Statist.*, **14**, 181–197.

Dabrowska, D. M. (1989). Uniform consistency of the kernel conditional Kaplan-Meier estimate, *Ann. Statist.*, **17**, 1157–1167.

Dabrowska, D. M. (1992). Variable bandwidth conditional Kaplan-Meier estimate, *Scand. J. Statist.*, **19**, 351–361.

Efron, B. (1981). Censored data and the bootstrap, *J. Amer. Statist. Assoc.*, **76**, 312–319.

González-Manteiga, W. and Cadarso-Suarez, C. (1994). Asymptotic properties of a generalized Kaplan-Meier estimator with some applications, *J. Nonparametric Statistics*, **4**, 65–78.

González-Manteiga, W., Cao, R. and Marron, J. S. (1996). Bootstrap selection of the smoothing parameter in nonparametric hazard rate estimation, *J. Amer. Statist. Assoc.*, **91**, 1130–1140.

Härdle, W. and Mammen, E. (1991). Bootstrap methods in nonparametric regression, *Nonparametric Functional Estimation and Related Topics* (ed. G. Roussas), 111–123, Kluwer, Dordrecht.

Kaplan, E. L. and Meier, P. (1958). Nonparametric estimation from incomplete observations, *J. Amer. Statist. Assoc.*, **53**, 457–481.

Lo, S.-H. and Singh, K. (1986). The product-limit estimator and the bootstrap: some asymptotic representations, *Probab. Theory Related Fields*, **71**, 455–465.

Sánchez-Sellero, C., González-Manteiga, W. and Cao, R. (1995). Bandwidth selection in density estimation with truncated and censored data (submitted).

Serfling, R. J. (1980). *Approximation Theorems of Mathematical Statistics*, Wiley, New York.

Singh, R. S. (1975). On the Glivenko-Cantelli theorem for weighted empiricals based on independent random variables, *Ann. Probab.*, **3**, 371–374.

Stone, C. J. (1977). Consistent nonparametric regression, *Ann. Statist.*, **5**, 595–645.

Van Keilegom, I. and Veraverbeke, N. (1996). Uniform strong convergence results for the conditional Kaplan-Meier estimator and its quantiles, *Comm. Statist. Theory Methods*, **25**, 2251–2265.

Van Keilegom, I. and Veraverbeke, N. (1997). Weak convergence of the bootstrapped conditional Kaplan-Meier process and its quantile process, *Comm. Statist. Theory Methods*, **26**(4), 853–869.