



Published in final edited form as:

Scand Stat Theory Appl. 2005 March ; 32(1): 77–91. doi:10.1111/j.1467-9469.2005.00416.x.

Kernel Estimation of Rate Function for Recurrent Event Data

CHIN-TSANG CHIANG,

Department of Mathematics, National Taiwan University

MEI-CHENG WANG, and

Department of Biostatistics, Johns Hopkins University

CHIUNG-YU HUANG

Division of Biostatistics, University of Minnesota

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 a 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 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 is studied through Monte Carlo simulations and an epidemiological example on intravenous drug users.

Keywords

bootstrap; independent censoring; intensity function; kernel estimator; Poisson process; rate function; recurrent events

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 & Cai (1993) and score equations were proposed for the estimation of regression parameters. Lin *et al.* (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 paper of Nelson (1988).

Under the Poisson process assumption, Bartoszyński *et al.* (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 the Poisson assumption, Lawless & Nadeau (1995) and Nelson (1995) studied non-parametric procedures for estimating the cumulative rate function and developed the corresponding robust variance estimates. Although counting process methods for estimating the cumulative rate function have been 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

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

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 to 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 of follow-up or end of the study. For the i th subject, let $N_i(t)$ denote the recurrent event process in $[0, T_0]$, and let $Y_i, 0 \leq Y_i \leq T_0$, 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 assumption:

(A1) Each recurrent event process $\{N_i(t)\}$ satisfies $E[dN_i(t) | \lambda(t)] = \lambda(t) I_{[Y_i, t]}$ dt , where $\lambda(t) = I_{[Y_i, t]}$ is the at risk indicator.

Note that the above model is also a model of Lin *et al.* (2000), and Scheike (2002) without considering covariates. As one can see, assumption (A1) is a practically unrestrictive assumption which does not place any specific distributional condition (e.g. Poisson assumption) on the recurrent event process. Based on the above model, we develop the moment and least squares methods for estimating the rate function and study the properties of the proposed estimators. In addition, bootstrap procedures are proposed to establish the criteria for bandwidth selection and to 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 in the curve at the censoring times. Secondly, under some regularity conditions, the asymptotic variance of the least squares estimator is smaller than that of the moment estimator. Thirdly, 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 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

In this section, two types of smoothing estimation methods for $\lambda(t)$ are proposed. To avoid edge effects in the boundary region, the k th order boundary kernel weight function of Gasser & Müller (1978) with adjustment for the censoring time is used in our estimation procedures. The reason for doing so is mainly to reduce and estimate the bias terms in the proposed estimators. In the next section, we will use the difference of the second order and the fourth order kernel estimators to estimate the dominant biases of the proposed second order kernel estimators.

The first kernel estimator, which is the improvement of the window type estimation of Bartoszyński *et al.* (1981), is given by

$$\widehat{\lambda}_{h_t,2}(t) = \sum_{i=1}^n \frac{\delta_i(t)}{\delta_{\cdot}(t)} \left(\int_0^{Y_i} K_{r_i,2} \left(\frac{t-u}{h_t} \right) dN_i(u) \right) = \sum_{i=1}^n \frac{\xi_i(t)}{\delta_{\cdot}(t)}, \quad t \in [0, T_0], \quad (1)$$

where $K_{s,k}(\cdot)$ is the k th order boundary kernel function with adjustment for s , h_t is a positive valued bandwidth, $\delta_{\cdot}(t) = \sum_{i=1}^n \delta_i(t)$, T_0 is the maximum value of the censoring times or the

recurrent event times, and the term $\lambda_i(t) = \left(\int_0^{Y_i} K_{Y_i,2}((t-u)/h_t) dN_i(u) \right)$ in (1) is an estimator of the subject-specific rate function $\lambda_i(t)$ for t in the interval $[0, Y_i]$. Similar to estimators considered by Ramlau-Hansen (1983), Tanner & Wong (1983) and Yandell (1983) for intensity or hazard function, another estimator is obtained by smoothing the

Nelson–Aalen type estimator $\widehat{\Lambda}_n(t) = \int_0^t \sum_{i=1}^n \delta_i(u) / \delta_{\cdot}(u) dN_i(u)$ as below,

$$\widetilde{\lambda}_{h_t,2}(t) = \int_0^{T_0} K_{r_0,2} \left(\frac{t-u}{h_t} \right) d\widehat{\Lambda}_n(u) = \sum_{i=1}^n \int_0^{T_0} K_{r_0,2} \left(\frac{t-u}{h_t} \right) \frac{\delta_i(u) dN_i(u)}{\delta_{\cdot}(u)}. \quad (2)$$

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

$$S(\lambda(t); h_t) = \sum_{i=1}^n \int_0^{T_0} \left(\frac{\delta_i(u) dN_i(u)}{\delta_{\cdot}(u)} - \lambda(t) \right)^2 K_{r_0,2} \left(\frac{t-u}{h_t} \right) du$$

with respect to $\lambda(t)$, it herein is called the least squares estimator of $\lambda(t)$.

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. The least squares method produces a smooth curve but the moment estimator has some nicks occurring at the censoring times. To avoid the problem of nicks, the estimated curve can be re-smoothed by using a smaller bandwidth. 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. As subjects are independently selected, a natural bootstrap sampling scheme following the work of Efron (1979) is to re-sample the entire measurements $\mathbf{y}_i = \{(N_i(\cdot), Y_i)\}$ of each subject with replacements from the original data. Let $\{\mathbf{y}_1^{nb}, \dots, \mathbf{y}_n^{nb}\}$ denote the bootstrap sample of $\{\mathbf{y}_1, \dots, \mathbf{y}_n\}$. Then, the bootstrap analogues, say, $\widehat{\lambda}_{h_i,2}^{nb}(t)$ and $\widetilde{\lambda}_{h_i,2}^{nb}(t)$ of $\widehat{\lambda}_{h_i,2}(t)$ and $\widetilde{\lambda}_{h_i,2}(t)$ are computed based on the bootstrap sample.

It can be seen that $\widehat{\lambda}_{h_i,2}^{nb}(t)$ can be directly computed from the condensed bootstrap data $\{\mathbf{u}_1^{nb}(t), \dots, \mathbf{u}_n^{nb}(t)\}$, which is the bootstrap data of $\{\mathbf{u}_1(t), \dots, \mathbf{u}_n(t)\}$ with $\mathbf{u}_i(t) = (\lambda_i(t), \lambda_i(t))$.

In contrast, $\widetilde{\lambda}_{h_i,2}^{nb}(t)$ needs to be computed from the raw longitudinal bootstrap sample $\{\mathbf{y}_1^{nb}, \dots, \mathbf{y}_n^{nb}\}$. Therefore, the bootstrap estimator $\widetilde{\lambda}_{h_i,2}^{nb}(t)$ requires more computational time and memory space than those for the bootstrap estimator $\widehat{\lambda}_{h_i,2}^{nb}(t)$. In the succeeding subsections, the bootstrap analogues will be used in the bandwidth selection and the construction of approximated confidence intervals.

3.1. Bandwidth selection

Conditioning on the recurrent event data, the bootstrap analogues $\widehat{\lambda}_{h_i,2}^{nb}(t)$ and $\widetilde{\lambda}_{h_i,2}^{nb}(t)$ both can be shown to be the approximately unbiased estimators of $\widehat{\lambda}_{h_i,2}(t)$ and $\widetilde{\lambda}_{h_i,2}(t)$. 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 biases of $\widehat{\lambda}_{h_i,2}(t)$ and $\widetilde{\lambda}_{h_i,2}(t)$ by

$$\widehat{b}_{h_i}(t) = \widehat{\lambda}_{h_i,2}(t) - \widehat{\lambda}_{h_i,4}(t) \quad \text{and} \quad \widetilde{b}_{h_i}(t) = \widetilde{\lambda}_{h_i,2}(t) - \widetilde{\lambda}_{h_i,4}(t). \quad (3)$$

In practical implementation, $K_{s,\lambda}(t-u)/h_i$ is often assigned to be $K_{s,\lambda}(t-u)/h_i = h_i^{-1} 1(s, (t-u)/h_i) K((t-u)/h_i)$ with $1(s, u)$ and $K(\cdot)$ being separately the $(l-1)$ th polynomial function of u and a symmetric density. The dominant bias term of each estimator at each time t can be estimated by using the estimators in equation (3) at the specified bandwidths as Schucany (1995). Under the considered data setting, we find that the method performs well in the numerical studies. Using the estimated dominant bias terms, say, $\widehat{b}_{h_i}^*$ and $\widetilde{b}_{h_i}^*$ for the bias adjustment and the sample variances of the bootstrap estimators, it is reasonable to approximate the mean squared errors, say, $mse(\widehat{\lambda}_{h_i,2}(t))$ and $mse(\widetilde{\lambda}_{h_i,2}(t))$ of $\widehat{\lambda}_{h_i,2}(t)$ and $\widetilde{\lambda}_{h_i,2}(t)$, respectively, by

$$mse^{nb}(\widehat{\lambda}_{h_i,2}(t)) = \widehat{b}_{h_i}^{*2}(t) + V^{nb}(\widehat{\lambda}_{h_i,2}^{nb}(t)) \quad \text{and} \quad mse^{nb}(\widetilde{\lambda}_{h_i,2}(t)) = \widetilde{b}_{h_i}^{*2}(t) + V^{nb}(\widetilde{\lambda}_{h_i,2}^{nb}(t)), \quad (4)$$

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

$$mise^{nb}(\widehat{\lambda}_{h,2}) = \int_0^{T_0} mse^{nb}(\widehat{\lambda}_{h,2}(t)) \pi(t) dt \quad \text{and} \quad mise^{nb}(\widetilde{\lambda}_{h,2}) = \int_0^{T_0} mse^{nb}(\widetilde{\lambda}_{h,2}(t)) \pi(t) dt, \quad (5)$$

where h is a positive time-independent bandwidth and $w(t)$ is a non-negative weight function. Generally, the mean integrated squared errors provide a variety of global distance measures to investigate the accuracy of the considered estimator. In implementation, the weight function $w(t)$ is often assigned to be a uniform density function.

3.2. Confidence intervals

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 to be 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 $(1 - \alpha)$ pointwise and simultaneous confidence intervals for $\lambda(t)$ with bias correction.

- 1.1 Draw the bootstrap sample $\{\mathbf{u}_1^{nb}, \dots, \mathbf{u}_n^{nb}\}$ of size n with replacement from the condensed data $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$.
- 1.2 Compute the bootstrap estimator $\widehat{\lambda}_{h,2}^{nb}(t)$ and $D_h^{nb}(t) = \widehat{\lambda}_{h,2}^{nb}(t) - \widehat{\lambda}_{h,2}(t)$ based on the condensed bootstrap data drawn in step 1.1.
- 1.3 Repeat steps 1.1–1.2 B times, and compute the standard errors and the percentiles of B bootstrap estimators $\widehat{\lambda}_{h,2}^{nb}(t)$ and $D_h^{nb}(t)$, and $M^{nb}(\widehat{\lambda}_h) = \sup_t |D_h^{nb}(t)/se^{nb}(\widehat{\lambda}_{h,2}^{nb}(t))|$, where $se^{nb}(\cdot)$ denotes the standard error of bootstrap estimators and t is a time period.
- 1.4 Construct separately the approximated $(1 - \alpha)$ pointwise and simultaneous confidence intervals for $\lambda(t)$ via

$$\left(\widehat{\lambda}_{h,2}(t) - \widehat{b}_{h_t}^*(t) - Q_{1-\alpha/2}^{nb}(\widehat{\lambda}_{h,2}^{nb}(t)), \widehat{\lambda}_{h,2}(t) - \widehat{b}_{h_t}^*(t) - Q_{\alpha/2}^{nb}(\widehat{\lambda}_{h,2}^{nb}(t))\right) \quad (6)$$

and

$$\widehat{\lambda}_{h,2}(t) - \widehat{b}_{h_t}^*(t) \pm se^{nb}(\widehat{\lambda}_{h,2}^{nb}(t)) Q_{1-\alpha}^{nb}(M^{nb}(\widehat{\lambda}_h)). \quad (7)$$

Here, $Q_p^{nb}(\cdot)$ is the p th quantile of B bootstrap estimators.

As an alternative to the above bootstrap percentile intervals in step 1.4, one can also use separately the pointwise and simultaneous intervals of the forms

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

and

$$\widehat{\lambda}_{h,2}(t) - \widehat{b}_{h_t}^*(t) \pm \ln(-\ln(\sqrt{1-\alpha})) se^{nb}(\widehat{\lambda}_{h,2}^{nb}(t)),$$

where $z_{(1-\alpha/2)}$ is the $(1 - \alpha/2)$ quantile value of the standard normal distribution. 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 $\{\mathbf{y}_1, \dots, \mathbf{y}_n\}$, and then compute the bootstrap estimator $\widehat{\lambda}_{h,2}^{nb}(t)$ based on the bootstrap data $\{\mathbf{y}_1^{nb}, \dots, \mathbf{y}_n^{nb}\}$. Same with steps 1.2–1.4, the approximated

$(1 - \alpha)$ pointwise and simultaneous confidence intervals for $\lambda(t)$ are constructed via substituting $\tilde{\lambda}_{h_t,2}(t)$, $\tilde{b}_{h_t}^*(t)$ and $\tilde{\lambda}_{h_t,2}^{nb}(t)$ for $\hat{\lambda}_{h_t,2}(t)$, $\hat{b}_{h_t}^*(t)$ and $\hat{\lambda}_{h_t,2}^{nb}(t)$, respectively.

4. Asymptotic properties

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

- (A2) Y_{fs} are independent and identically distributed with the cumulative distribution function $F_Y(y)$ and the probability measure $P_Y(y)$.
- (A3) $\lambda(t)$ is four times differentiable and bounded with $\lambda(t) > 0$ for some positive constant.
- (A4) $E[dN(u)dN(v)dN(w)] = c_0 du dv dw$ for some positive constant c_0 and for all $u, v, w \in [0, T_0]$.
- (A5) Define $\zeta_\nu(t, h_t, s) = \int_{(t-s)/h_t}^{t/h_t} |\lambda(s, u)K(u)|^\nu du$ with $\zeta_3(t, h_t, s) < \infty$.
- (A6) $h_t = n^{-1/5} h_{0t}$ (or $h = n^{-1/5} h_0$) for some positive bounded constant $h_0(h_0)$.

Let $\lambda(t) = \int_0^{T_0} K_{T_0,2}((t-u)/h_t) \lambda'(1 - F_Y(u))(u) dN_\lambda(u)$. By the law of large numbers, it be can shown that

$$\frac{\delta_n(t)}{n} \xrightarrow{p} (1 - F_Y(t)), \text{ as } n \rightarrow \infty. \tag{8}$$

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

$$\hat{\lambda}_{h_t,2}(t) = (1 - F_Y(t))^{-1} \left(\frac{1}{n} \sum_{i=1}^n \xi_i(t) \right) (1 + O_p(n^{-1/2})), \tag{9}$$

and

$$\tilde{\lambda}_{h_t,2}(t) = \left(\frac{1}{n} \sum_{i=1}^n \zeta_i(t) \right) (1 + O_p(n^{-1/2})). \tag{10}$$

Before establishing theorem 1, a technical lemma is stated below.

Lemma 1

Suppose that assumptions (A1)–(A6) and $F_Y(T_0) < 1$ are satisfied. Then,

$$E[\xi_i(t)] = (1 - F_Y(t)) (\lambda(t) + b_{1,h_t}(t) h_t^2) + o(h_t^2), \tag{11}$$

$$E[\zeta_i(t)] = \lambda(t) + b_{2,h_t}(t) h_t^2 + o(h_t^2), \tag{12}$$

$$E[\xi_i^\nu(t)] = \lambda(t) \left(\int_{\{s \geq t\}} \gamma_\nu(t, h_t, s) dP_Y(s) \right) h_t^{-\nu+1} + o(h_t^{-\nu+1}), \tag{13}$$

and

$$E \left[\zeta_i^\nu(t) \right] = \left(\frac{\lambda(t) \gamma_\nu(t, h_t, T_0)}{(1 - F_Y(t))^{\nu-1}} \right) h_t^{-\nu+1} + o(h_t^{-\nu+1}), \quad (14)$$

where

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

with $\beta_{2,2}(t, h_t, s) = \int_{(t-s)/h_t}^{t/h_t} u^2 \gamma_2(s, u) K(u) du$ and $\gamma_\nu(t, h_t, s) = \int_{(t-s)/h_t}^{t/h_t} \gamma_\nu(s, u) K(u) du$ 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)–(A6), and $F_Y(T_0) < 1$ are satisfied. Then, for all $z \in \mathbf{R}$ and n

$$\sup_z \left| P \left(\frac{\sqrt{nh_t} (\widehat{\lambda}_{h_t,2}(t) - \lambda(t) - b_{1,h_t}(t))}{\sigma_{1,h_t}(t)} \leq z \right) - \Phi(z) \right| \rightarrow 0 \quad (15)$$

and

$$\sup_z \left| P \left(\frac{\sqrt{nh_t} (\widetilde{\lambda}_{h_t,2}(t) - \lambda(t) - b_{2,h_t}(t))}{\sigma_{2,h_t}(t)} \leq z \right) - \Phi(z) \right| \rightarrow 0, \quad (16)$$

where

$$\sigma_{1,h_t}^2(t) = \frac{\lambda(t) \left(\int_{\{y \geq t\}} \mathcal{S}_2(t, h_t, y) dP_Y(y) \right)}{(1 - F_Y(t))^2}, \quad \sigma_{2,h_t}^2(t) = \frac{\lambda(t) \mathcal{S}_2(t, h_t, T_0)}{(1 - F_Y(t))}.$$

Proof. See appendix.

When the estimators are computed based on a linear function and a symmetric density separately for $\gamma_2(\cdot)$ and $K(\cdot)$, it can be derived algebraically that $\sigma_{1,h_t}^2(t) \sim \sigma_{2,h_t}^2(t)$ as h converges to 0. In addition, 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 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.

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

Lemma 2

Suppose that assumptions (A1)–(A3) and (A6), and $F_Y(T_0) < 1$. Then, as $n \rightarrow \infty$,

$$\sqrt{nh_t}(\widehat{b}_{h_t}(t) - b_{1,h_t}(t)) \xrightarrow{P} 0 \text{ and } \sqrt{nh_t}(\widetilde{b}_{h_t}(t) - b_{2,h_t}(t)) \xrightarrow{P} 0. \quad (17)$$

Before deriving the asymptotic normalities of the bootstrap estimators, let $P^{nb}(\cdot)$ denote the probability measure conditioning on the sample $\{\mathbf{y}_1, \dots, \mathbf{y}_n\}$. The next theorem and lemma 2 show the validity of the approximated bootstrap confidence intervals in section 3.2, i.e. the sampling distributions of $(\widehat{\lambda}_{h_i,2}^{nb}(t) - \widehat{\lambda}_{h_i,2}(t))$ and $(\widetilde{\lambda}_{h_i,2}^{nb}(t) - \widetilde{\lambda}_{h_i,2}(t))$ can be used to approximate the distributions of $(\widehat{\lambda}_{h_i,2}(t) - \lambda(t) - \widehat{b}_{h_i}^*(t))$ and $(\widetilde{\lambda}_{h_i,2}(t) - \lambda(t) - \widetilde{b}_{h_i}^*(t))$.

Theorem 2

Suppose that assumptions (A1)–(A6), and $F_Y(T_0) < 1$ are satisfied. Then, for all $z \in \mathbf{R}$ and $n \rightarrow \infty$,

$$\sup_z |P^{nb} \left(\frac{\sqrt{nh_t}(\widehat{\lambda}_{h_i,2}^{nb}(t) - \widehat{\lambda}_{h_i,2}(t))}{\sigma_{1,h_t}(t)} \leq z \right) - \Phi(z)| \xrightarrow{P} 0 \quad (18)$$

and

$$\sup_z |P^{nb} \left(\frac{\sqrt{nh_t}(\widetilde{\lambda}_{h_i,2}^{nb}(t) - \widetilde{\lambda}_{h_i,2}(t))}{\sigma_{2,h_t}(t)} \leq z \right) - \Phi(z)| \xrightarrow{P} 0. \quad (19)$$

Proof. See appendix.

Before showing the validity of the bootstrap simultaneous confidence intervals in section 3.2, the condition, $((nh)^{-1} + h)(\ln(n))^4 \rightarrow 0$ as $n \rightarrow \infty$, is examined first. It is indicated from assumption (A6) that the bandwidth selection satisfies this requirement. As for the function $K(\cdot)$, the Gaussian kernel density function will be assigned in estimation and is shown to satisfy the assumptions of Jhun (1988). Using the expressions in (9) and (10) with global bandwidth h , and paralleling the proof of Hall (1993), one can show that the sampling distributions of $M^{nb}(\widehat{\lambda}_h)$ and $M^{nb}(\widetilde{\lambda}_h)$ can be used to approximate the distributions of $M(\widehat{\lambda}_h) = \sqrt{nh} \sup_t |(\widehat{\lambda}_{h,2}(t) - \lambda(t) - b_{1,h}(t)) / \sigma_{1,h}(t)|$ and $M(\widetilde{\lambda}_h) = \sqrt{nh} \sup_t |(\widetilde{\lambda}_{h,2}(t) - \lambda(t) - b_{2,h}(t)) / \sigma_{2,h}(t)|$ respectively.

Theorem 3

Suppose that assumptions (A1)–(A6), and $F_Y(T_0) < 1$ are satisfied. Then, for all $z \in \mathbf{R}^+$ and $n \rightarrow \infty$,

$$\sup_z |P^{nb} \left(\sqrt{2 \ln(h^{-1})} M^{nb}(\tilde{\lambda}_h) \leq z \right) - P \left(\sqrt{2 \ln(h^{-1})} M(\tilde{\lambda}_h) \leq z \right)| \xrightarrow{P} 0 \quad (20)$$

and

$$\sup_z |P^{nb} \left(\sqrt{2 \ln(h^{-1})} M^{nb}(\tilde{\lambda}_h) \leq z \right) - P \left(\sqrt{2 \ln(h^{-1})} M(\tilde{\lambda}_h) \leq z \right)| \xrightarrow{P} 0. \quad (21)$$

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:

$$\text{Case 1. } \lambda_i(t) = Z_i + \phi_{10}(t) \quad \text{with} \quad \phi_{10}(t) = 0.5(\sin(t\pi/6) - 1).$$

$$\text{Case 2. } \lambda_i(t) = Z_i \phi_{20}(t) \quad \text{with} \quad \phi_{20}(t) = 0.5(\sin(t\pi/6) + 1).$$

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

$$\lambda(t) = 0.5(\sin(t\pi/6) + 1).$$

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 process is repeated 500 times. For each simulated data set, the kernel estimators $\hat{\lambda}_{h,2}(t)$ and $\tilde{\lambda}_{h,2}(t)$ are computed by (1) and (2) with the normal density for $K(\cdot)$. Moreover, the local and global optimal bandwidths are selected separately via the bootstrap mean squared errors in (4 and 5). The 500 simulation averages of both estimated curves with bias correction are very close to the true rate function. Here, we find that the bias correction improves the accuracy of both estimators, and the variation of $\tilde{\lambda}_{h,2}(t)$ is smaller than that of $\hat{\lambda}_{h,2}(t)$ 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 at 39 equally spaced time points within the period (0, 4). As the performances of both methods are comparable, we present only the empirical coverage probabilities of the moment estimation method. Table 1 summarizes the empirical coverage probabilities of 95% pointwise confidence intervals for $\lambda(t)$ at selected time points and different bandwidths. It is shown

that the coverage probabilities for the procedures constructed using the optimal local bandwidths are generally close to the nominal level. However, poor coverage probabilities appear at extremely small or large bandwidths. The same conclusion can also be drawn for the empirical coverage probabilities of 95% simultaneous confidence intervals for (t) in Table 2.

6. A data example

The data set used here involves 450 HIV-negative intravenous drug users who entered the study before 1 August 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 1 August 1993 and 31 December 1997. Details of this study can be found in Vlahov *et al.* (1991).

Let y_i be the time length from 1 August 1993 to the date of the last visit for the i th 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 (4) and (5) are used separately to select local and global optimal bandwidths for $\hat{\lambda}_{h_i,2}(t)$ and $\tilde{\lambda}_{h_i,2}(t)$. Moreover, the corresponding 95% pointwise and simultaneous confidence intervals of the hospitalization rate estimators are computed at 430 equal spaced time points within the period (0, 4.31). It appears in Figs. 1(a),(b) that both the estimated curves $\hat{\lambda}_{h_i,2}(t)$ and $\tilde{\lambda}_{h_i,2}(t)$ imply the same biological explanation except wider confidence intervals for the moment method. From the figures, we can see that the hospitalization occurrence rates are lower than 1 and the peak of the curve occurs in about the later half of the study. This might indicate that the intravenous drug users received some medical intervention during the study period, and, hence, tend to have a decreasing trend in frequencies of hospitalization.

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 have pointed 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 these 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.

Acknowledgments

The first author's research was partially supported by the National Science Council grant 9-2118-M-002-001 (Taiwan), and the second author's research was partially supported by the National Institute of Health grants CA098252 (USA). We would like to thank the referee for valuable comments. We are also grateful to Dave Vlahov

and Stratheed at Johns Hopkins University for providing the anonymous ALIVE data. Provision of the data was supported by National Institute on Drug Abuse grants DA04334 and DA08009.

Appendix

Proof of lemma 1

In the following proof, we only derive the statements of (11) and (13). Same arguments can be used in the derivation of (12) and (14). By assumptions (A1)–(A3) and the Taylor expansion, we can get

$$\begin{aligned}
 E[\xi_i(t)] &= E\left[\delta_i(t) \int_0^{Y_i} K_{Y_i,2}\left(\frac{t-u}{h_t}\right) dN_i(u)\right] = \int_{\{s \geq t\}} \left(\int_0^s K_{s,2}\left(\frac{t-u}{h_t}\right) E[dN_i(u)]\right) dP_Y(s) \\
 &= \int_{\{s \geq t\}} \left(\int_0^s K_{s,2}\left(\frac{t-u}{h_t}\right) \lambda(u) du\right) dP_Y(s) \\
 &= \int_{\{s \geq t\}} \left(\int_{t-s/h_t}^{t/h_t} \alpha_2(s, u) K(u) \left(\sum_{k=0}^2 \frac{\lambda^{(k)}(t)(-uh_t)^k}{k!} + o((uh)^2)\right) du\right) dP_Y(s) \quad (22) \\
 &= (1 - F_Y(t)) \lambda(t) + \left(\frac{\lambda^{(2)}(t) \int_{\{s \geq t\}} \beta_{2,2}(t, h_t, s) dP_T(s)}{2}\right) h_t^2 + o(h_t^2).
 \end{aligned}$$

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

$$E[\xi_i^2(t)] = I_A + II_A, \quad (23)$$

where

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

and

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

By assumptions (A1)–(A3), we can show that

$$\begin{aligned}
 I_A &= E\left[\delta_i(t) \int_0^{Y_i} K_{Y_i,2}^2\left(\frac{t-u}{h_t}\right) dN_i(u)\right] = \int_{\{s \geq t\}} \int_0^s K_{s,2}^2\left(\frac{t-u}{h_t}\right) E[dN_i(u)] dP_Y(s) \\
 &= \int_{\{s \geq t\}} \left(\int_0^s K_{s,2}^2\left(\frac{t-u}{h_t}\right) \lambda(u) du\right) dP_Y(s) = \lambda(t) \left(\int_{\{s \geq t\}} \gamma_2(t, h_t, s) dP_Y(s)\right) h_t^{-1} + o(h_t^{-1}). \quad (24)
 \end{aligned}$$

Assumption (A4) implies that there exists a positive constant c_1 such that

$$\begin{aligned}
 II_A &= \int_{\{s \geq t\}} \left(\int \int_{\{u \neq v\}} K_{s,2}\left(\frac{t-u}{h_t}\right) K_{s,2}\left(\frac{t-v}{h_t}\right) E[dN_i(u) dN_i(v)]\right) dP_Y(s) \\
 &\leq c_1 \int_{\{s \geq t\}} \left(\int \int_{\{u \neq v\}} K_{s,2}\left(\frac{t-u}{h_t}\right) K_{s,2}\left(\frac{t-v}{h_t}\right) dudv\right) dP_Y(s) \quad (25) \\
 &= c_1 (1 - F_T(t)) (1 + o(1)).
 \end{aligned}$$

Substituting (24) and (25) into (23), $E[\xi_i^2(t)]$ is obtained. Along the same lines as the derivation of $E[\xi_i^2(t)]$, statement (15) for $v = 3$ follows.

Proof of theorem 1

From (9), we make the linear approximation

$$\frac{\widehat{\lambda}_{h,2}(t) - E[\widehat{\lambda}_{h,2}(t)]}{\sqrt{V(\widehat{\lambda}_{h,2}(t))}} = \frac{\sum_{i=1}^n (\xi_i(t) - E[\xi_i(t)])}{\sqrt{nV(\xi_i(t))}} (1 + o_p(1)). \quad (26)$$

By the Berry–Esséen theorem, we obtain the following inequality:

$$\sup_z \left| P \left(\frac{\sum_{i=1}^n (\xi_i(t) - E[\xi_i(t)])}{\sqrt{nV(\xi_i(t))}} \leq z \right) - \Phi(z) \right| \leq d \frac{(nE[|\xi_i(t) - E[\xi_i(t)]|^3])}{(nV(\xi_i(t)))^{3/2}}, \quad (27)$$

where d is a positive constant independent of n and $\lambda(t)$. As $|\lambda(t)|$, $\lambda(t)$ and $N_\lambda(t)$ s are positive random variables, it implies that

$$E[|\xi_i(t) - E[\xi_i(t)]|^3] \leq E[(\eta_i(t))^3] + 6[(\eta_i(t))^2]E[\eta_i(t)] + (E[\eta_i(t)])^3, \quad (28)$$

where $\lambda(t) = \int_0^t |K_{t,2}((t-u)/h)| dN_\lambda(u)$. Similar to the derivation of lemma 1, we can obtain, for $v = 1, 2, 3$,

$$E[(\eta_i(t))^v] = \lambda(t) \left(\int_{\{s \geq t\}} S_v(t, h, s) dP_\tau(s) \right) h^{-v+1} + o(h^{v+1}). \quad (29)$$

From (28), (29) and lemma 1, we can imply

$$\frac{nE[|\xi_i(t) - E[\xi_i(t)]|^3]}{(nV(\xi_i(t)))^{3/2}} \leq (nh_t)^{-1/2} \left(\frac{\int_{\{s \geq t\}} S_3(t, h_t, s) dP_\tau(s)}{\int_{\{s \geq t\}} \gamma_2(t, h_t, s) dP_\tau(s)} \right) \rightarrow 0, \quad (30)$$

By assumption (A6) and substituting (30) into (27), we derive

$$\sup_z \left| P \left(\frac{\sum_{i=1}^n (\xi_i(t) - E[\xi_i(t)])}{\sqrt{nV(\xi_i(t))}} \leq z \right) - \Phi(z) \right| \rightarrow 0, \text{ as } n \rightarrow \infty. \quad (31)$$

Finally, the proof of (15) is completed by (26) together with (31),

$$E[\widehat{\lambda}_{h,2}(t)] = \lambda(t) + b_{1,h}(t) h_t^2 + o(h_t^2) \quad (32)$$

and

$$V(\widehat{\lambda}_{h,2}(t)) = \sigma_{1,h_t}^2(t) (nh_t)^{-1} + o((nh_t)^{-1}). \quad (33)$$

For the proof of (16), we first make the linear approximation

$$\frac{\tilde{\lambda}_{h_t,2}(t) - E[\tilde{\lambda}_{h_t,2}(t)]}{\sqrt{V(\tilde{\lambda}_{h_t,2}(t))}} = \frac{\sum_{i=1}^n (\xi_i(t) - E[\xi_i(t)])}{\sqrt{nV(\xi_i(t))}} (1 + o_p(1)). \quad (34)$$

Then, paralleling the steps from (27–31) with (34),

$$E[\tilde{\lambda}_{h_t,2}(t)] = \lambda(t) + b_{2,h_t}(t) h_t^2 + o(h_t^2) \quad (35)$$

and

$$V(\tilde{\lambda}_{h_t,2}(t)) = \sigma_{2,h}^2(t) (nh_t)^{-1} + o((nh_t)^{-1}), \quad (36)$$

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

Proof of theorem 2

As the derivation of the proof for the bootstrap estimator $\tilde{\lambda}_{h_t,2}^{nb}(t)$ is similar those of $\widehat{\lambda}_{h_t,2}^{nb}(t)$, for the space of presentation, we derive only the asymptotic normalities of $\widehat{\lambda}_{h_t,2}^{nb}(t)$. Let $\xi_i^{nb}(t) = \delta_i^{nb}(t) \int_0^{Y_i^{nb}} K_{Y_i^{nb},2}((t-u)/h) dN_i^{nb}(u)$. By the law of large numbers, we can show that, for any $\varepsilon > 0$,

$$P^{nb} \left(\left| \frac{\delta_i^{nb}(t)}{n} - \frac{\delta_i(t)}{n} \right| > \varepsilon \right) \rightarrow 0. \quad (37)$$

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

$$\widehat{\lambda}_{h_t,2}^{nb}(t) \left(\frac{\delta_i^{nb}(t)}{n} \right)^{-1} \left(\frac{1}{n} \sum_{i=1}^n \xi_i^{nb}(t) \right) = \left(\frac{\delta_i(t)}{n} \right)^{-1} \left(\frac{1}{n} \sum_{i=1}^n \xi_i^{nb}(t) \right) (1 + o_{p^{nb}}(1)). \quad (38)$$

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

$$\frac{\widehat{\lambda}_{h_t,2}^{nb}(t) - E^{nb}[\widehat{\lambda}_{h_t,2}(t)]}{\sqrt{V^{nb}(\widehat{\lambda}_{h_t,2}(t))}} = \frac{\sum_{i=1}^n (\xi_i^{nb}(t) - E^{nb}[\xi_i(t)])}{\sqrt{nV^{nb}(\xi_i^{nb}(t))}} (1 + o_{p^{nb}}(1)). \quad (39)$$

In addition, by the Berry–Essé theorem, there exists a positive constant c_2 such that

$$\sup_z \left| P^{nb} \left(\frac{\sum_{i=1}^n (\xi_i^{nb}(t) - E^{nb}[\xi_i^{nb}(t)])}{\sqrt{nV^{nb}(\xi_i^{nb}(t))}} \leq z \right) - \Phi(z) \right| \leq c_2 \frac{nE^{nb} \left[|\xi_i^{nb}(t) - E[\xi_i^{nb}(t)]|^3 \right]}{(nV^{nb}(\xi_i^{nb}(t)))^{3/2}}. \quad (40)$$

As

$$E^{nb} \left[(\xi_i^{nb}(t))^v \right] = \frac{\sum_{i=1}^n (\xi_i(t))^v}{n}, \text{ for } v=1, 2, 3, \quad (41)$$

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

$$hV^{nb} \left(\xi_i^{nb}(t) \right) = \left(\frac{1}{n} \sum_{i=1}^n (\xi_i(t))^2 - \left(\frac{1}{n} \sum_{i=1}^n \xi_i(t) \right)^2 \right) \xrightarrow{p} \sigma_{1,h_t}^2(t) \quad (42)$$

and

$$h_t^2 E^{nb} \left[\left| \xi_i^{nb}(t) - E \left[\xi_i^{nb}(t) \right] \right|^3 \right] \leq h^2 \left(\frac{1}{n} \sum_{i=1}^n (\eta_i(t))^3 + 6 \bar{\eta}(t) \left(\frac{1}{n} \sum_{i=1}^n (\eta_i(t))^2 \right) + \left(\bar{\xi}(t) \right)^3 \right) \xrightarrow{p} \lambda(t) \left(\int_{\{s \geq t\}} \mathcal{S}_3(t, h_t, s) dP_Y(s) \right). \quad (43)$$

Substituting (42) and (43) into (41), we get

$$\sup_z \left| P^{nb} \left(\frac{\sum_{i=1}^n \left(\xi_i^{nb}(t) - \left(\sum_{i=1}^n \xi_i(t) / n \right) \right)}{\sqrt{nh_t^{-1} \sigma_{1,h_t}^2(t)}} \leq z \right) - \Phi(z) \right| \xrightarrow{p} 0. \quad (44)$$

Finally, (20) is obtained from (39) and (44).

References

- Bartoszy ski R, Brown BW, McBride CM, Thompson JR. Some nonparametric techniques for estimating the intensity function of a cancer related nonstationary Poisson process. *Ann. Statist.* 1981; 9:1050–1060.
- Efron B. Bootstrap methods: another look at the jackknife. *Ann. Statist.* 1979; 7:1–26.
- Gasser, Th.; Müller, H-G. Kernel estimation of regression functions. Smoothing techniques for curve estimation. Gasser, Th.; Rosenblatt, M., editors. Springer-Verlag; Berlin: 1979. p. 23-68. *Springer Lecture Notes in Mathematics No. 757*
- Hall P. On edgeworth expansion and bootstrap confidence bands in nonparametric curve estimation. *J. Roy. Statist. Soc. Ser. B.* 1993; 55:291–304.
- Jhun M. Bootstrap density estimates. *Comm. Statist. Theory Methods.* 1988; 17:61–78.
- Lawless JF, Nadeau C. Some simple robust method for the analysis of recurrent events. *Technometrics.* 1995; 37:158–168. Lin; Fleming, TR. Springer-Verlag; New York: p. 37-49.
- Lin DY, Wei LJ, Yang I, Ying Z. Semiparametric regression for the mean and rate functions of recurrent events. *J. Roy. Statist. Soc. Ser. B.* 2000; 62:711–730.
- Nelson WB. Graphical analysis of system repair data. *J. Qual. Technol.* 1988; 20:24–35.
- Nelson WB. Confidence limits for recurrence data – applied to cost or number of product repairs. *Technometrics.* 1995; 37:147–157.
- Pepe MS, Cai J. Some graphical displays and marginal regression analyses for recurrent failure times and time dependent covariates. *J. Amer. Statist. Assoc.* 1993; 88:811–820.
- Ramlau-Hansen H. Smoothing counting process intensities by means of kernel functions. *Ann. Statist.* 1983; 11:453–466.
- Schucany WR. Adaptive bandwidth choice for kernel regression. *J. Amer. Statist. Assoc.* 1995; 90:535–540.
- Scheike TH. The additive nonparametric and semiparametric Aalen model as the rate function for a counting process. *Lifetime Data Anal.* 2002; 8:229–246. [PubMed: 12182120]
- Tanner MA, Wong WH. The estimation of the hazard function from randomly censored data by the kernel method. *Ann. Statist.* 1983; 11:989–993.
- Vlahov D, Anthony JC, Muñoz A, Margolick J, Nelson KE, Celentano DD, Solomon L, Polk BF. The ALIVE study: a longitudinal study of HIV-1 infection in intravenous drug users: description of methods. *J. Drug Issues.* 1991; 21:759–776.

Yandell BS. Nonparametric inference for rates with censored survival data. *Ann. Statist.* 1983; 11:1119–1135.

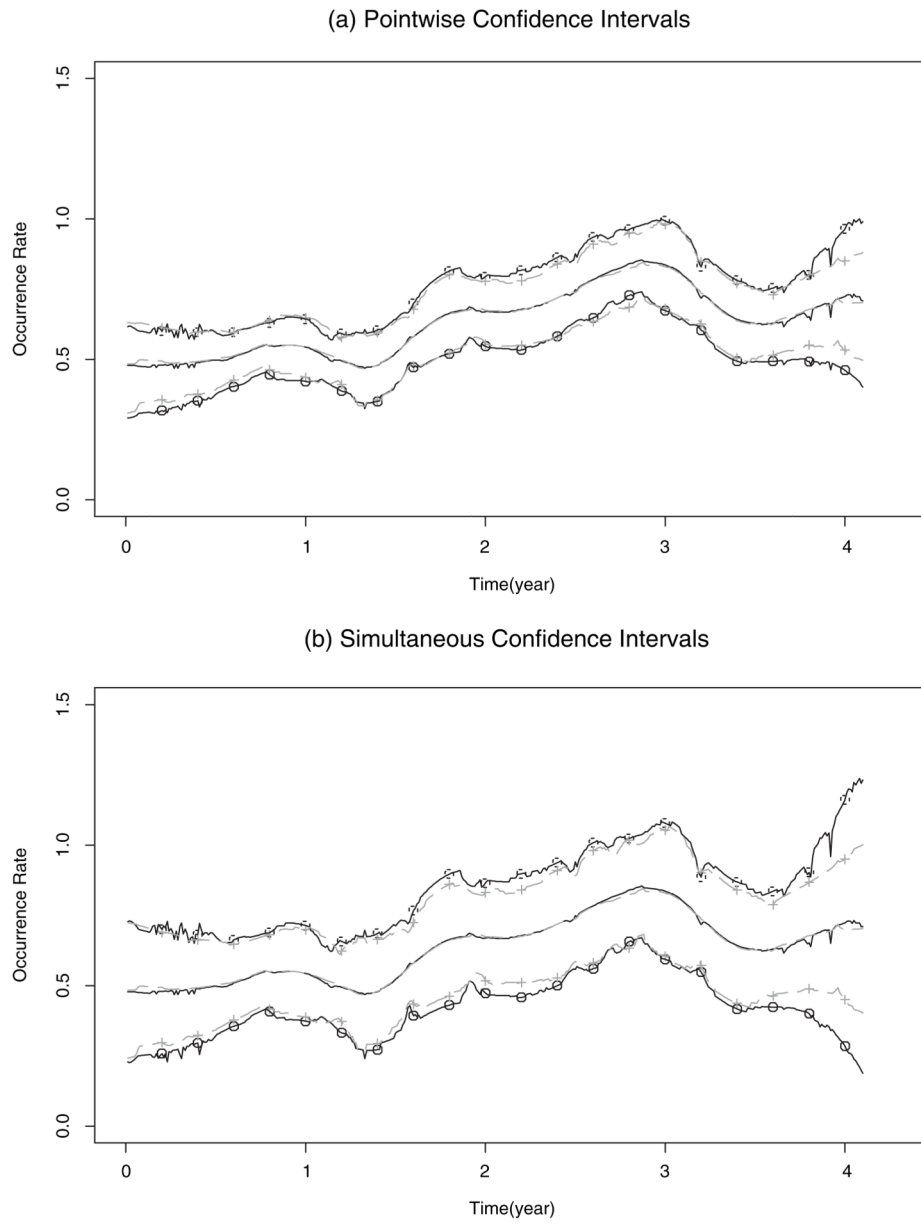


Fig. 1. The solid and dashed curves represent respectively the estimated rate functions $\widehat{\lambda}_{h,2}(t)$ and $\widetilde{\lambda}_{h,2}(t)$ of HIV-negative intravenous drug users with the corresponding 95% bootstrap intervals indicated by o and $+$.

Table 1
The empirical coverage probabilities of the 95% pointwise bootstrap confidence intervals for (t) , based on the moment estimation method

Time point	0.5	1.0	1.5	2.0	2.5	3.0	3.5
(Case 1) \hat{h}_t	0.956	0.914	0.918	0.928	0.938	0.956	0.986
$h_t = 0.05$	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$h_t = 0.37$	0.852	0.806	0.830	0.796	0.824	0.868	0.910
(Case 2) \hat{h}_t	0.954	0.920	0.940	0.946	0.926	0.942	0.976
$h_t = 0.05$	1.000	0.998	1.000	0.998	0.998	0.996	1.000
$h_t = 0.37$	0.846	0.816	0.848	0.692	0.828	0.842	0.898

Table 2
The empirical coverage probabilities of the 95% simultaneous bootstrap confidence intervals for (t) , based on the moment estimation method

Time period	(0,4)	[0.5, 3.5]	[1.0, 3.0]	[1.5, 3.5]
(Case 1) \hat{h}	0.986	0.970	0.972	0.958
$h = 0.05$	1.000	1.000	1.000	1.000
$h = 0.37$	0.660	0.514	0.538	0.592
(Case 2) \hat{h}	0.986	0.968	0.960	0.968
$h = 0.05$	1.000	1.000	1.000	1.000
$h = 0.37$	0.638	0.550	0.588	0.650