T E C H N I C A L
R E P O R T

0406

CONDITIONAL RESIDUAL LIFE UNDER RANDOM CENSORSHIP

Noël VERAVERBEKE

I A P S T A T I S T I C S
N E T W O R K

INTERUNIVERSITY ATTRACTION POLE

http://www.stat.ucl.ac.be/IAP

CONDITIONAL RESIDUAL LIFE UNDER RANDOM CENSORSHIP

Noël VERAVERBEKE
Limburgs Universitair Centrum, Belgium
noel.veraverbeke@luc.ac.be

Abstract. In this paper we study the nonparametric estimation of two important functionals of the residual lifetime beyond some fixed or random timepoint: the mean and any quantile. The observations are subject to random right censoring and covariate information is taken into account. In this generality we obtain asymptotic representations and asymptotic normality results for the proposed estimators. Several results in the literature are thereby generalized to the regression context.

AMS Subject Classification: Primary 62G05; Secondary 62G20.

Key Words: asymptotic representation, Beran estimator, censoring, quantiles, residual lifetime, smoothing.

1 Introduction

The mean and median of the residual lifetime have been studied extensively in the reliability and survival analysis literature. They are defined as the mean and median of the remaining lifetime after a given time t, conditional on survival upon that time t. Several papers deal with the nonparametric estimation of these functionals, based on a random sample of possibly right censored lifetimes. For example Yang (1977, 1978), Hall and Wellner (1981), Ghorai et al (1982), Gill (1983) studied the mean residual lifetime estimator in the general right random censorship model. Ghorai and Rejtő (1987) considered the situation where the censoring scheme follows a Koziol-Green model (Koziol and Green (1976)). For the estimation of quantiles of the residual lifetime under random censorship, we mention Csörgő (1987). Also the situation where the value of t is unknown and has to be estimated from the data, is important (see Ahmad (1999), Veraverbeke (2001)).
In the present paper, several of the results in the references above will be generalized to a regression context. This situation often occurs in practice when together with the (possibly censored) lifetime, also another variable (covariate) is measured for each observation, e.g. the blood pressure in the case of a medical study. Our results allow to estimate the mean or median of the residual lifetime, conditonal on survival upon a given time t and at a given value of the covariate. It is clear that our method will involve smoothing in the covariate space: lifetimes with covariates close to the given covariate value of interest should have a large contribution in the estimation procedure.
After some preliminaries in Sections 2 and 3, we deal with quantiles in Sections 4 and 5 and with the mean in Sections 6 and 7.

2 Nonparametric regression with censored data

We will describe our regression results in the situation of fixed design points $0 \leq x_1 \leq x_2 \leq \ldots \leq x_2 \leq 1$ at which we observe independent and nonnegative responses $Y_1, \ldots, Y_n$, respresenting lifetimes. Denote $F_{x_i}(t) = P(Y_i \leq t)$ for the distribution function of the lifetime Y_i at x_i. It will often occur that the responses $Y_1, \ldots, Y_n$ are subject to random right censoring, i.e. the observed random variables at design points x_i are in fact T_i and δ_i $(i = 1, \ldots, n)$ with $T_i = \min(Y_i, C_i)$ are $\delta_i = I(Y_i \leq C_i)$, where $C_1, \ldots, C_n$ are independent and nonnegative censoring variables with distribution functions $G_{x_i}(t) = P(C_i \leq t)$. We will assume independence of Y_i and

C_i for each i.
Consequently we have that the distribution function $H_{x_i}(t) = P(T_i \le t)$ satisfies the relation $1 - H_{x_i}(t) = (1 - F_{x_i}(t))(1 - G_{x_i}(t))$. At a given design value $x \in]0,1[$, we put F_x, G_x, H_x for the distribution function of respectively the lifetime Y_x at x, the censoring variable C_x at x and $T_x = \min(Y_x, C_x)$. We will also 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 T_{x_i}, C_{x_i}, T_{x_i}, δ_{x_i}).
The conditional residual lifetime distribution is defined as $F_x(y \mid t) = P(Y_x - t \le y \mid Y_x > t)$, i.e. the distribution function of the residual lifetime, conditional on survival upon a given time t and at a given value of the covariate x. For any distribution function F, we denote by T_F the right endpoint of the support of F. Then for $0 < y < T_{F_x}$ we have that

$$F_x(y \mid t) = \frac{F_x(t+y) - F_x(t)}{1 - F_x(t)}.$$

The mean conditional residual lifetime is defined as

$$\mu_x(t) = E(Y_x - t \mid Y_x > t) = \int_t^\infty y \frac{dF_x(y)}{1 - F_x(t)} - t = \int_t^\infty \frac{1 - F_x(y)}{1 - F_x(t)} dy. \tag{1}$$

An alternative to $\mu_x(t)$ is the median conditional residual lifetime. We define, more generally, for $0 < p < 1$, the p-th quantile of $F_x(y \mid t)$:

$$Q_x(t) = F_x^{-1}(p \mid t) = \inf\{y : F_x(y \mid t) \ge p\} = -t + F_x^{-1}(p + (1-p)F_x(t)) \tag{2}$$

where, for any $0 < q < 1$, we write $F_x^{-1}(q) = \inf\{y : F_x(y) \ge q\}$ for the q-th quantile of F_x.

3 Nonparametric estimation of the conditonal distribution function and quantile fucntion

Estimation of $\mu_x(t)$ and $Q_x(t)$ on the basis of the observations (T_i, δ_i) $(i = 1, \dots, n)$ will be done by replacing F_x and F_x^{-1} in (1) and (2) by corresponding empirical versions F_{xh} and F_{xh}^{-1}, where F_{xh} is the Beran estimator (Beran (1981)). This estimator for $F_x(t)$ is a generalization of the Kaplan-Meier estimator (Kaplan and Meier (1958)) and is sometimes called the conditional Kaplan-Meier estimator. The

Beran estimator F_{xh} involves smoothing weights $w_{ni}(x;h_n)$, which we take of the Gasser-Müller type:

$$w_{ni}(x;h_n) = \int_{x_{i-1}}^{x_i} \frac{1}{h_n} K\left(\frac{x-z}{h_n}\right) dz \Big/ \int_{0}^{x_n} \frac{1}{h_n} K\left(\frac{x-z}{h_n}\right) dz \quad (i = 1, \ldots, n),$$

$x_0 = 0$, K is a known probability density function, called the kernel, and $\{h_n\}$ is a sequence of positive constants, tending to 0 as $n \to \infty$, called the bandwidth sequence. The Beran estimator is now given as

$$F_{xh}(t) = 1 - \left\{ \prod_{T_{(i)} \le t} \left(1 - \frac{w_{n(i)}(x;h_n)}{1 - \sum_{j=1}^{i-1} w_{n(j)}(x;h_n)} \right)^{\delta_{(i)}} \right\} I(t < T_{(n)}) \tag{3}$$

where $T_{(1)} \le \ldots \le T_{(n)}$ are the ordered T_i and $\delta_{(1)}, \ldots, \delta_{(n)}$ and $w_{n(1)}(x;h_n), \ldots,$ $w_{n(n)}(x;h_n)$ are the corresponding δ_i and $w_{ni}(x;h_n)$. Note that this estimator is a generalization of the Kaplan-Meier estimator if we think the weights all equal to n^{-1}. Also note that in the case of no censoring (all $T_i = Y_i$, all $\delta_i = 1$), the formula (3) reduces to $F_{xh}(t) = \sum_{i=1}^{n} w_{ni}(x;h_n) I(Y_i \le t)$, which is the nonparametric estimator for $F_x(t)$ based on an uncensored sample $Y_1, \ldots, Y_n$ in the fixed design regression case (Stone (1977)).

We will formulate the regularity conditions which will be assumed throughout. First some notations: for the design points $x_1, \ldots, x_n$ we write $\underline{\Delta}_n = \min_{1 \le i \le n} (x_i - x_{i-1})$ and $\overline{\Delta}_n = \max_{1 \le i \le n} (x_i - x_{i-1})$ and for the kernel K we write $\|K\|_\infty = \sup_{u \in I\!R} K(u)$, $\|K\|_2^2 = \int_{-\infty}^{\infty} K^2(u)du$, $\mu_1^K = \int_{-\infty}^{\infty} uK(u)du$ and $\mu_2^K = \int_{-\infty}^{\infty} u^2 K(u)du$.

On the design and on the kernel, we will assume the following conditions:

(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 $[-L_0, L_0]$ of some $L_0 > 0$, $\mu_1^K = 0$, and K is Lipschitz of order 1.

The results will also require typical smoothness conditions on the underlying distribution functions of the model. We formulate them on the distribution function $H_x(t) = P(T_x \le t)$ and on the subdistribution function $H_x^u(t) = P(T_x \le t, \delta_x = 1)$:

(C3) $H_x(t)$ and $H_x^u(t)$ have second order partial derivatives with respect to t and with respect to x, which are continuous in $(x,t) \in [0,1] \times [0,T]$, for some $T > 0$.

Partial derivatives with respect to t will be denoted as $H_x^{'}(t)$, $H_x^{''}(t)$, ... and partial derivatives with respect to x as $\dot{H}_x(t)$, $\ddot{H}_x(t)$, ...

Below we will use asymptotic representations for the Beran estimator F_{xh} and the corresponding quantile estimator F_{xh}^{-1}. The representation for F_{xh} in Lemma 1 is taken over from Theorem 2.1 in Van Keilegom and Veraverbeke (1997). The representation for $F_{xh}^{-1}(p_n)$ in Lemma 2 is formulated here for random p_n, tending to p at a certain rate. The proof of Lemma 2 is not given since it parallels that of a similar result in Gijbels and Veraverbeke (1988, Theorem 2.1).

Lemma 1. Assume (C1), (C2), (C3) in $[0,T]$ with $T < T_{H_x}$, $h_n \to 0$, $\dfrac{\log n}{nh_n} \to 0$, $\dfrac{nh_n^5}{\log n} = O(1)$. Then, for $t < T_{H_x}$,

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

where

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

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

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

Lemma 2. Assume (C1), (C2), (C3) in $[0,T]$ with $T < T_{H_x}$, $h_n \to 0$, $\dfrac{\log n}{nh_n} \to 0$, $\dfrac{nh_n^5}{\log n} = O(1)$. Assume that $F_x^{-1}(p) < T$ and that $f_x(F_x^{-1}(p)) > 0$ (where $f_x = F_x'$).

If $\{p_n\}$ is a sequence of random variables ($0 < p_n < 1$) with $p_n - p = O_P((nh_n)^{-1/2})$, then, as $n \to \infty$,

$$F_{xh}^{-1}(p_n) = F_x^{-1}(p) + \frac{1}{f_x(F_x^{-1}(p))}\left(p_n - F_{xh}(F_x^{-1}(p))\right) + o_p((nh_n)^{-1/2}).$$

4 Estimation of quantiles of the conditional residual lifetime

From (2) it follows that the obvious estimator for $Q_x(t)$ is given by

$$Q_{xh}(t) = -t + F_{xh}^{-1}(p + (1-p)F_{xh}(t)) \tag{4}$$

where F_{xh} is the Beran estimator given in (3).
Denote $q_x = p + (1-p)F_x(t)$ and $q_{xh} = p(1-p)F_{xh}(t)$. Using Lemma 2 first and then Lemma 1, we have that

$$Q_{xh}(t) - Q_x(t) = \frac{1}{f_x(F_x^{-1}(q_x))}(q_{xh} - F_{xh}(F_x^{-1}(q_x)) + o_p((nh_n)^{-1/2})$$

$$= \frac{1}{f_x(F_x^{-1}(q_x))}\{q_{xh} - q_x - [F_{xh}(F_x^{-1}(q_x)) - F_x(F_x^{-1}(q_x))]\} + o_p((nh_n)^{-1/2}$$

$$= \frac{1}{f_x(F_x^{-1}(q_x))}\sum_{i=1}^{n} w_{ni}(x;h_n\left[(1-p)\psi_x(T_i,\delta_i,t) - \psi_x(T_i,\delta_i,F_x^{-1}(q_x))\right] + o_p((nh_n)^{-1/2}).$$

The mean and variance of the main term in this representation can be calculated and the following result can be derived from this asymptotic representation (very similar as in Van Keilegom and Veraverbeke (1997)).

Theorem 1. Assume (C1), (C2), (C3) in $[0,T]$ with $T < T_{H_x}$. Assume that $F_x^{-1}(q_x) < T$ and that $f_x(F_x^{-1}(q_x)) > 0$.

(a) If $nh_n^5 \to 0$ and $(\log n)^3/(nh_n) \to 0$:

$$(nh_n)^{1/2}(Q_{xh}(t) - Q_x(t)) \xrightarrow{d} N(0;\sigma_x^2(t))$$

(b) If $h_n = Cn^{-1/5}$ for some $C > 0$:

$$(nh_n)^{1/2}(Q_{xh}(t) - Q_x(t)) \xrightarrow{d} N(\beta_x(t);\sigma_x^2(t)).$$

Here,

$$\sigma_x^2(t) = \|K\|_2^2 \frac{(1-p)^2(1-F_x(t))^2}{f_x^2(t+Q_x(t))} \int_t^{t+Q_x(t)} \frac{dH_x^u(s)}{(1-H_x(s))^2}$$

$$\beta_x(t) = (1-p)b_x(t) - b_x(t+Q_x(t))$$

with

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

This theorem generalizes the result of Csörgő (1987) to the regression case.

5 Estimation of quantiles of the duration of old age

In many studies like sociology, insurance, ... it is interesting to replace t in $Q_x(t)$ by some estimator $\widehat{t}$. The variable t is then considered as an unknown parameter, usually the starting point of 'old age'. For example, t could be defined through the proportion of retired people in the population, that is $t = F_x^{-1}(p_0)$ for some known p_0. This unknown t could then be estimated by $\widehat{t} = F_{xh}^{-1}(p_0)$.
Let $\widehat{t}$ be some general estimator for t and consider the estimator (4) with t replaced by $\widehat{t}$:

$$Q_{xh}(\widehat{t}) = -\widehat{t} + F_{xh}^{-1}(p + (1-p) F_{xh}(\widehat{t})).$$

Denote $\widehat{q}_{xh} = p + (1-p) F_{xh}(\widehat{t})$. Then, $\widehat{q}_{xh} - q_x = (1-p)(F_{xh}(\widehat{t}) - F_x(t))$, and $Q_{xh}(\widehat{t}) - Q_x(t) = -(\widehat{t} - t) + (F_{xh}^{-1}(\widehat{q}_{xh}) - F_x^{-1}(q_x))$. Now write

$$\begin{aligned} F_{xh}(\widehat{t}) - F_x(t) &= \{[F_{xh}(\widehat{t}) - F_{xh}(t)] - [F_x(\widehat{t}) - F_x(t)]\} \\ &\quad + \{F_{xh}(t) - F_x(t)\} + \{F_x(\widehat{t}) - F_x(t)\}. \end{aligned}$$

To the first term on the right hand side we can apply the modulus of continuity result in Van Keilegom and Veraverbeke (1997). To the second term we apply our Lemma 1 and to the third term we apply a first order Taylor expansion. This gives that

$$\widehat{q}_{xh} - q_x = (1-p)\{f_x(t)(\widehat{t} - t) + \sum_{i=1}^n w_{ni}(x; h_n) \psi_x(T_i, \delta_i, t)\} + o_p((nh_n)^{-1/2})$$

This, together with Lemma 2, leads to the following asymptotic representation for $Q_{xh}(\widehat{t}) - Q_x(t)$.

Theorem 2. Assume (C1), (C2), (C3) on $[0, T]$ with $T < T_{H_x}$. Assume that $t \leq T$,

$F_x^{-1}(q_x) < T$ and that $f_x(F_x^{-1}(q_x)) > 0$. Assume $h_n \to 0$, $(\log n)^3/(nh_n) \to 0$, $nh_n^5/\log n = O(1)$. Also assume that $\widehat{t} - t = O_P((nh_n)^{-1/2})$. Then, as $n \to \infty$,

$$Q_{xh}(\widehat{t}) - Q_x(t) = \left(-1 + (1-p)\frac{f_x(t)}{f_x(F_x^{-1}(q_x))}\right)(\widehat{t} - t)$$

$$+\frac{1}{f_x(F_x^{-1}(q_x))}\sum_{i=1}^{n} w_{ni}(x;h_n)[(1-p)\psi_x(T_i,\delta_i,t) - \psi_x(T_i,\delta_i,F_x^{-1}(q_x))] + o_p((nh_n)^{1/2}).$$

Example. If $t = F_x^{-1}(p_0)$ and $\widehat{t} = F_{xh}^{-1}(p_0)$, for some known p_0, we can apply Lemma 2 to $\widehat{t} - t$ and from Theorem 2 we obtain that

$$Q_{xh}(\widehat{t}) - Q_x(t) = \sum_{i=1}^{n} w_{ni}(x;h_n)\left[\frac{\psi_x(T_i,\delta_i,F_x^{-1}(p_0))}{f_x(F_x^{-1}(p_0))} - \frac{\psi_x(T_i,\delta_i,F_x^{-1}(q_x)}{f_x(F_x^{-1}(q_x))}\right] + o_p(nh_n)^{-1/2}).$$

Bias and variance of the main term can be calculated and we obtain by standard asymptotics the following limiting result.

Corollary. Let $t = F_x^{-1}(p_0)$, $\widehat{t} = F_{xh}^{-1}(p_0)$, $q = p + (1-p)p_0$. Assume (C1), (C2), (C3) in $[0,T]$ with $T < T_{H_x}$, $h_n \to 0$, $F_x^{-1}(q) < T$, $f_x(F_x^{-1}(q)) > 0$, $f_x(F_x^{-1}(p_0)) > 0$.

(a) If $nh_n^5 \to 0$ and $(\log n)^3/(nh_n) \to 0$:

$$(nh_n)^{1/2}(Q_{xh}(\widehat{t}) - Q_x(t)) \xrightarrow{d} N(0; \widetilde{\sigma}_x^2(t))$$

(b) If $h_n = Cn^{-1/5}$ for some $C > 0$:

$$(nh_n)^{1/2}(Q_{xh}(\widehat{t}) - Q_x(t)) \xrightarrow{d} N(\widetilde{\beta}_x(t), \widetilde{\sigma}_x^2(t)).$$

Here,

$$\begin{aligned}\widetilde{\sigma}_x^2(t) \quad &= \quad \|K\|_2^2(1-p_0)^2\left\{\left(\frac{1}{f_x(F_x^{-1}(p_0))} - \frac{1}{f_x(F_x^{-1}(q)}\right)^2\gamma_x(F_x^{-1}(p_0))\right.\\ &\quad \left. + \frac{(1-p)^2}{f_x^2(F_x^{-1}(q_x))}\left(\gamma_x(F_x^{-1}(q)) - \gamma_x(F_x^{-1}(p_0))\right)^2\right\}\end{aligned}$$

with

$$\gamma_x(t) = \int_0^t \frac{dH_x^u(s)}{(1-H_x(s))^2}; \tag{5}$$

$$\widetilde{\beta}_x(t) = \frac{b_x(F_x^{-1}(p_0))}{f_x(F_x^{-1}(p_0))} - \frac{b_x(F_x^{-1}(q))}{f_x(F_x^{-1}(q))}$$

with $b_x(t)$ as in Theorem 1.

6 Estimation of the mean conditional residual lifetime

If we want to estimate $\mu_x(t)$ by plugging in $F_{xh}(t)$ for $F_x(t)$ in (1), we encounter serious problems due to the right censoring and $\mu_x(t)$ cannot be estimated without additional conditions. A typical way out is to estimated truncated versions

$$\mu_x^T(t) = \int_t^T \frac{1 - F_x(s)}{1 - F_x(t)} ds \tag{6}$$

where $T < T_{F_x}$. We will only consider fixed T, but we note that it would also be possible to take $T = T_n$, some numerical sequence converging to T_{F_x} at a certain rate. This rate would depend heavily on F_x and G_x and would therefore never be known in practice.

Replacing F_x in (6) by the Beran estimator F_{xh} (see (3)) we obtain the following estimator

$$\mu_{xh}^T(t) = \int_t^T \frac{1 - F_{xh}(s)}{1 - F_{xh}(t)} ds.$$

The following asymptotic normality result generalizes that of Ghorai et al. (1980).

Theorem 3. Assume (C1), (C2), (C3) in $[0, T]$ with $T < T_{H_x}$. Assume that $t \leq T$.

(a) If $nh_n^5 \to 0$ and $(\log n)^3/(nh_n) \to 0$:

$$(nh_n)^{1/2}(\mu_{xh}^T(t) - \mu_x^T(t)) \to N(0; \tilde{\tilde{\sigma}}_x^2(t)).$$

(b) If $h_n = Cn^{-1/5}$ for some $C > 0$:

$$(nh_n)^{1/2}(\mu_{xh}^T(t) - \mu_x^T(t)) \to N(\tilde{\tilde{\beta}}_x(t); \tilde{\tilde{\sigma}}_x^2(t)).$$

Here,

$$\tilde{\tilde{\sigma}}_x^2(t) = \|K\|_2^2 \frac{1}{(1 - F_x(t))^2} \int_t^T \left(\int_u^T (1 - F_x(v)) dv \right)^2 d\gamma_x(u)$$

$$\tilde{\tilde{\beta}}_x(t) = \frac{1}{2} C^{5/2} \mu_2^K \frac{1}{1 - F_x(t)} \left\{ \int_t^T b_x(u) du - b_x(t) \int_t^T (1 - F_x(u)) du \right\}$$

with $\gamma_x(t)$ as in (5) and $b_x(t)$ as in Theorem 1.

Proof. We have that $\mu_{xh}^T(t) - \mu_x^T(t) = \int_t^T \left\{ \frac{1-F_{xh}(s)}{1-F_{xh}(t)} - \frac{1-F_x(s)}{1-F_x(t)} \right\} ds$. Write

$$\frac{1-F_{xh}(s)}{1-F_{xh}(t)} - \frac{1-F_x(s)}{1-F_x(t)} = -\frac{F_{xh}(s)-F_x(s)}{1-F_x(t)} + \frac{1-F_x(s)}{(1-F_x(t))^2}(F_{xh}(t)-F_x(t))$$

$$-\frac{(F_{xh}(t)-F_x(t))(F_{xh}(s)-F_x(s))}{(1-F_x(t))^2} + \frac{(F_{xh}(t)-F_x(t))^2(1-F_{xh}(s))}{(1-F_{xh}(t))(1-F_x(t))^2}$$

and use the asymptotic representation of Lemma 1 on the factors $F_{xh}(s) - F_x(s)$ and $F_{xh}(t) - F_x(t)$ in the first terms of the right hand side. This gives

$$\mu_{xh}^T(t) - \mu_x^T(t) =$$

$$\sum_{i=1}^n w_{ni}(x;h_n) \left\{ -\frac{1}{1-F_x(t)} \int_t^T \psi_x(T_i,\delta_i,s)ds + \frac{\int_t^T (1-F_x(s))ds}{(1-F_x(t))^2} \psi_x(T_i,\delta_i,t) \right\} \quad (7)$$

$$+R_{n1}(x,t) + R_{n2}(x,t) + R_{n3}(x,t) + R_{n4}(x,t)$$

where

$$\begin{aligned}
R_{n1}(x,t) &= -\frac{1}{1-F_x(t)} \int_t^T r_n(x,s)ds, \\
R_{n2}(x,t) &= \frac{\int_t^T (1-F_x(s))ds}{(1-F_x(t)^2} r_n(x,t), \\
R_{n3}(x,t) &= \frac{F_{xh}(t)-F_x(t)}{(1-F_x(t))^2} \int_t^T (F_{xh}(s)-F_x(s))ds, \\
R_{n4}(x,t) &= \frac{(F_{xh}(t)-F_x(t))^2}{(1-F_{xh}(t))(1-F_x(t))^2} \int_t^T (1-F_{xh}(s))ds.
\end{aligned}$$

We have that $R_{n1}(x,t)$ and $R_{n2}(x,t)$ are $O((nh_n)^{-3/4}(\log n)^{3/4})$ a.s. and that both $R_{n3}(x,t)$ and $R_{n4}(x,t)$ are $O_P((nh_n)^{-1})$. Hence the sum of the four remainder terms in (7) is $o_p((nh_n)^{-1/2})$ if $(\log n)^3/(nh_n) \to 0$. Using standard methods on the main term in the asymptotic representation (7) leads to the asymptotic normality result.

Remark. If there is no censoring, we have that $d\gamma_x(t) = dF_x(t)/(1-F_x(t))^2 = d(1/(1-F_x(t)))$ and $\widetilde{\widetilde{\sigma}}^2_x(t)$ becomes equal to

$$\|K\|_2^2 \frac{1}{(1-F_x(t))^2} \left\{ (1-F_x(T))(T-t)^2 + \int\limits_t^T (v-t)^2 dF_x(v) - \frac{\left(\int\limits_t^T (1-F_x(v))dv \right)^2}{1-F_x(t)} \right\}$$

and if we could let T tend to $+\infty$:

$$\|K\|_2^2 \left\{ \frac{\int_t^\infty (v-t)^2 dF_x(v)}{(1-F_x(t))^2} - \frac{\mu_x^2(t)}{1-F_x(t)} \right\} = \|K\|_2^2 \frac{\mathrm{Var}(Y_x - t \mid Y_x > t)}{1-F_x(t)}.$$

7 Estimation of the mean conditonal residual lifetime beyond some random timepoint

As explained in Section 5, it makes also sense to consider the estimator $\mu_{xh}^T(\widehat{t})$ where $\widehat{t}$ is some estimator for t. We only sketch how to proceed. The starting point is the following decomposition.

$$\begin{aligned} \mu_{xh}^T(\widehat{t}) - \mu_x^T(t) &= \left\{ \left[\mu_{xh}^T(\widehat{t}) - \mu_{xh}^T(t) \right] - \left[\mu_x^T(\widehat{t}) - \mu_x^T(t) \right] \right\} \\ &+ \left\{ \mu_{xh}^T(t) - \mu_x^T(t) \right\} + \left\{ \mu_x^T(\widehat{t}) - \mu_x^T(t) \right\}. \end{aligned}$$

For the second term in the right hand side we use the representation (7) in the proof of Theorem 3 and for the last term we use a one term Taylor expansion. Some easy algebra and Taylor expansion also shows that the first term equals

$$\frac{1}{1-F_x(t)} \int\limits_t^{\widehat{t}} (F_x(s) - F_{xh}(s))ds + o_p((nh_n)^{-1/2})$$

and using again the modulus of continuity result in Van Keilegom and Veraverbeke (1997) we obtain that this is $o_p((nh_n)^{-1/2})$. Therefore we can obtain the following asymptotic representation

$$\begin{aligned} \mu_{xh}^T(\widehat{t}) - \mu_x^T(t) &= \sum_{i=1}^n w_{ni}(x;h_n) \left\{ -\frac{\int_t^T \psi_x(T_i,\delta_i,s)ds}{1-F_x(t)} + \frac{\int_t^T (1-F_x(s))ds}{(1-F_x(t))^2} \psi_x(T_i,\delta_i,t) \right\} \\ &\quad + \mu_x^{T'}(t)(\widehat{t} - t) + o_p((nh_n)^{-1/2}). \end{aligned}$$

Example. If $t = F_x^{-1}(p_0)$ and $\widehat{t} = F_{xh}^{-1}(p_0)$, then we obtain that $(nh_n)^{1/2}(\mu_{xh}^T(\widehat{t}) - \mu_x^T(t))$ is asymptotically normal with asymptotic variance

$$\|K\|_2^2 \left\{ \sigma_x^2(F_x^{-1}(p_0)) + \mu_x^{T'}(t)^2 \frac{p_0(1-p_0)}{f_x^2(F_x^{-1}(p_0))} \right\}.$$

This is a generalization (and correction) of a result of Ahmad (1999) in the situation without censoring and without regression.

Acknowledgement. The research was supported by the Project BIL00/28, Bilateral and Scientific Technological Cooperation and by IAP Research Network P5/24 of the Belgian State (Federal Office for Scientific, Technical and Cultural Affairs).

References

Ahmad, I.A. (1999). Nonparametric estimation of the expected duration of old age. J. Statist. Planning Inf. **81**, 223-227.

Beran, R. (1981). Nonparametric regression with randomly censored survival data. Technical Report, Univ. California, Berkeley.

Csörgő, S. (1987). Estimating percentile residual life under random censorship. In: Contributions to Stochastics (W. Sandler, ed.) 19-27. Physica-Verlag, Heidelberg.

Ghorai, J., Susarla, A., Susarla, V. and van Ryzin, J. (1980). Non-parametric estimation of mean residual life time with censored data. In: Colloquia Mathematica Societatis Janos Bolyai, 32. Nonparametric Statistical Inference (B.V. Gnedenko et al, eds.) 269-291. North-Holland, Amsterdam.

Ghorai, J. and Rejtő, L. (1987). Estimation of mean residual life with censored data under the proportional hazards model. Commun. Statist., - Theory Meth., **16**, 2097-2114.

Gijbels, I. and Veraverbeke, N. (1988). Weak asymptotic representations for quantiles of the product-limit estimator. J. Statist. Planning Inf. **22**, 213-222.

Gill, R.D. (1983). Large sample behavior of the product-limit estimator on the whole line. Ann. Statist. **19**, 1457-1470.

Hall, W.J. and Wellner, J.A. (1981). Mean residual life. In: Statistics and Related Topics (M. Csörgő et al, eds.) 169-184. North-Holland, Amsterdam.

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

Koziol, J.A. and Green, S.B. (1976). A Cramér-von Mises statistic for randomly censored data. Biometrika **63**, 465-474.

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

Van Keilegom, I. and Veraverbeke, N. (1997). Estimation and bootstrap with censored data in fixed design nonparametric regression. Ann. Inst. Statist. Math. **49**, 467-491.

Veraverbeke, N. (2001). Estimation of the quantiles of the duration of old age. J. Statist. Planning Inf. **98**, 165-170.

Yang, G.L. (1977). Life expectancy under random censorship. Stochastic Processes Appl. **6**, 33-39.

Yang, G.L. (1978). Estimation of biometric function. Ann. Statist. **6**, 112-116.