12-2-2003

# Kernel Estimation of Rate Function for Recurrent Event Data

Chin-Tsang Chiang
*Department of Mathematics, National Taiwan University*

Mei-Cheng Wang
*Johns Hopkins Bloomberg School of Public Health*, mcwang@jhsph.edu

Chiung-Yu Huang
*Division of Biostatistics, University of Minnesota*

# Kernel Estimation of Rate Function for Recurrent Event Data

Chin-Tsang Chiang
Department of Mathematics, National Taiwan University

Mei-Cheng Wang
Department of Biostatistics, Johns Hopkins University

Chiung-Yu Huang
Division of Biostatistics, University of Minnesota

December 2, 2003

## Abstract

Recurrent event data are largely characterized by the rate function but smoothing techniques for estimating the rate function have never been rigorously developed or studied in statistical literature. This paper considers the moment and least squares methods for estimating the rate function from recurrent event data. With an independent censoring assumption on the recurrent event process, we study statistical properties of the proposed estimators and propose bootstrap procedures for the bandwidth selection and for the approximation of confidence intervals in the estimation of the occurrence rate function. It is identified that the moment method without resmoothing via a smaller bandwidth will produce curve with nicks occurring at the censoring times, whereas there is no such problem with the least squares method. Furthermore, the asymptotic variance of the least squares estimator is shown to be smaller under

1

regularity conditions. However, in the implementation of the bootstrap procedures, the moment method is computationally more efficient than the least squares method because the former approach uses condensed bootstrap data. The performance of the proposed procedures are studied through Monte Carlo simulations and an epidemiological example on intravenous drug users.

# 1   Introduction

Recurrent event data are frequently encountered in longitudinal follow-up studies when study individuals experience multiple events repeatedly over time. In this paper, we consider recurrent events of the same type, and develop methods and theory of smoothing procedures for estimating the rate function of recurrent event processes.

Recurrent event data are largely characterized by the rate function. In the regression context, semi-parametric marginal rate models were considered by Pepe and Cai (1993) and score equations were proposed for the estimation of regression parameters. Lin, Wei, Yang and Ying (2000) further provided a rigorous justification of the marginal model through the empirical process theory. Recent work on the estimation of the cumulative rate function, which is formulated in the framework of counting processes, can be tracked back to the papers of Nelson (1988) and Andersen, Borgan, Gill and Keiding (1993), among others. Under the Poisson process assumption, Bartoszyński, Brown, McBride and Thompson (1981) considered a class of smoothing methods to estimate the rate function. In their theoretical development, the censoring times are assumed to be pre-fixed constants. For recurrent event data without Poisson assumption, Lawless and Nadeau (1995) and Nelson (1995) studied non-parametric

2

procedures for estimating the cumulative rate function and developed the corresponding robust variance estimates. Although counting process methods for estimating the cumulative rate function were thoroughly studied, smoothing techniques for estimating the rate function have never been rigorously developed or studied without the Poisson assumption.

Let $N(t)$ denote the recurrent event process in a finite interval $[0, T_0]$, where $T_0$ is a positive constant. The rate function of a continuous recurrent event process at $t$, $t \in [0, T_0]$, is defined as

$$\lambda(t) = \lim_{\Delta \to 0^+} \frac{P(N(t + \Delta) - N(t) > 0)}{\Delta} \, .$$

Note that the definition of the rate function is different from the conventional intensity function where the intensity function is defined as the occurrence probability of recurrent events *conditional* on the event history up to $t$. The rate function is defined as the population average of occurrence probability of recurrent events at time point $t$ *unconditionally* on the event history. Because its marginal interpretation is useful for risk factor comparison, the rate function is preferred over the intensity function as the tool for analysis in many public health and biomedical applications.

Suppose the data are collected from $n$ independent subjects experiencing recurrent events and the observation of recurrent events from each subject could be terminated due to loss to follow-up or end of study. For the *ith* subject, let $N_i(t)$ denote the recurrent event process in $[0, T_0]$, and let $Y_i$, $0 \le Y_i \le T_0$, be the censoring time at which the observation of $N_i(t)$ is terminated. In this paper, we propose smoothing methods of the rate function under the following model assumptions:

3

(A1) Each recurrent event process $\{N_i(t)\}$ satisfies $E_i[dN_i(t)] = \lambda_i(t)dt$ with $\lambda(t) = E[\lambda_i(t)]$, where $E_i[\cdot]$ is the expectation conditioning on the $i$th subject, $\lambda_i(t)$ is the subject-specific rate function, and $\lambda(t)$ is the rate function of recurrent events in the target population.

(A2) $N_i(\cdot)$ is independent of $Y_i$.


Note that assumption (A1) is a practically unrestrictive assumption which does not place any specific distributional condition (e.g. Poisson assumption) on the recurrent event process. Assumption (A2) is essentially an independent censoring condition which assumes the censoring mechanism does not carry information about the recurrent event process. Based on assumptions (A1-2), we develop the moment and least squares methods for estimating the rate function and study the properties of the proposed estimators. Moreover, the bootstrap procedures are proposed to establish the criteria for bandwidth selection and construct the practical confidence intervals for the rate function. In this study, the pros and cons of these two estimation methods are identified and explored: First, unlike the moment method, the least squares method does not produce estimates with nicks occurring at the censoring times. Second, under some regularity conditions, the asymptotic variance of the least squares estimator is smaller than that of the moment estimator. Third, in the implementation of bootstrap procedures, the bootstrap analogue of the moment estimator can be computed via the condensed bootstrap data instead of the original raw data, and is thus faster in computation. The contents of this paper are organized as follows: Section 2 introduces the moment and least squares methods for

4

the estimation of the rate function. In Section 3, we propose bootstrap procedures for selecting bandwidths and constructing practical confidence intervals. Section 4 establishes asymptotic properties of the proposed estimators and the corresponding bootstrap analogues. The consistency properties of the estimators for bias correction are also derived in this section. Monte Carlo simulations are conducted in Section 5 to examine the performance of the proposed procedures. In Section 6, the methods are applied to data collected in an intravenous drug user study. A discussion of the estimation methods will be provided in Section 7.

## 2    Estimation Methods

Let $K_{s,l}\left(\frac{t-u}{h_t}\right) = \frac{1}{h}\alpha_l\left(s, \frac{t-u}{h_t}\right)K\left(\frac{t-u}{h_t}\right)$ be the $l$th order boundary kernel function of Gasser and Müller (1978) with adjustment for the boundary time $s$ which, for $1 < m < l$, satisfies

$$\beta_{0,l}(t, h_t, s) = 1, \beta_{m,l}(t, h_t, s) = 0 \ and \ \beta_{l,l}(t, h_t, s) < \infty,$$

where $h_t$ is a positive valued bandwidth, $K(\cdot)$ is a kernel density, and $\beta_{j,l}(t, h_t, s) = \int_{\frac{(t-s)}{h_t}}^{\frac{t}{h_t}} u^j \alpha_l(s, u)K(u)du$. In practical implementation, the second order boundary kernel function is often assigned with $\alpha_2(s, u)$ and $K(\cdot)$ being separately a linear function of $u$ and symmetric density.

In this section, two types of estimation methods for $\lambda(t)$ are proposed. The first kernel estimator, which is the improvement of the window type estimation of

5

Bartoszyński, Brown, McBride and Thompson (1981), is given by

$$\widehat{\lambda}_{h_t,2}(t) \;=\; \sum_{i=1}^{n} \frac{\delta_i\,(t)}{\delta_{\cdot}\,(t)} \left( \int_0^{Y_i} K_{Y_i,2}\left(\frac{t-u}{h_t}\right) dN_i(u) \right)$$

$$= \sum_{i=1}^{n} \frac{\xi_i(t)}{\delta_{\cdot}(t)}, \quad t \in [0, T_0], \tag{1}$$

where $\delta_i\,(t) = 1_{[Y_i \geq t]}$ is an indicator function, $\delta_{\cdot}\,(t) = \sum_{i=1}^{n}\delta_i\,(t)$, $T_0$ is the maximum value of the censoring times or the recurrent event times, and the term $\xi_i(t) = \left(\delta_i(t)\int_0^{Y_i} K_{Y_i,2}\left(\frac{t-u}{h_t}\right) dN_i(u)\right)$ in (1) is an estimator of the subject-specific rate function $\lambda_i(t)$ for $t$ in the interval $[0, Y_i]$. In their estimation, the estimator of $\lambda(t)$ may inappropriately use the information of subjects whose recurrent event processes are terminated before $t$. Under model assumptions (A1-2), the estimator $\widehat{\lambda}_{h_t,2}(t)$ uses information of subjects who are at risk at $t$, i.e., $Y_i \geq t$. This is a reasonable approach because the risk set at each $t$ forms a random sample from the population under the independent censoring assumption. Thus, the estimator $\widehat{\lambda}_{h,2}(t)$ uses the risk set as the base for kernel estimation, and is termed as the moment estimator of $\lambda(t)$.

A drawback of the moment method is that nicks could occur at the censoring times in the estimated curve $\widehat{\lambda}_{h_t,2}(t)$. The presence of nicks is mainly caused by the empirical weights $(1/\delta_{\cdot}(t))$ which operate on the subject-specific smoothers. To avoid the problem of nicks, the estimated curve can be re-smoothed via the use of a smaller bandwidth. By using the equality $E[dN_i(t)\delta_i(t)] = E[\delta_i(t)]\lambda(t)dt$ and substituting the consistent estimator $\bar{\delta}_{\cdot}(t) = \sum_{i=1}^{n}\delta_i(t)/n$ for $E[dN_i(t)]$, another smoothing estimator,

6

say, $\widetilde{\lambda}_{h_t,2}(t)$ can be obtained via minimizing the sum of squares

$$S(\lambda(t); h_t) = \sum_{i=1}^{n} \int_0^{T_0} (\frac{\delta_i(u)dN_i(u)}{\bar{\delta}.(u)du} - \lambda(t))^2 K_{T_0,2}(\frac{t-u}{h_t})du \qquad (2)$$

with respect to $\lambda(t)$. Since $\left(\int_0^{T_0} K_{T_0,2}(\frac{t-u}{h_t})du\right)$ is equal to 1, the solution of (2) leads to

$$\widetilde{\lambda}_{h_t,2}(t) = \sum_{i=1}^{n} \int_0^{T_0} K_{T_0,2}(\frac{t-u}{h_t})\frac{\delta_i(u)dN_i(u)}{\delta.(u)}. \qquad (3)$$

As we can see, the least squares estimator in (3) can also be computed by smoothing the Nelson-Aalen type estimator $\widehat{\Lambda}_n(t) = \int_0^t \sum_{i=1}^{n} \frac{\delta_i(u)dN_i(u)}{\delta.(u)}$ as below,

$$\widetilde{\lambda}_{h_t,2}(t) = \int_0^{T_0} K_{T_0,2}(\frac{t-u}{h_t})d\widehat{\Lambda}_n(u). \qquad (4)$$

Note that $\widetilde{\lambda}_{h_t,2}(t)$ is computed by smoothing the empirical subject-specific rate estimator, whereas $\widehat{\lambda}_{h_t,2}(t)$ takes the average of the subject-specific rate smoothers. Thus, the least squares method produces a smooth curve but the moment method does not. When there is no censoring on the recurrent event processes (i.e., the censoring times equal a pre-fixed constant), it is easy to see that these two estimators are the same. In the succeeding sections, we will investigate the properties of $\widehat{\lambda}_{h_t,2}(t)$ and $\widetilde{\lambda}_{h_t,2}(t)$ through the empirical process theory of recurrent event processes.

# 3   Bootstrap Methods

In this section, we will propose and study the bootstrap analogues of the two kernel estimators. Since subjects are independently selected, a natural bootstrap sampling scheme following the work of Efron (1979) is to re-sample the entire measurements

7

$\mathbf{y}_i = \{(N_i(\cdot), Y_i)\}$ of each subject with replacement from the original data. Let $\{\mathbf{y}_1^{nb}, \cdots, \mathbf{y}_n^{nb}\}$ denote the bootstrap sample of $\{\mathbf{y}_1, \cdots, \mathbf{y}_n\}$. Then, the bootstrap analogues, say, $\widehat{\lambda}_{h_t,2}^{nb}(t)$ and $\widetilde{\lambda}_{h_t,2}^{nb}(t)$ of $\widehat{\lambda}_{h_t,2}(t)$ and $\widetilde{\lambda}_{h_t,2}(t)$ can be expressed as

$$\widehat{\lambda}_{h_t,2}^{nb}(t) = \sum_{i=1}^{n} \frac{\xi_i^{nb}(t)}{\delta_\cdot^{nb}(t)} \tag{5}$$

and

$$\widetilde{\lambda}_{h_t,2}^{nb}(t) = \int_0^{T_0} K_{T_0,2}\left(\frac{t-u}{h_t}\right) d\widehat{\Lambda}_n^{nb}(u), \tag{6}$$

where $\delta_i^{nb}(t) = 1_{[Y_i^{nb} \geq t]}$, $\delta_\cdot^{nb}(t) = \sum_{i=1}^{n} \delta_i^{nb}(t)$, $\xi_i^{nb}(t) = \left(\delta_i^{nb}(t) \int_0^{Y_i^{nb}} K_{Y_i^{nb},2}\left(\frac{t-u}{h_t}\right) dN_i^{nb}(u)\right)$ and

$$\widehat{\Lambda}_n^{nb}(t) = \int_0^t \frac{\sum_{i=1}^{n} \delta_i^{nb}(u) dN_i^{nb}(u)}{\delta_\cdot^{nb}(u)}. \tag{7}$$

It can be seen that $\widehat{\lambda}_{h_t,2}^{nb}(t)$ can be directly computed from the condensed bootstrap data $\{\mathbf{u}_1^{nb}(t), \cdots, \mathbf{u}_n^{nb}(t)\}$, which is the bootstrap data of $\{\mathbf{u}_1(t), \cdots, \mathbf{u}_n(t)\}$ with $\mathbf{u}_i(t) = (\xi_i(t), \delta_i(t))$. In contrast, $\widetilde{\lambda}_{h_t,2}^{nb}(t)$ needs to be computed from the raw longitudinal bootstrap sample $\{\mathbf{y}_1^{nb}, \cdots, \mathbf{y}_n^{nb}\}$. Therefore, the bootstrap estimator $\widetilde{\lambda}_{h_t,2}^{nb}(t)$ requires more computational time and memory space than those for the bootstrap estimator $\widehat{\lambda}_{h_t,2}^{nb}(t)$. Conditioning on the recurrent event data, the bootstrap analogues $\widetilde{\lambda}_{h_t,2}^{nb}(t)$ and $\widehat{\lambda}_{h_t,2}^{nb}(t)$ both can be shown to be the approximately unbiased estimators of $\widehat{\lambda}_{h_t,2}(t)$ and $\widetilde{\lambda}_{h_t,2}(t)$. Thus, these bootstrap estimators fail to mimic the biases of the estimators because the bootstrap bias is relatively negligible. To remedy this problem, the bias correction of Schucany (1995), using the difference of the second and the fourth order kernel estimators, is extended in this case to approximate the

8

biases of $\widehat{\lambda}_{h_t,2}(t)$ and $\widetilde{\lambda}_{h_t,2}(t)$ by

$$\widehat{b}_{h_t}(t) = \widehat{\lambda}_{h_t,2}(t) - \widehat{\lambda}_{h_t,4}(t) \qquad and \qquad \widetilde{b}_{h_t}(t) = \widetilde{\lambda}_{h_t,2}(t) - \widetilde{\lambda}_{h_t,4}(t). \qquad (8)$$

The dominant bias term of each estimator at each time $t$ can be estimated via using the estimators in (8) at specified bandwidths as Schucany (1995). Under the considered data setting, the method is performed well in the numerical studies. Using the dominant bias terms, say, $\widehat{b}_{h_t}^*$ and $\widetilde{b}_{h_t}^*$ for the bias adjustment and the sample variances of the bootstrap estimators, it is reasonable to approximate the mean squared errors, say, $MSE(\widehat{\lambda}_{h_t,2}(t))$ and $MSE(\widetilde{\lambda}_{h_t,2}(t))$ of $\widehat{\lambda}_{h_t,2}(t)$ and $\widetilde{\lambda}_{h_t,2}(t)$, respectively, by

$$MSE^{nb}(\widehat{\lambda}_{h_t,2}(t)) = \widehat{b}_{h_t}^{*2}(t) + V^{nb}(\widehat{\lambda}_{h_t,2}^{nb}(t)) \qquad (9)$$

and

$$MSE^{nb}(\widetilde{\lambda}_{h_t,2}(t)) = \widetilde{b}*_{h_t}^{2}(t) + V^{nb}(\widetilde{\lambda}_{h_t,2}^{nb}(t)), \qquad (10)$$

where $V^{nb}(\cdot)$ denotes the sample variance of a bootstrap estimator. The local bandwidth estimators, say, $\widehat{h_t}$ and $\widetilde{h_t}$ at time $t$ are then defined to be the minimizers of $MSE^{nb}(\widehat{\lambda}_{h_t,2}(t))$ and $MSE^{nb}(\widetilde{\lambda}_{h_t,2}(t))$. Moreover, the global bandwidth estimators for $\widehat{\lambda}_{h_t,2}(t)$ and $\widetilde{\lambda}_{h_t,2}(t)$ can be obtained by minimizing the mean integrated squared errors

$$MISE^{nb}(\widehat{\lambda}_{h,2}) = \int_0^{T_0} MSE^{nb}(\widehat{\lambda}_{h,2}(t))\pi(t)dt \qquad (11)$$

and

$$MISE^{nb}(\widetilde{\lambda}_{h,2}) = \int_0^{T_0} MSE^{nb}(\widetilde{\lambda}_{h,2}(t))\pi(t)dt, \qquad (12)$$

9

where $h$ is a positive time-independent bandwidth and $\pi(t)$ is a non-negative weight function. Generally, the mean integrated squared errors provides a variety of global distance measures to investigate the accuracy of the considered estimator. In implementation, the weight function $\pi(t)$ is often assigned to be a uniform density function. Since the global bandwidth cannot reflect the curvature of the estimated curve, we will use the criteria (9) and (10) to select the optimal local bandwidths in the numerical study.

For the construction of confidence intervals, it is well known that the plug-in asymptotic procedure often provides poor estimates. Here, the bootstrap method is considered as a good alternative for the construction of confidence intervals. The basic idea is to use the empirical distributions of the bootstrap quantities to approximate the sampling distributions of the corresponding estimators. The validity of the proposed procedures will be verified in the next section. The bootstrap procedure for the moment method below describes the steps for constructing the approximated $(1 - \alpha)$ confidence intervals for $\lambda(t)$ with bias correction.

**1.1.** Draw the bootstrap sample $\{\mathbf{u}_1^{nb}, \cdots, \mathbf{u}_n^{nb}\}$ of size $n$ with replacement from the condensed data $\{\mathbf{u}_1, \cdots, \mathbf{u}_n\}$.

**1.2.** Compute the bootstrap estimator $\widehat{\lambda}_{h,2}^{nb}(t)$ based on the condensed bootstrap data drawn in Step 1.1.

**1.3.** Repeat Steps 1.1-1.2 B times and construct the approximated $(1 - \alpha)$ confidence intervals for $\lambda(t)$ via either the bootstrap-normal confidence interval, which is

10

based on the normal approximation,

$$\widehat{\lambda}_{h_t,2}(t) - \widehat{b}^*_{h_t}(t) \pm z_{(1-\alpha/2)} se^{nb}(\widehat{\lambda}^{nb}_{h_t,2}(t)) \tag{13}$$

or the bootstrap-percentile confidence interval, which is based on the quantile approximation,

$$(\widehat{\lambda}_{h_t,2}(t) - \widehat{b}^*_{h_t}(t) - U^{nb}_{\alpha/2}, \widehat{\lambda}_{h_t,2}(t) - \widehat{b}^*_{h_t}(t) - L^{nb}_{\alpha/2}). \tag{14}$$

Here, $z_{(1-\alpha/2)}$ is $(1-\alpha/2)$th percentile of the univariate standard normal distribution, and $se^{nb}(\cdot)$, $L^{nb}_{\alpha/2}$ and $U^{nb}_{\alpha/2}$ are respectively the standard error, the $(\alpha/2)$th and $(1-\alpha/2)$th percentiles of $B$ bootstrap estimators.

As for the least squares method, we first re-sample the entire measurements $\mathbf{y}_i$ of each subject with replacement from the original data $\{\mathbf{y}_1, \cdots, \mathbf{y}_n\}$, and then compute the bootstrap estimator $\widetilde{\lambda}^{nb}_{h_t,2}(t)$ based on the bootstrap data $\{\mathbf{y}^{nb}_1, \cdots, \mathbf{y}^{nb}_n\}$. Same with the Step 1.3, the approximated $(1-\alpha)$ confidence intervals for $\lambda(t)$ can be constructed via either

$$\widetilde{\lambda}_{h_t,2}(t) - \widetilde{b}^*_{h_t}(t) \pm z_{(1-\alpha/2)} se^{nb}(\widetilde{\lambda}^{nb}_{h_t,2}(t)) \tag{15}$$

or

$$(\widetilde{\lambda}_{h_t,2}(t) - \widetilde{b}^*_{h_t}(t) - U^{nb}_{\alpha/2}, \widetilde{\lambda}_{h_t,2}(t) - \widetilde{b}^*_{h_t}(t) - L^{nb}_{\alpha/2}). \tag{16}$$

# 4    Asymptotic Properties

We assume the following regularity conditions for the rest of this paper:

11

(A3) $Y_i$'s are independent and identically distributed with the cumulative distribution function $F_Y(y)$ and the probability measure $P_Y(y)$.

(A4) $\lambda(t)$ is four times differentiable and bounded.

(A5) $E[dN(u)dN(v)dN(w)] \leq c_0 dudvdw$ for some positive constant $c_0$ and for all $u, v, w \in [0, T_0]$.

(A6) Define $\varsigma_\nu(t, h_t, s) = \int_{\frac{t-s}{h_t}}^{\frac{t}{h_t}} |\alpha_2(s, u)K(u)|^\nu du$ with $\varsigma_3(t, h_t, s) < \infty$.

(A7) $h_t = n^{\frac{-1}{5}} h_{0t}$ for some positive bounded constant $h_{0t}$.

Let $\zeta_i(t) = \int_0^{T_0} K_{T_0,2}(\frac{t-u}{h_t}) \frac{\delta_i(u)dN_i(u)}{(1-F_Y(u))}$. By the law of large numbers, it can be shown that

$$\frac{\delta_.(t)}{n} \xrightarrow{p} (1 - F_Y(t)), \; as \; n \to \infty. \tag{17}$$

Thus, the kernel estimators $\widehat{\lambda}_{h_t,2}(t)$ and $\widetilde{\lambda}_{h_t,2}(t)$ can be expressed as

$$\widehat{\lambda}_{h_t,2}(t) = (1 - F_Y(t))^{-1} \left( \frac{1}{n} \sum_{i=1}^n \xi_i(t) \right) \left( 1 + O_p(n^{-1/2}) \right), \tag{18}$$

and

$$\widetilde{\lambda}_{h_t,2}(t) = \left( \frac{1}{n} \sum_{i=1}^n \zeta_i(t) \right) \left( 1 + O_p(n^{-1/2}) \right). \tag{19}$$

Before establishing Theorem 1, one technical lemma is stated first in the following lemma.

12

**Lemma 1.** Suppose that assumptions (A1)-(A7) and $F_Y(T_0) < 1$ are satisfied. Then,

$$E[\xi_i(t)] = (1 - F_Y(t))\left(\lambda(t) + b_{1,h_t}(t)h_t^2\right) + o(h_t^2), \tag{20}$$

$$E[\zeta_i(t)] = \lambda(t) + b_{2,h_t}(t)h_t^2 + o(h_t^2), \tag{21}$$

$$E[\xi_i^\nu(t)] = \lambda(t)\left(\int_{\{s \geq t\}} \gamma_\nu(t, h_t, s)dP_Y(s)\right)h_t^{-\nu+1} + o(h_t^{-\nu+1}), \tag{22}$$

and

$$E[\zeta_i^\nu(t)] = \left(\frac{\lambda(t)\gamma_\nu(t, h_t, T_0)}{(1 - F_Y(t))^{\nu-1}}\right)h_t^{-\nu+1} + o(h_t^{-\nu+1}), \tag{23}$$

where

$$b_{1,h_t}(t) = \frac{\lambda^{(2)}(t)\int_{\{y \geq t\}}\beta_{2,2}(t, h_t, y)dP_Y(y)}{2(1 - F_Y(t))}, \qquad b_{2,h_t}(t) = \frac{\lambda^{(2)}(t)\beta_{2,2}(t, h_t, T_0)}{2},$$

and $\gamma_\nu(t, h_t, s) = \int_{\frac{(t-s)}{h_t}}^{\frac{t}{h_t}}(\alpha_2(s, u)K(u))^\nu du$ for $\nu = 2, 3$.

PROOF. See Appendix.□

Let $\Phi(\cdot)$ denote the cumulative distribution function of the univariate standard normal distribution. The asymptotic normalities of the estimators are established in the following theorem.

**Theorem 1.** Suppose that assumptions (A1) through (A7), $\lambda(t) > 0$, and $F_Y(T_0) < 1$ are satisfied. Then, for all $z \in \mathbf{R}$ and $n \to \infty$,

13

$$\sup_z |P\left( \frac{\sqrt{nh_t}(\widehat{\lambda}_{h_t,2}(t) - \lambda(t) - b_{1,h_t}(t))}{\sigma_{1,h_t}(t)} \le z \right) - \Phi(z)| \to 0 \qquad (24)$$

and

$$\sup_z |P\left( \frac{\sqrt{nh_t}(\widetilde{\lambda}_{h_t,2}(t) - \lambda(t) - b_{2,h_t}(t))}{\sigma_{2,h_t}(t)} \le z \right) - \Phi(z)| \to 0, \qquad (25)$$

where

$$\sigma_{1,h_t}^2(t) = \frac{\lambda(t)(\int_{\{y \ge t\}} \varsigma_2(t, h_t, y)dP_Y(y))}{(1 - F_Y(t))^2}, \qquad \sigma_{2,h_t}^2(t) = \frac{\lambda(t)\varsigma_2(t, h_t, T_0)}{(1 - F_Y(t))}.$$

PROOF. See Appendix.□

When the estimators are computed based on a linear function and a symmetric density separately for $\alpha_2(\cdot)$ and $K(\cdot)$, it can be derived algebraically that $\sigma_{1,h_t}^2(t) \ge \sigma_{2,h_t}^2(t)$ as $h$ converges to 0. Also, if the distribution of the censoring time is continuous and the kernel function is assigned without boundary adjustment, paralleling the proof of Theorem 1, these two estimators can be shown to have the same asymptotic bias and variance. As for the assumptions on the recurrent event process, the variances of $\widehat{\lambda}_{h_t,2}(t)$ and $\widetilde{\lambda}_{h_t,2}(t)$ under the general empirical process assumptions may not be equal to those under the Poisson-type assumption. However, each estimator has the same asymptotic variance with or without the Poisson assumption. This phenomenon can be explained by the fact that the kernel smoothers are locally smooth and the event history information carried in the kernel smoothers is ignorable; thus, the asymptotic variances of the estimators are mainly dominated by the rate function $\lambda(t)$ regardless of the correlation structure on the recurrent event process.

14

In the next lemma, we state the consistency of the estimators $\widehat{b}_{h_t}(t)$ and $\widetilde{b}_{h_t}(t)$. The properties can be derived directly by using Lemma 1 and the similar proof procedure of Theorem 1.

**Lemma 2.** Suppose that assumptions (A1)-(A4) and (A7), and $F_Y(T_0) < 1$. Then, as $n \to \infty$,

$$\sqrt{nh_t}\left(\widehat{b}_{h_t}(t) - b_{1,h_t}(t)\right) \overset{p}{\to} 0 \qquad and \qquad \sqrt{nh_t}\left(\widetilde{b}_{h_t}(t) - b_{2,h_t}(t)\right) \overset{p}{\to} 0. \qquad (26)$$

Before deriving the asymptotic normalities of the bootstrap estimators, let $P^{nb}(\cdot)$ denote the probability measure conditioning on the sample $\{\mathbf{y}_1, \cdots, \mathbf{y}_n\}$. The next theorem and Lemma 2 show the validity of the approximated bootstrap confidence intervals in (13)-(16), i.e., the sampling distributions of $(\widehat{\lambda}_{h_b,2}^{nb}(t) - \widehat{\lambda}_{h_t,2}(t))$ and $(\widetilde{\lambda}_{h_t,2}^{nb}(t) - \widetilde{\lambda}_{h_t,2}(t))$ can be used to approximate the distributions of $(\widehat{\lambda}_{h_t,2}(t) - \lambda(t) - \widehat{b}_{h_t}^*(t))$ and $(\widetilde{\lambda}_{2,h_t}(t) - \lambda(t) - \widetilde{b}_{h_t}^*(t))$.

**Theorem 2.** Suppose that assumptions (A1) through (A7), $\lambda(t) > 0$, and $F_Y(T_0) < 1$ are satisfied. Then, for all $z \in \mathbf{R}$ and $n \to \infty$,

$$\sup_z |P^{nb}\left(\frac{\sqrt{nh_t}(\widehat{\lambda}_{h_t,2}^{nb}(t) - \widehat{\lambda}_{h_t,2}(t))}{\sigma_{1,h_t}(t)} \leq z\right) - \Phi(z)| \overset{p}{\to} 0 \qquad (27)$$

and

$$\sup_z |P^{nb}\left(\frac{\sqrt{nh_t}(\widetilde{\lambda}_{h_t,2}^{nb}(t) - \widetilde{\lambda}_{h_t,2}(t))}{\sigma_{2,h_t}(t)} \leq z\right) - \Phi(z)| \overset{p}{\to} 0. \qquad (28)$$

15

PROOF. See Appendix.□

# 5  Monte Carlo Simulations

In this numerical study, we simulate data from subject-specific non-stationary Poisson processes. Note that, due to the independent increment property of Poisson processes, the intensity function of a Poisson process is also the rate function. Assume that, given the value of the random variable $Z_i$, $N_i(t)$ is a non-stationary Poisson process which has the following subject-specific rate (or intensity) functions:

Case 1. $\lambda_i(t) = Z_i + \phi_{10}(t)$ with $\phi_{10}(t) = 0.5\left(\sin(\frac{t\pi}{6}) - 1\right)$.

Case 2. $\lambda_i(t) = Z_i\phi_{20}(t)$ with $\phi_{20}(t) = 0.5\left(\sin(\frac{t\pi}{6}) + 1\right)$.

In the above additive and multiplicative models, the random variable $Z_i$ can be considered as the random effect or frailty, and is designed to be distributed as the uniform distribution $U(0.9, 1.1)$. Thus, both of the rate functions are equal to

$$\lambda(t) = 0.5\left(\sin(\frac{t\pi}{6}) + 1\right).$$

Under the independent censoring assumption (A2), the simulated data are generated from 450 independent non-stationary Poisson processes $\{N_i(t)\}$ with recurrent event times ranging from 0 to 4. Moreover, the censoring times are set to be distributed as

$$f_Y(y) = \frac{2\exp(2y)}{(\exp(8) - \exp(2))}, \quad y \in [1, 4].$$

Here, the assumed conditions are similar in nature to those found in the Intravenous drug user study which will be studied in the next section. The simulation

16

Table 1: The empirical coverage probabilities of the 95% bootstrap confidence intervals for $\lambda(t)$, based on the moment estimation method, at seven time points.

| Time point | 0.5 | 1.0 | 1.5 | 2.0 | 2.5 | 3.0 | 3.5 |
|---|---|---|---|---|---|---|---|
| (Case 1) Cov. Prob. | 0.950 | 0.948 | 0.952 | 0.934 | 0.950 | 0.918 | 0.960 |
| (Case 2) Cov. Prob. | 0.954 | 0.964 | 0.946 | 0.938 | 0.928 | 0.918 | 0.934 |

Table 2: The empirical coverage probabilities of the 95% bootstrap confidence intervals for $\lambda(t)$, based on the least squares estimation method, at seven time points.

| Time point | 0.5 | 1.0 | 1.5 | 2.0 | 2.5 | 3.0 | 3.5 |
|---|---|---|---|---|---|---|---|
| (Case 1) Cov. Prob. | 0.922 | 0.926 | 0.948 | 0.942 | 0.926 | 0.952 | 0.926 |
| (Case 2) Cov. Prob. | 0.933 | 0.913 | 0.940 | 0.923 | 0.943 | 0.923 | 0.900 |

process is repeated 500 times. For each simulated data set, the kernel estimators $\widehat{\lambda}_{h_t,2}(t)$ and $\widetilde{\lambda}_{h_t,2}(t)$ are computed by (1) and (4) with the normal density for $K(\cdot)$. Moreover, the local optimal bandwidths are selected via the bootstrap mean squared errors in (9) and (10). The 500 simulation averages of both estimated curves are very close to the true rate function. However, the variation of $\widetilde{\lambda}_{h_t,2}(t)$ is smaller than that of $\widehat{\lambda}_{h_t,2}(t)$ although the difference is not apparent. To evaluate the validity of the estimators, the approximated 95% bootstrap confidence intervals are constructed based on 200 bootstrap replications. For each estimation method, we only present the bootstrap-normal procedure since the bootstrap-percentile procedure has a similar conclusion. Table 1 and Table 2 summarize the empirical coverage probabilities of 95% bootstrap confidence intervals for $\lambda(t)$, based on two estimation methods, at selected time points. It appears that the coverage probabilities are generally close to the nominal level.

17

# 6 A Data Example

The data set used here involves 450 HIV-negative intravenous drug users, who entered the study before August 1, 1993, from the AIDS Link to Intravenous Experiences cohort study. This study was initiated in 1988 and started to systematically collect health service data in July, 1993. The repeated hospitalizations for each drug user here were observed between August 1, 1993 and December 31, 1997. Details of this study can be found in Vlahov, et al. (1991).

Let $y_i$ be the time length from August 1, 1993 to the date of the last visit for the *ith* drug user, and $T_0$ the maximum time of $y_i$'s. Among these patients, the median of the number of recurrent events is 1 and the number ranges from 0 to 19. The mean of the censoring time is 3.734 years and the censoring time ranges from 0.275 to 4.394 years. The main objectives of our analysis are to estimate the rate function of hospitalizations over time for HIV-negative drug users and to evaluate the accuracy of the estimated curve.

In this study, two estimators are computed based on the Gaussian kernel density for $K(\cdot)$. For the bandwidth selection, the bootstrap mean squared errors in (9) and (10) are used separately to select local optimal bandwidths for $\widehat{\lambda}_{h_t,2}(t)$ and $\widetilde{\lambda}_{h_t,2}(t)$. Moreover, the corresponding $\pm 1.96$ estimated standard error bars of the hospitalization rate estimator at the selected time points are provided. It appears in Figure 1 that both of the estimated curves $\widehat{\lambda}_{h_t,2}(t)$ and $\widetilde{\lambda}_{h_t,2}(t)$ imply the same biological explanation except wider confidence intervals for the moment method. From the figure,

18

we can see that the hospitalization rates are lower than 1 and the peak of the curve occurs roughly at the later half of the study. This reflects the effect of detoxification on the frequencies of hospitalization for intravenous drug users.

# 7    Discussion

In this study, two smoothing methods are proposed for the estimation of the rate function under the independent censoring model without the Poisson assumption on the recurrent event process. We point out that the moment estimator has the disadvantage of having nicks in the estimates. When a linear function and a symmetric density are designed for the second order kernel smoother, the asymptotic variance of the moment estimator is shown to be larger than that of the least squares estimator, although the differences are expected to be small. Regardless of the mentioned disadvantages, the moment estimation possesses important computational advantages as we discussed in section 3. In general, when sample size $(n)$ is appropriate, the drawbacks of the moment estimator appear insignificant and the computational advantage becomes an attractive feature for adopting the moment estimation approach.

## REFERENCES

Andersen, P. K., Borgan, O., Gill, R. D. and Keiding N. (1993). *Statistical Models Based on Counting Processes.* Springer-Verlag New York, Inc. 76.

Bartoszyński, R., Brown, B. W., McBride, C. M. and Thompson, J. R. (1981). Some

19

Nonparametric Techniques for Estimating the Intensity Function of a Cancer Related Nonstationary Poisson Process. *The Annals of Statistics.* **9**, 1050-1060.

Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *Ann. Statist.* **7**, 1-26.

Gasser, Th. and Müller, H.-G. (1978). *Kernel estimation of regression functions. Smoothing Techniques for Curve Estimation* (Th. Gasser and M. Rosenblatt, eds.), Spring Lecture Notes in Mathematics No.757, Springer-Verlag, Berlin, 23-68.

Lawless, J. F. and Nadeau, C. (1995). Some simple robust method for the analysis of recurrent events. *Technometrics.* **37**, 158-168. Lin and T. R. Fleming, Springer-Verlag New York, Inc. 37-49.

Lin, D. Y., Wei, L. J., Yang, I. and Ying, Z. (2000). Semiparametric regression for the mean and rate functions of recurrent events. *Journal of the Royal Statistical Society.* **B62**, 711-730.

Nelson, W. B. (1988). Graphical Analysis of System Repair Data. *Journal of Quality Technology.* **20**, 24-35.

Nelson, W. B. (1995). Confidence Limits for Recurrence Data-Applied to Cost or Number of Product Repairs. *Technometrics.* **37**, 147-157.

Pepe, M. S. and Cai, J. (1993). Some graphical displays and marginal regression analyses for recurrent failure times and time dependent covariates. *Journal of the American Statistical Association.* **88**, 811-820.

20

Schucany, W. R. (1995). Adaptive bandwidth choice for kernel regression. *Journal of the American Statistical Association.* **90**, 535-540.

Sun, J and Wei, L. J. (2000). Regression analysis of panel count data with covariate-dependent observation and censoring times. *Journal of the Royal Statistical Society.* **B62**, 293-302.

Vlahov, D., Anthony, J. C., Muñov, A., Margolick J., Nelson K. E., Celentano D. D., Solomon, L. and Polk, B. F. (1991). The ALIVE Study: A Longitudinal Study of HIV-1 Infection in Intravenous Drug Users: Description of Methods. *The Journal of Drug Issues.* **21**, 759-776.

# APPENDIX

**Proof of Lemma 1**

In the following proof, we only derive the statements of (20) and (22). Same arguments can be used in the derivation of (21) and (23). By assumptions (A1)-(A4) and the Taylor expansion, we can get

$$
\begin{aligned}
E[\xi_i(t)] &= E[\delta_i(t) \int_0^{Y_i} K_{Y_i,2}(\frac{t-u}{h_t})dN_i(u)] \\
&= \int_{\{s \geq t\}} \left( \int_0^s K_{s,2}(\frac{t-u}{h_t})E[dN_i(u)] \right) dP_Y(s) \\
&= \int_{\{s \geq t\}} \left( \int_0^s K_{s,2}(\frac{t-u}{h_t})\lambda(u)du \right) dP_Y(s)
\end{aligned}
$$

21

$$
= \int_{\{s \geq t\}} \left( \int_{\frac{t-s}{h_t}}^{\frac{t}{h_t}} \alpha_2(s,u) K(u) \left( \sum_{k=0}^{2} \frac{\lambda^{(k)}(t)(-uh_t)^k}{k!} + o((uh)^2) \right) du \right) dP_Y(s)
$$

$$
= (1 - F_Y(t)) \lambda(t) + \left( \frac{\lambda^{(2)}(t) \int_{\{s \geq t\}} \beta_{2,2}(t,h_t,s) dP_T(s)}{2} \right) h_t^2 + o(h_t^2). \quad (29)
$$

For $E[\xi_i^2(t)]$, it can be decomposed into two components

$$
E[\xi_i^2(t)] = I_A + II_A, \quad (30)
$$

where

$$
I_A = E[\delta_i(t) \int_0^{Y_i} K_{Y_i,2}^2(\frac{t-u}{h_t}) dN_i(u)]
$$

and

$$
II_A = E[\delta_i(t) \int_{\{u \neq v\}} K_{Y_i,2}(\frac{t-u}{h_t}) K_{Y_i,2}(\frac{t-v}{h_t}) dN_i(u) dN_i(v)].
$$

By assumptions (A1)-(A4), we can show that

$$
I_A = E[\delta_i(t) \int_0^{Y_i} (K_{Y_i,2}(\frac{t-u}{h_t}))^2 dN_i(u)]
$$

$$
= \int_{\{s \geq t\}} \left( \int_0^s (K_{s,2}(\frac{t-u}{h_t}))^2 E[dN_i(u)] \right) dP_Y(s)
$$

$$
= \int_{\{s \geq t\}} \left( \int_0^s (K_{s,2}(\frac{t-u}{h_t}))^2 \lambda(u) du \right) dP_Y(s)
$$

$$
= \lambda(t) \left( \int_{\{s \geq t\}} \gamma_2(t,h_t,s) dP_Y(s) \right) h_t^{-1} + o(h_t^{-1}). \quad (31)
$$

Assumption (A5) implies that there exists a positive constant $c_1$ such that

$$
II_A = \int_{\{s \geq t\}} \left( \int \int_{\{u \neq v\}} K_{s,2}(\frac{t-u}{h_t}) K_{s,2}(\frac{t-v}{h_t}) E[dN_i(u) dN_i(v)] \right) dP_Y(s)
$$

22

$$\leq c_1 \int_{\{s \geq t\}} \left( \int \int_{\{u \neq v\}} K_{s,2}(\frac{t-u}{h_t}) K_{s,2}(\frac{t-v}{h_t}) du dv \right) dP_Y(s)$$

$$= c_1 \left(1 - F_T(t)\right) \left(1 + o(1)\right). \tag{32}$$

Substituting (31) and (32) into (30), $E[\xi_i^2(t)]$ is obtained. Along the same lines as the derivation of $E[\xi_i^2(t)]$, the statement of (22) for $\nu = 3$ follows.

**Proof of Theorem 1**

From (18), we make the linear approximation

$$\frac{\widehat{\lambda}_{h_t,2}(t) - E[\widehat{\lambda}_{h_t,2}(t)]}{\sqrt{V(\widehat{\lambda}_{h_t,2}(t))}} = \frac{\sum_{i=1}^n (\xi_i(t) - E[\xi_i(t)])}{\sqrt{nV(\xi_i(t))}} \left(1 + o_p(1)\right). \tag{33}$$

By the Berry-Esséen theorem, we get the following inequality.

$$\sup_z \left| P\left( \frac{\sum_{i=1}^n (\xi_i(t) - E[\xi_i(t)])}{\sqrt{nV(\xi_i(t))}} \leq z \right) - \Phi(z) \right| \leq d \frac{(nE[|\xi_i(t) - E[\xi_i(t)]|^3])}{(nV(\xi_i(t))^{3/2})}, \tag{34}$$

where $d$ is a positive constant independent of $n$ and $\xi_i(t)$. Since $|\xi_i(t)| \leq \eta_i(t)$ and $N_i(t)$'s are positive random variables, it implies that

$$E[|\xi_i(t) - E[\xi_i(t)]|^3] \leq E[(\eta_i(t))^3] + 6[(\eta_i(t))^2] E[\eta_i(t)] + (E[\eta_i(t)])^3, \tag{35}$$

where $\eta_i(t) = \delta_i(t) \int_0^{t_i} |K_{t_i,2}(\frac{t-u}{h_t})| dN_i(u)$. Similar to the derivation of Lemma 1, we can get that, for $\nu = 1, 2, 3$,

$$E[(\eta_i(t))^\nu] = \lambda(t) \left( \int_{\{s \geq t\}} \varsigma_\nu(t, h, s) dP_T(s) \right) h^{-\nu+1} + o(h^{\nu+1}). \tag{36}$$

From (35), (36) and Lemma 1, it implies that

$$\frac{nE[|\xi_i(t) - E[\xi_i(t)]|^3]}{(nV(\xi_i(t))^{\frac{3}{2}}} \leq (nh_t)^{\frac{-1}{2}} \left( \frac{\int_{\{s \geq t\}} \varsigma_3(t, h_t, s) dP_Y(s)}{\int_{\{s \geq t\}} \gamma_2(t, h_t, s) dP_Y(s)} \right) \rightarrow 0, \tag{37}$$

23

By assumption (A7) and substituting (37) into (34), we derive that

$$\sup_z |P\left(\frac{\sum_{i=1}^n (\xi_i(t) - E[\xi_i(t)])}{\sqrt{nV(\xi_i(t))}} \le z\right) - \Phi(z)| \to 0, \ as \ n \to \infty. \tag{38}$$

Finally, the proof of (24) is completed by (33) together with (38),

$$E[\widehat{\lambda}_{h_t,2}(t)] = \lambda(t) + b_{1,h}(t)h_t^2 + o(h_t^2) \tag{39}$$

and

$$V(\widehat{\lambda}_{h_t,2}(t)) = \sigma_{1,h_t}^2(t)(nh_t)^{-1} + o\left((nh_t)^{-1}\right). \tag{40}$$

For the proof of (25), we first make the linear approximation

$$\frac{\widetilde{\lambda}_{h_t,2}(t) - E[\widetilde{\lambda}_{h_t,2}(t)]}{\sqrt{V(\widetilde{\lambda}_{h_t,2}(t))}} = \frac{\sum_{i=1}^n (\zeta_i(t) - E[\zeta_i(t)])}{\sqrt{nV(\zeta_i(t))}} \left(1 + o_p(1)\right). \tag{41}$$

Then, paralleling the steps from (34) to (38) with (41),

$$E[\widetilde{\lambda}_{h_t,2}(t)] = \lambda(t) + b_{2,h_t}(t)h_t^2 + o(h_t^2) \tag{42}$$

and

$$V(\widetilde{\lambda}_{h_t,2}(t)) = \sigma_{2,h}^2(t)(nh_t)^{-1} + o\left((nh_t)^{-1}\right), \tag{43}$$

the asymptotic normality of $\widetilde{\lambda}_{h_t,2}(t)$ is obtained.

**Proof of Theorem 2**

Since the derivation of the proof for the bootstrap estimator $\widetilde{\lambda}_{h_t,2}^{nb}(t)$ are similar to those of $\widehat{\lambda}_{h_t,2}^{nb}(t)$, for the space of presentation, we only derive the asymptotic

24

normalities of $\widehat{\lambda}_{h_t,2}^{nb}(t)$. Let $\xi_i^{nb}(t) = \delta_i^{nb}(t) \int_0^{Y_i^{nb}} K_{Y_i^{nb},2}(\frac{t-u}{h_t}) dN_i^{nb}(u)$. By the law of large numbers, we can show that, for any $\varepsilon > 0$,

$$P^{nb}\left(|\frac{\delta_.^{nb}(t)}{n} - \frac{\delta_.(t)}{n}| > \varepsilon\right) \to 0. \tag{44}$$

Then, the bootstrap kernel $\widehat{\lambda}_{h_t,2}^{nb}(t)$ can be expressed as

$$\widehat{\lambda}_{h_t,2}^{nb}(t) = \left(\frac{\delta_.^{nb}(t)}{n}\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^n \xi_i^{nb}(t)\right)$$

$$= \left(\frac{\delta_.(t)}{n}\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^n \xi_i^{nb}(t)\right)(1 + o_{p^{nb}}(1)). \tag{45}$$

Similar to the proof of Theorem 1, we can derive that

$$\frac{\widehat{\lambda}_{h_t,2}^{nb}(t) - E^{nb}[\widehat{\lambda}_{h_t,2}(t)]}{\sqrt{V^{nb}(\widehat{\lambda}_{h_t,2}(t))}} = \frac{\sum_{i=1}^n (\xi_i^{nb}(t) - E^{nb}[\xi_i(t)])}{\sqrt{nV^{nb}(\xi_i^{nb}(t))}}\left(1 + o_{p^{nb}}(1)\right). \tag{46}$$

Also, by the Berry-Esséen theorem, there exists a positive constant $c_2$ such that

$$\sup_z |P^{nb}(\frac{\sum_{i=1}^n (\xi_i^{nb}(t) - E^{nb}[\xi_i^{nb}(t)])}{\sqrt{nV^{nb}(\xi_i^{nb}(t))}} \leq z) - \Phi(z)| \leq c_2 \frac{nE^{nb}[|\xi_i^{nb}(t) - E[\xi_i^{nb}(t)]|^3]}{(nV^{nb}(\xi_i^{nb}(t))^{\frac{3}{2}}}. \tag{47}$$

Since

$$E^{nb}[(\xi^{nb}(t))^\nu] = \frac{\sum_{i=1}^n (\xi(t))^\nu}{n}, \; for \; \nu = 1, 2, 3, \tag{48}$$

, by the law of large numbers and Lemma 1, it implies

$$hV^{nb}(\xi_i^{nb}(t)) = \left(\frac{1}{n}\sum_{i=1}^n (\xi_i(t))^2 - (\frac{1}{n}\sum_{i=1}^n \xi_i(t))^2\right) \xrightarrow{p} \sigma_{1,h_t}^2(t) \tag{49}$$

and

$$h_t^2 E^{nb}[|\xi_i^{nb}(t) - E[\xi_i^{nb}(t)]|^3] \leq h^2\left(\frac{1}{n}\sum_{i=1}^n (\eta_i(t))^3 + 6\bar{\eta}(t)(\frac{1}{n}\sum_{i=1}^n (\eta_i(t))^2) + (\bar{\xi}(t))^3\right)$$

$$\xrightarrow{p} \lambda(t)\left(\int_{\{s \geq t\}} \varsigma_3(t, h_t, s) dP_Y(s)\right). \tag{50}$$

25

Substituting (49) and (50) into (48), we get

$$\sup_z |P^{nb}(\frac{\sum_{i=1}^n (\xi_i^{nb}(t) - \frac{\sum_{i=1}^n \xi_i(t)}{n})}{\sqrt{n h_t^{-1} \sigma_{1,h_t}^2(t)}} \le z) - \Phi(z)| \xrightarrow{p} 0. \tag{51}$$
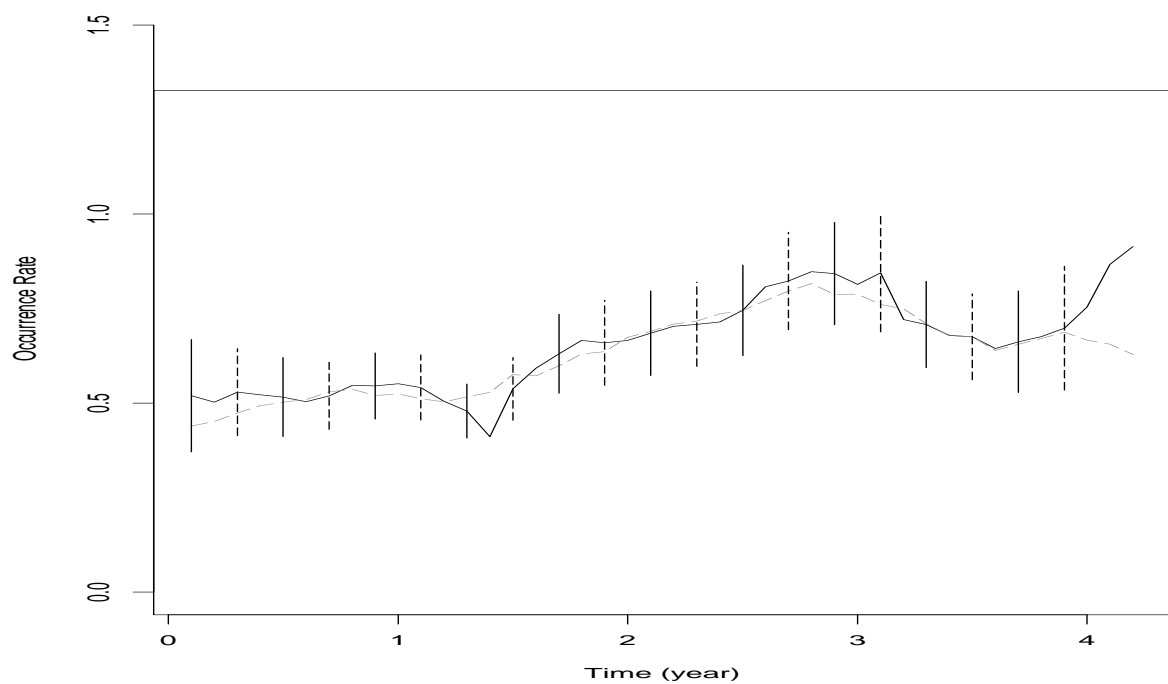
Finally, (27) is obtained from (46) and (51).

26

Figure 1: The solid and dashed curves represent separately the estimated rate functions $\widehat{\lambda}_{h,2}(t)$ and $\widetilde{\lambda}_{h,2}(t)$ of HIV-negative intravenous drug users with the corresponding 95% bootstrap intervals.

27