行政院國家科學委員會專題研究計畫　成果報告

遞迴事件發生率函數之信賴區間建立(2/2)

計畫類別：個別型計畫
計畫編號：NSC93-2118-M-002-001-
執行期間：93年08月01日至94年07月31日
執行單位：國立臺灣大學數學系暨研究所

計畫主持人：江金倉

報告類型：完整報告
報告附件：出席國際會議研究心得報告及發表論文
處理方式：本計畫可公開查詢

中　華　民　國　94年8月11日

Random Weighted Bootstrap Method For Recurrent Events With Informative Censoring

August 11, 2005

Abstract

Using the data from the AIDS Link to Intravenous Experiences cohort study as an example, an informative censoring model was used to characterize the repeated hospitalization process of a group of patients. Under the informative censoring assumption, the estimators of the baseline rate function and the regression parameters were shown to be influenced by a latent variable in the considered model. It becomes impractical to directly estimate the unknown quantities in the moments of the estimators for the bandwidth selection of a smoothing estimator and the construction of confidence intervals, which are respectively based on the asymptotic mean squared errors and the asymptotic distributions of the estimators. To overcome these difficulties, we develop a random weighted bootstrap procedure to select appropriate bandwidths and to construct approximated confidence intervals. One can see that our method is simple and faster to implement from a practical point of view, and is at least as accurate as other bootstrap methods. In this article, it is shown that the proposed method is useful through the performance of a Monte Carlo simulation. An application of our procedure is also illustrated by a recurrent event sample of intravenous drug users for inpatient cares over time.

1 Introduction

In this study, we consider recurrent events of the same type, which are frequently occurring in longitudinal studies and are collected from a group of independent subjects experiencing recurrent events, with the information of time-independent covariates. For the ith subject, $i = 1, \cdots, n$, let $N_i(t)$, $\mathbf{X}_i = (X_{i1}, \cdots, X_{ip})^T$, Y_i, and $\{T_{ij}\}_{j=1}^{m_i}$ separately denote the recurrent event process, the $p \times 1$ covariate vector, the minimum value of the censoring time (i.e., the time to the end of follow-up) and the end of study time T_0, and the ordered event times

Key words and phrases: asymptotic normality, bandwidth, confidence interval, cross-validation, independent censoring, informative censoring, kernel estimator, longitudinal study, naive bootstrap, occurrence rate function, Poisson process, wild bootstrap, weighted bootstrap.

in the time interval $[0, Y_i]$ with m_i being the number of the recurrent events occurring at or prior to Y_i. Generally, the main interest in this type of data is to estimate the occurrence rate function in the target population and to evaluate the effects of possible risk factors on the recurrent event process. In some applications, it is unreasonable to assume the censoring time to be independent of the recurrent event process because each subject's recurrent events could be terminated before or at the end of study by loss to follow-up or informative drop-out. Using the data from the AIDS Link to Intravenous Experiences cohort study as an example, Wang, Qin and Chiang (2001) proposed an informative censoring model. The model uses a latent variable to characterize the correlation between the recurrent event process and the censoring mechanism, yet the distributions of the latent variable and censoring time are both left as nonparametric components in the model. The model consists of the following assumptions.

(A1) Suppose there exists a nonnegative-valued latent variable Z_i so that, conditioning on $(\mathbf{x}_i, z_i)$, $N_i(t)$ is a non-stationary Poisson process with the subject-specific rate function $\lambda_i(t) = z_i \lambda_0(t) \exp(\boldsymbol{\beta}^T \mathbf{x}_i)$, where the baseline rate function $\lambda_0(t)$ is a continuous function, $\boldsymbol{\beta}$ is a $p \times 1$ parameter vector, and Z_i satisfies $E[Z_i | \mathbf{x}_i] = 1$.

(A2) Conditioning on $(\mathbf{x}_i, z_i)$, $N_i(\cdot)$ is independent of Y_i.

Based on the above model with or without using the information of covariates, Chiang and Wang (2003), Wang, Qin and Chiang (2001), and Wang and Chiang (2002) suggested the estimation methods for the baseline rate function $\lambda_0(t)$ and the regression parameters $\boldsymbol{\beta}$. From their works, it was shown that the moments of the estimators are influenced by a latent variable. Thus, the selection of bandwidths and the construction of confidence intervals, which are separately based on the asymptotic mean squared errors and the asymptotic distributions of the estimators, become impractical to directly estimate the unknown quantities in the moments of the estimators. To overcome these difficulties, we develop a random weighted bootstrap procedure which is simple and fast to implement from a practical point of view, and is at least as accurate as other bootstrap methods. Since the considered estimators involve vectors of observations, the traditional bootstrap methods, such as the naive bootstrap and the wild bootstrap procedures, in this setting are impractically slow in implementation. However, in our proposed approach, the random weighted bootstrap framework identifies alternative scheme without resampling from the data or generating the bootstrap data from the shape function of the recurrent event times and is markedly superior in terms of practical computational efficiency. One can find that the random weighted bootstrap framework discussed here encompasses many of the known bootstrap methods and can be applied to many other similar settings as well. The method which we highlight here can be seen as an extension of the Bayesian bootstrap and Bayesian Bootstrap clones (BBC) discussed in Rubin (1981), Lo (1987, 1991), among others.

In Sections 2, we summarize the estimation methods for the baseline rate function $\lambda_0(t)$ and the regression parameters $\boldsymbol{\beta}$. In Section 3, the random weighted bootstrap analogues are proposed and used to approximate the sampling quantities of interest related to the estimators. Based on the behaviors of the proposed random weighted bootstrap estimators, a class of approximated confidence intervals are constructed in Section 4. As for the selection of appropriate bandwidths, it was indicated from simulation experiments that a "leave one subject out" least squares cross-validation criterion is often inadequate and computationally inefficient. However, our random weighted bootstrap procedure provides an explicit and fast criterion for bandwidth selection on local and global bases. In Section 5, we conduct Monte Carlo simulations and apply the proposed procedure to the intravenous drug user data. Finally, proofs of the distributional properties of the random weighted bootstrap estimators are given in the Appendix.

2 Estimation

In this section, we summarize the estimation methods for the baseline rate function $\lambda_0(t)$ and the regression parameters $\boldsymbol{\beta}$ in the considered informative censoring regression model. Let $\Lambda_0(t)$ be the cumulative function of $\lambda_0(t)$ and define the density function $f(t) = \lambda_0(t)/\Lambda_0(T_0)$ for $0 \le t \le T_0$ with the cumulative distribution function $F(t)$. Under model assumptions (A1)-(A2), Wang, Qin and Chiang (2001) showed that, conditioning on $(m_i, y_i, z_i, \mathbf{x}_i)$, the unordered recurrent event times of $(T_{i1}, \cdots, T_{im_i})$ are independent and identically distributed with the truncated density function $f_i(t) = f(t)/F(y_i)$ for $t \in [0, y_i]$. Thus, $F(t)$ can be estimated by the nonparametric maximum likelihood estimator $\widehat{F}(t)$ of the form

$$\widehat{F}(t) = \prod_{\{s_{(l)} > t\}} \left(1 - \frac{d_{(l)}}{N_{(l)}}\right), \tag{1}$$

where $\left\{s_{(l)}\right\}$ are the order statistics of the event times $\{T_{ij}\}$, $d_{(l)} = \sum_{i=1}^n \sum_{j=1}^{m_i} I(T_{ij} = s_{(l)})$, and $N_{(l)} = \sum_{i=1}^n \sum_{j=1}^{m_i} I(T_{ij} \le s_{(l)} \le Y_i)$. Using the equality $E[m_i(F(Y_i))^{-1}|\mathbf{X}_i = \mathbf{x}_i] = \Lambda_0(T_0)\exp(\boldsymbol{\beta}^T\mathbf{x}_i)$ and substituting the nonparametric estimator $\widehat{F}(t)$ for $F(t)$, their estimators $\widehat{\beta}_0$ and $\widehat{\boldsymbol{\beta}}$ for $\beta_0(=\Lambda_0(T_0))$ and $\boldsymbol{\beta}$ are obtained from the estimating equations

$$\frac{1}{n}\sum_{i=1}^n (1, \mathbf{X}_i^T)^T \left(\frac{m_i}{\widehat{F}(Y_i)} - \beta_0 \exp(\boldsymbol{\beta}^T\mathbf{X}_i)\right) \exp(\boldsymbol{\beta}^T\mathbf{X}_i) = 0. \tag{2}$$

Under assumptions (A1)-(A2),

(**A3**) $\beta_0 > 0$ and $P(Z > 0, Y \ge T_0) > 0$, and

(**A4**) $G(t) = \int zI(y \ge t)dP_{Y,Z}(y, z)$ is a continuous function,

they also showed that the estimators converge to the multivariate normal distribution.

For the estimation of the baseline rate function $\lambda_0(t)$, we extend the smoothing estimation method of Chiang and Wang (2003), and Wang and Chiang (2002) to our data setting.

Multiplying the smoothing estimator of $f_i(t)$ by the estimator $\widehat{\Lambda}_0(t) = \widehat{\beta}_0 \widehat{F}(t)$ of $\Lambda_0(t)$ and using the information of subjects who are still at risk, our suggested kernel estimator $\widetilde{\lambda}_{h_t,2}(t)$ is given by

$$\widetilde{\lambda}_{h_t,2}(t) = \frac{\sum_{i=1}^{n} \frac{1}{n}\delta_i(t) D_{n,h_t,2}(\mathbf{U}_i, \widehat{\Lambda}_0, t)}{\delta.(t)}, \quad t \in [0, T_0], \tag{3}$$

where $\delta.(t) = \sum_{i=1}^{n} \delta_i(t)$, $\delta_i(t) = I(Y_i \geq t, m_i \geq 1)$, $\mathbf{U}_i = ((T_{i1}, \cdots, T_{im_i}), m_i, Y_i, \mathbf{X}_i)$, and

$$D_{n,h_t,2}(\mathbf{U}_i, \widehat{\Lambda}_0, t) = \left(\frac{\widehat{\Lambda}_0(Y_i)}{m_i} \sum_{j=1}^{m_i} K_{Y_i,2}(\frac{t - T_{ij}}{h_t}) \right).$$

The weight function $K_{y,l}(\frac{t-u}{h_t}) = \frac{1}{h}\alpha_l(y, \frac{t-u}{h_t})K(\frac{t-u}{h_t})$ in (3) is the lth order boundary kernel function of Gasser and Müller (1978) with adjustment for the boundary time y which, for $1 < m < l$, satisfies

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

where h_t is a positive valued bandwidth, which may be selected on the local or global basis, $K(\cdot)$ is a kernel density, and $\beta_{j,l}(t, h_t, y) = \int_{\frac{(t-s)}{h_t}}^{\frac{t}{h_t}} u^j \alpha_l(y, u) K(u) du$. In practical implementation, $\alpha_l(y, u)$ is often assigned to be the lth order polynomial function of u. When the information of covariates is not used in the informative censoring model, the estimator $\widehat{\Lambda}_0(t)$ in (3) can be computed by the explicit expression

$$\widehat{\Lambda}_0(t) = \left(\frac{1}{n} \sum_{l=1}^{n} \frac{m_l}{\widehat{F}(Y_l)} \right) \widehat{F}(t).$$

Another estimator for $\lambda_0(t)$ can also be obtained via directly smoothing $\widehat{\Lambda}_0(t)$ as below.

$$\widehat{\lambda}_{h_t,2}(t) = \int_0^{T_0} K_{T_0,2}(\frac{t - u}{h_t}) d\widehat{\Lambda}_0(u).$$

From the theoretical derivation and numerical studies, we found that there is no apparent difference between $\tilde{\lambda}_{h_t,2}(t)$ and $\hat{\lambda}_{h_t,2}(t)$. However, in the development of bootstrap procedures, it is easy to see that $\tilde{\lambda}_{h_t,2}(t)$ possesses an important computational advantage.

This advantage leads our focus on the first smoothing estimator $\tilde{\lambda}_{h_t,2}(t)$ in the succeeding discussion. When the following conditions

(A5) $\lambda_0(t)$ is twice differentiable and bounded,

(A6) the kernel function $K(u)$ is continuous, bounded, and satisfies $\gamma_4(t, h_t, y) < \infty$ with $\gamma_l(t, h_t, y) = \int_{\frac{(t-y)}{h_t}}^{\frac{t}{h_t}} |\alpha_l(y, u)K(u)|^l du$, and

(A7) $h_t = n^{-1/5} h_{t_0}$ for some bounded positive constant h_{t_0},

are further satisfied, Chiang and Wang (2003) derived the asymptotic normality of $(\tilde{\lambda}_{h_t,2}(t) - \lambda_0(t))$ as below.

$$\sup_u |P\left(\frac{(nh_t)^{1/2}(\tilde{\lambda}_{h_t,2}(t) - \lambda_0(t) - b_{h_t}(t))}{\sigma_\lambda(t)} \leq u\right) - \Phi(u)| \rightarrow 0, \ as\ n \rightarrow \infty, \tag{4}$$

where $\Phi(\cdot)$ denotes the cumulative distribution function of the standard normal,

$$b_{h_t}(t) = \frac{\lambda_0^{(2)}(t)(\int_z \int \int_{\{k \geq 1, y \geq t\}} \beta_{2,2}(t, h_t, y) dP_{mYZ}(k, y, z))h_t^2}{2\mu_\delta(t)},$$

and

$$\sigma_\lambda^2(t) = \frac{\lambda_0(t)(\int_z \int \int_{\{k \geq 1, y \geq t\}} (\gamma_2(t, h_t, y)\Lambda_0(y)/k) dP_{mYZ}(k, y, z))}{\mu_\delta^2(t)}$$

with $\mu_\delta(t) = E[\delta_1(t)]$. From the articles of Wang, Qin and Chiang (2001) and (4), we find that the moments of the proposed estimators are all influenced by a latent variable Z. It becomes difficult to directly estimate the unknown quantities in the asymptotic distributions of the estimators. In the following sections, we will discuss possible bootstrap methods to overcome this difficulty.

3 Random Weighted Bootstrap

A natural approach following the work of Efron (1979) is simply to resample the data with replacement and create bootstrap analogues of $\widehat{\boldsymbol{\beta}}$ and $\tilde{\lambda}_{h_t,2}(t)$. Since the random vectors

$\mathbf{U}_1, \cdots, \mathbf{U}_n$ are assumed to be independent across subjects, it seems reasonable to draw independent bootstrap random vectors $\mathbf{U}_1^b, \cdots, \mathbf{U}_n^b$ from the empirical distribution

$$P_{n,\mathbf{U}} = \frac{1}{n} \sum_{i=1}^{n} I_{\mathbf{U}_i}$$

and use the bootstrap sample $\left\{ \mathbf{U}_i^b : \mathbf{U}_i^b = \left((T_{i1}^b, \cdots, T_{im_i^b}^b), m_i^b, Y_i^b, \mathbf{X}_i^b \right) \right\}_{i=1}^{n}$ to play the role of the $\mathbf{U}_i$'s to form bootstrap analogues of $\widehat{\boldsymbol{\beta}}$ and $\widetilde{\lambda}_{h_t,2}(t)$. As we stressed in the introduction, another problem concerns practical issues involving implementation and what we feel to be computationally inefficient. This is essentially due to the high dimension of the data and the use of the resampling mechanism to rebuild bootstrap replicates at each stage. To avoid the existing difficulty in the naive bootstrap procedure, we propose general weighted bootstrap approximations for the sampling quantities of interest related to the estimators $\widehat{\boldsymbol{\beta}}$ and $\tilde{\lambda}_{h_t,2}(t)$. In addition, we identify a specific weighting scheme which is simpler and much faster than the naive bootstrap approach considered above. The proposed method involves a novel yet very simple approximation to the unknown quantities related to the estimators but does not require drawing observations or "really" generating random weights.

Let $\{\mathcal{D}_n, n > 1\}$ be a sequence of distribution functions which may depend on the sample and define a triangular array $\{D_{i:n} : 1 \leq i \leq n\}$ of random variables such that each row is an independent and identically distributed vector with distribution $\mathcal{D}_n$. Furthermore, let $n\bar{D}_n = D_{1:n} + ... + D_{n:n}$, $W_{i:n} = D_{i:n}/(n\bar{D}_n)$, and $\rho = E[D]/(V[D])^{1/2}$ be a scale factor modification which occurs due to variability in the weights. One can define a random weighted bootstrap empirical measure, for fixed $\{\mathbf{U}_1, \cdots, \mathbf{U}_n\}$, as

$$P_{n,\mathbf{U}}^{w} = \sum_{i=1}^{n} W_{i:n} I_{\mathbf{U}_i}.$$

An algorithm for the general weighted bootstrap procedures is then based on simulated independent non-negative exchangeable weights, and holds the data fixed rather than resampling the data and recomputing the entire statistics at each iteration. In this study, we choose a Bayesian Bootstrap clone (BBC) procedure of Lo (1991) with $D_i = \Gamma_i$ being independent and identically distributed $Gamma(4, 2)$ random variables. It is known that ρ is equal to 2 in this case and the choice of Γ_i has second order performance that is equivalent to Efron's bootstrap in terms of Edgeworth expansions in case of the sample mean in the complete data situation of Weng (1989) and Lo (1993), and for the Kaplan-Meier and cumulative hazard estimators for censored data (cf. James (1997)). Second order means that one is doing better than simply using a normal approximation and was one of the first of many victories for the bootstrap over the normal approximation (see Singh (1981)). Naturally, there are other possibilities for D_i. A general guideline is to choose D_i's with skewness 1, which is approximately the same as Efron's bootstrap based on resampling. For the general practioner, Efron's bootstrap is the most frequently used and understood method. However, if the practioner then considers Efron's bootstrap as one among many essentially equivalent methods, then the question should now become that of efficient practical implementation. This issue becomes more and more important as we begin to analyze increasingly complex statistical models and data settings. The answer is that indeed there are schemes which do outperform Efron's procedure in implementation. We highlight one such method here based on the BBC-type weights.

Without drawing bootstrap data as the naive bootstrap method, our random weighted bootstrap estimators are obtained by fixing $\{\mathbf{U}_1, \cdots, \mathbf{U}_n\}$ and replacing $\frac{1}{n}$ by BBC-type weights $W_{i:n}$'s with $D_i = \Gamma_i$'s being independent and identically distributed $Gamma(4, 2)$ random variables. Our random weighted bootstrap analogues, say, $\widehat{\beta}_0^{rwb}$ and $\widehat{\boldsymbol{\beta}}^{rwb}$ of $\widehat{\beta}_0$ and $\widehat{\boldsymbol{\beta}}$ are obtained from the following random weighted bootstrap estimating equations

$$\sum_{i=1}^{n} W_{i:n}(1, \mathbf{X}_i^T)^T \left(\frac{m_i}{\widehat{F}^{rwb}(t)} - \beta_0 \exp(\boldsymbol{\beta}^T \mathbf{X}_i) \right) \exp(\boldsymbol{\beta}^T \mathbf{X}_i) = 0, \tag{5}$$

where

$$\widehat{F}^{rwb}(t)=\prod_{\{s_{(l)}>t\}}\left(1-\frac{d_{(l)}^{rwb}}{N_{(l)}^{rwb}}\right)$$

with $d_{(l)}^{rwb}=\sum_{i=1}^{n}W_{i:n}\sum_{j=1}^{m_i}I(T_{ij}=s_{(l)})$ and $N_{(l)}^{rwb}=\sum_{i=1}^{n}W_{i:n}\sum_{j=1}^{m_i}I(T_{ij}\leq s_{(l)}\leq Y_i)$. Conditioning on $\{\mathbf{U}_1,\cdots,\mathbf{U}_n\}$, $\widehat{F}^{rwb}(t)$ is derived to be

$$\widehat{F}^{rwb}(t)=\widehat{F}(t)(1+\sum_{i=1}^{n}(W_{i:n}-\frac{1}{n})\widehat{b}_i(t))+o_{p^*}(n^{-1/2}), \tag{6}$$

where $P^*(\cdot)$ represents the probability measure conditioning on $\{\mathbf{U}_1,\cdots,\mathbf{U}_n\}$, $\bar{\Gamma}_n=n^{-1}\sum_{i=1}^{n}\Gamma_i$ and

$$\widehat{b}_i(t)=\sum_{j=1}^{m_i}\left\{\int_t^{T_0}\frac{I(T_{ij}\leq u\leq Y_i)d\widehat{Q}(u)}{\widehat{R}^2(u)}-\frac{I(t<T_{ij}\leq T_0)}{\widehat{R}(T_{ij})}\right\}$$

with $\widehat{Q}(u)=n^{-1}\sum_{i=1}^{n}\sum_{j=1}^{m_i}I(T_{ij}\leq u)$ and $\widehat{R}(u)=n^{-1}\sum_{i=1}^{n}\sum_{j=1}^{m_i}I(T_{ij}\leq u\leq Y_i)$. Since the vector of $W_{i:n}=\Gamma_i/\sum_{j=1}^{n}\Gamma_j$'s follows a Dirichlet distribution with parameters $(4,\cdots,4)$, it is straightforward to show that $E^*[\widehat{F}^{rwb}(t)]\doteq\widehat{F}(t)$ and $V^*[\widehat{F}^{rwb}(t)]\doteq\widehat{F}^2(t)\widehat{\sigma}_b^2/4n$, where $E^*[\cdot]$ and $V^*[\cdot]$ denote the expectation and variance conditioning on $\{\mathbf{U}_1,\cdots,\mathbf{U}_n\}$, and $\widehat{\sigma}_b^2=n^{-1}\sum_{i=1}^{n}\widehat{b}_i^2(t)$. By using the property of $\widehat{F}^{rwb}(t)$ in (6), we will show in the next section that the quantities $2(\widehat{\boldsymbol{\beta}}^{rwb}-\widehat{\boldsymbol{\beta}})$ are asymptotically equivalent to the sampling distribution of $(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta})$.

In this section, we propose one random weighted bootstrap analogue of $\widetilde{\lambda}_{h_t,2}(t)$ without using the random weighted bootstrap estimator $\widehat{\Lambda}_0^{rwb}(t)=\widehat{\beta}_0^{rwb}\widehat{F}^{rwb}(t)$ of $\widehat{\Lambda}_0(t)$ in $D_{n,h_t,l}(\mathbf{U}_i,\widehat{\Lambda}_0,t)$. Our random weighted bootstrap estimator is defined to be

$$\tilde{\lambda}_{h_t,2}^{rwb}(t)=\frac{\sum_{i=1}^{n}W_{i:n}\delta_i(t)D_{n,h_t,2}(\mathbf{U}_i,\widehat{\Lambda}_0,t)}{\sum_{j=1}^{n}W_{j:n}\delta_j(t)}. \tag{7}$$

Similar to the proof for the asymptotic behavior of $\tilde{\lambda}_{h_t,2}^{rwb}(t)$, we can show that the asymptotic distribution of another bootstrap analogue below is same with that of our random weighted bootstrap estimator.

$$\widehat{\lambda}_{h_t,2}^{rwb}(t)=\frac{\sum_{i=1}^{n}W_{i:n}\delta_i(t)D_{n,h_t,2}(\mathbf{U}_i,\widehat{\Lambda}_0^{rwb},t)}{\sum_{j=1}^{n}W_{j:n}\delta_j(t)}. \tag{8}$$

However, the estimator in (7) is computationally more efficient than that in (8) in practical implementation. Furthermore, without really generating random weights, it follows from the exchangeability of the weights that the explicit formulas for the conditional mean and variance of $\widetilde{\lambda}^{rwb}_{h_t,2}(t)$ can be derived to be $E^*[\widetilde{\lambda}^{rwb}_{h_t,2}(t)] = \widetilde{\lambda}_{h_t,2}(t)$ and $V^*[\widetilde{\lambda}^{rwb\Gamma}_{h_t,2}(t)] \doteq \frac{1}{4\delta.(t)}\hat{\sigma}^2_\lambda(t;h_t)$ with

$$\hat{\sigma}^2_\lambda(t;h_t) = \frac{1}{\delta.(t)+1/4}\left[\sum_{i=1}^{n}\delta_i(t)\left(D_{n,h_t,2}(\mathbf{U}_i,\widehat{\Lambda}_0,t) - \frac{\sum_{j=1}^{n}\delta_j(t)D_{n,h_t,2}(\mathbf{U}_j,\widehat{\Lambda}_0,t)}{\delta.(t)}\right)^2\right].$$

To those who are familiar with the literature on bootstrap curve estimation, the bootstrap estimator of $\widetilde{\lambda}_{h_t,2}(t)$ fails to mimic the non-negligible bias $b_{h_t}(t)$ of $\widetilde{\lambda}_{h_t,2}(t)$ with respect to $\lambda_0(t)$. To remedy this problem, the bias correction of Schucany (1995) can be extended in this case by using the difference between $\widetilde{\lambda}_{h_t,2}(t)$ and $\widetilde{\lambda}_{h_t,4}(t)$, which is computed via using the fourth order boundary kernel function $K_{y,4}(\frac{t-u}{h_t})$ in (3), to estimate the parameter function $b_{h_t}(t)$. Another remedy for this problem is known as a wild bootstrap. It was shown that, conditioning on $(m_i, y_i, z_i, \mathbf{x}_i)$, the unordered observations of T_{ij}'s are independent and identically distributed with density function $f_i(t)$. Thus, the truncated density function $f_i(t)$ can be naturally estimated by

$$\widetilde{f}_{i,g_t}(t) = \widetilde{\lambda}_{g_t,2}(t)/\widehat{\Lambda}_0(Y_i).$$

Here, the bandwidth g_t is set to be greater than h_t. The wild bootstrap procedure is implemented, for fixed $(m_i, Y_i, \mathbf{X}_i)$ in $\mathbf{U}_i$, by drawing observations $\{(T^{wb}_{ij})\}_{1\le j\le m_i}$ from $\widetilde{f}_{i,g_t}(t)$, and using the bootstrap values $\mathbf{U}^{wb}_i = \Big((T^{wb}_{i1},\cdots,T^{wb}_{im_i}), m_i, Y_i, \mathbf{X}_i\Big)$, $i = 1,\cdots,n$, to form the wild bootstrap estimator $\widetilde{\lambda}_{h_t,2}(t)$.

$$\widetilde{\lambda}^{wb}_{h_t,2}(t) = \frac{\sum_{i=1}^{n}\frac{1}{n}D_{n,h_t,2}(\mathbf{U}^{wb}_i,\widehat{\Lambda}_0,t)}{\delta.(t)}.$$

We find that the wild bootstrap procedure still exists the problem of computational efficiency in our data setting. Thus, the extension of Schucany (1995) for bias adjustment without drawing observations from the data is suggested in this article.

4 Random Weighted Bootstrap Procedure

In this section, the distributional behavior of the proposed random weighted bootstrap estimators $\tilde{\lambda}^{rwb}_{h_t,2}(t)$ and $\widehat{\boldsymbol{\beta}}^{rwb}$ will be used to construct the approximated confidence intervals for the baseline rate function $\lambda_0(t)$ and the regression parameters $\boldsymbol{\beta}$. In addition, the random weighted bootstrap estimator of the mean squared errors of $\tilde{\lambda}_{h_t,2}(t)$ will be proposed and suggested to select bandwidths.

4.1 Weighted Bootstrap Confidence Intervals

Without relying on the asymptotic distributions, the approximated random weighted bootstrap confidence intervals can be constructed by the following steps.

Step 1. Simultaneously generate B times $\mathbf{W}_n = (W_{1:n}, \cdots, W_{n:n})$ based on independent and identically distributed $Gamma(4, 2)$ random variables.

Step 2. Compute $\widehat{\boldsymbol{\beta}}$, $\tilde{\lambda}_{h,l}(t)$, l=2, 4, and B weighted bootstrap estimators $\widehat{\boldsymbol{\beta}}^{rwb}$ and $\tilde{\lambda}^{rwb}_{h_t,2}(t)$ based on (2), (3), (5), (7) and $\mathbf{W_n}$ generated in Step 1.

Step 3. Construct an approximated $(1-\alpha)$ random weighted bootstrap confidence regions for β_l, $l = 1, \cdots p$, and $\lambda_0(t)$ separately by

$$\widehat{\beta}_l + 2\left(U^*_{\alpha/2}(\widehat{\beta}^{rwb}_l - \widehat{\beta}_l), L^*_{\alpha/2}(\widehat{\beta}^{rwb}_l - \widehat{\beta}_l)\right) \tag{9}$$

and

$$(\tilde{\lambda}_{h_t,2}(t) - \widehat{b}_{h_t}(t)) + 2\left(U^*_{\alpha/2}(\tilde{\lambda}^{rwb}_{h_t,2}(t) - \tilde{\lambda}_{h_t,2}(t)), L^*_{\alpha/2}(\tilde{\lambda}^{rwb}_{h_t,2}(t) - \tilde{\lambda}_{h_t,2}(t))\right), \tag{10}$$

where $L^*_{\alpha/2}(\cdot)$ and $U^*_{\alpha/2}(\cdot)$ denote the $100(\alpha/2)$th and $100(1-\alpha/2)$th percentiles of B weighted bootstrap estimators.

An alternative to the above random weighted bootstrap percentile intervals in Step 3, one can also use the bootstrap intervals of the forms

$$\widehat{\beta}_l \pm 2z_{(1-\alpha/2)}se^*(\widehat{\beta}_l^{rwb}) \tag{11}$$

and

$$\left(\tilde{\lambda}_{h_t,2}(t) - \widehat{b}_{h_t}(t)\right) \pm 2z_{(1-\alpha/2)}se^*(\widetilde{\lambda}^{rwb}_{h_t,2}(t)), \tag{12}$$

where $z_{(1-\alpha/2)}$ is the $(1-\alpha/2)$ quantile value of the standard normal distribution, and $se^*(\cdot)$ is the standard error of B estimators computed in Step 2. The validity of the above normal approximated confidence intervals can be verified by the asymptotic normalities of the estimators and the corresponding random weighted bootstrap estimators. Moreover, without really generating random weights $\mathbf{W}_n$, a simple and fast way is directly to construct the bootstrap confidence intervals for $\lambda_0(t)$ by

$$\left(\tilde{\lambda}_{h_t,2}(t) - \widehat{b}_{h_t}(t)\right) \pm z_{(1-\alpha/2)}\frac{\widehat{\sigma}_\lambda(t;h_t)}{(\delta.(t))^{1/2}}. \tag{13}$$

In the following two theorems, we derive the asymptotic distributions of $(\widehat{\boldsymbol{\beta}}^{rwb} - \widehat{\boldsymbol{\beta}})$ and $(\tilde{\lambda}^{rwb}_{h_t,2}(t) - \tilde{\lambda}_{h_t,2}(t))$ with the bias adjustment. Based on the distributional relation between the estimators and their random weighted bootstrap analogues, the approximated confidence regions are then constructed.

Theorem 1. Suppose that assumptions (A1)-(A4) and (A8) $E[|\mathbf{e}_{il_1}|^3] < \infty$ and $E[|\left(\frac{\partial \mathbf{e}_i}{\partial \mathbf{X}_i}\right)_{l_1 l_2}|^3] < \infty$ for $l_1, l_2 = 1, \cdots, p$, are satisfied. Then, conditioning on $\{\mathbf{U}_1, \cdots, \mathbf{U}_n\}$,

$$2n^{1/2}\left(\widehat{\boldsymbol{\beta}}^{rbw} - \widehat{\boldsymbol{\beta}}\right) \xrightarrow{d} N_p\left(\mathbf{0}, (\Psi^{-1}\Sigma(\Psi^T)^{-1})_{2:(p+1)\times 2:(p+1)}\right) \textit{ in probability, as } n \to \infty, \tag{14}$$

where $\mathbf{e}$, Ψ, and Σ are defined in the paper of Wang, Qin and Chiang (2002).

Proof. See Appendix.□

Theorem 2. Suppose that assumptions (A1)-(A7) are satisfied. Then, for all $u \in R$,

$$\sup_u |P^* \left(\frac{2(nh_t)^{1/2}(\tilde{\lambda}^{rwb}_{h_t,2}(t) - \tilde{\lambda}_{h_t,2}(t))}{\sigma_\lambda(t)} \le u \right) - \Phi(u)| \xrightarrow{p} 0, \ as\ n \to \infty \tag{15}$$

and

$$(nh_t)^{1/2} \left(\widehat{b}_{h_t}(t) - b_{h_t}(t) \right) \xrightarrow{p} 0, \ as\ n \to \infty. \tag{16}$$

Proof. See Appendix.□

The property in (16) can be obtained by verifying the following two properties:

$$(nh_t)^{1/2} \left(\widetilde{\lambda}_{h_t,2}(t) - \lambda_0(t) - b_{h_t}(t) \right) \xrightarrow{p} 0, \tag{17}$$

and

$$(nh_t)^{1/2} \left(\widetilde{\lambda}_{h_t,4}(t) - \lambda_0(t) \right) \xrightarrow{p} 0. \tag{18}$$

The proofs for (17) and (18) are along the same lines as the proof in Theorem 1.

4.2 Bandwidth Selection - Baseline Rate Estimator

In kernel estimation, the selection of bandwidths is usually more important than the selection of kernel functions (cf. Silverman (1986) and Härdle (1990)). Although the appropriate bandwidths may be selected subjectively by examining the plots of the fitted curves, there is still no standard criterion for doing so. Thus, an appropriate automatic procedure urgently becomes necessary both for simplifying the process and preventing the investigator from arbitrarily driving the process.

To select an appropriate global bandwidth for the kernel estimator $\tilde{\lambda}_{h_t,2}(t)$, it is reasonable to extend the criterion of a least squares cross-validation to a "leave one subject out" least squares cross-validation criterion. This criterion is similar to the intuitive procedure suggested by Rice and Silverman (1991). Generally speaking, we are interested in finding a bandwidth, say, h_{cv} that minimizes the cross-validation score

$$CV(h) = \int_{\{y\geq 0\}} (\int_0^y \tilde{\lambda}_{h,2}^2(u)du)d\widehat{F}_Y(y) - \frac{2}{n}\sum_{i=1}^{n} \frac{\widehat{\Lambda}_{0(-i)}(Y_i)}{m_i} \sum_{j=1}^{m_i} \tilde{\lambda}_{h,2,(-i)}(t_{ij}), \tag{19}$$

where

$$\widehat{\lambda}_{h,2,(-i)}(t) = \sum_{l_1\neq i} \frac{\delta_{l_1}(t)}{\delta_{\cdot(-i)}(t)} (\frac{\widehat{\Lambda}_{0(-i)}(Y_{l_1})}{m_i} \sum_{l_2=1}^{m_{l_1}} K_{Y_{l_1},2}(\frac{t-t_{l_1l_2}}{h})),$$

$\widehat{\Lambda}_{0(-i)}(t)$ is computed as $\widehat{\Lambda}_0(t)$ with the ith subject being deleted, $\delta_{\cdot(-i)}(t) = \sum_{l_1\neq i}\delta_{l_1}(t)$, and $\widehat{F}_Y(y)$ is the empirical estimator of $F_Y(y)$. One can show that $CV(h)+\int(\int_0^y \lambda_0^2(u)du)dF_Y(y)$ is an asymptotically unbiased estimator of the integrated mean squared error (IMSE) of $\widetilde{\lambda}_{h,2}(t)$,

$$IMSE(\tilde{\lambda}_{h,2}) = \int_{\{y\geq 0\}} (\int_0^y E[(\tilde{\lambda}_{h,2}(u) - \lambda_0(u))^2]du)dF_Y(y).$$

Since the second term is independent of the choice of the bandwidth h, the minimizer h_{cv} of $CV(h)$ will approximately minimize $IMSE(\tilde{\lambda}_{h,2})$.

Although a cross-validation criterion is an acceptable bandwidth selection method, simulation experiments indicate that the bandwidth selected from this criterion is often inadequate and computationally inefficient. Furthermore, the cross-validation bandwidth cannot adapt to the local behavior of the estimated curve since it is selected on the global basis. As we mentioned before, we are dealing with the problem with a latent variable Z and a higher-dimensional vector $\mathbf{U}$, one can use the random weighted bootstrap approach proposed here for the purpose of bandwidth selection. Since the bias and the variance of the occurrence rate estimator $\tilde{\lambda}_{h_t,2}(t)$ can be estimated by the estimators $\widehat{b}_{h_t}(t)h_t^2$ and $\widehat{\sigma}_\lambda^2(t;h_t)/\delta.(t)$, it is reasonable to approximate $MSE(\tilde{\lambda}_{h_t,2}(t))$ by

$$MSE^*(\tilde{\lambda}_{h_t,2}(t)) = \widehat{b}_{h_t}^2(t) + \frac{\widehat{\sigma}_\lambda^2(t;h_t)}{\delta.(t)}. \tag{20}$$

The local bandwidth estimator, say, h_t^{rwb} at time t is then defined to be the minimizer of $MSE^*(t)$. Similarly, the global bandwidth estimator, say, h^{rwb} can be obtained by minimizing $IMSE^*(\tilde{\lambda}_{h,2}) = \int_0^{T_0} MSE^*(\tilde{\lambda}_{h,2}(u))du$.

5 Numerical Studies

To examine the performance of the proposed random weighted bootstrap procedure, a series of Monte Carlo simulations are conducted. The proposed procedure is also applied to a recurrent event sample of intravenous drug users for inpatient cares over time.

5.1 Monte Carlo Simulations

In the simulation study, the recurrent event data are repeatedly generated 500 times from 400 independent non-stationary Poisson processes under two informative censoring models. The data generated in the simulation study of Wang, Qin and Chiang (2002) and the simulated data without using the information of covariates will be used separately to investigate the proposed random weighted bootstrap procedure for the regression parameters $\boldsymbol{\beta}$ and the rate function $\lambda_0(t)$.

In the first setting informative censoring model, the subject-specific rate function of each recurrent event process $N_i(t)$ is assumed to be $\lambda_i(t) = Z_i\lambda_0(t)\exp(X_i)$, where

$$\lambda_0(t) = 0.6 + \frac{(t-6)^3}{360}, \qquad t \in [0, 10],$$

the covariate X_i is specified to follow a Bernoulli distribution with a parameter 0.5, and the latent variables Z_i is defined as $Z_i = \exp(-X_i \ln(2.75))Z_i^*$ with Z_i^* being distributed as

$$f_{Z_i^*|x_i}(z^*) = (1 - x_i)I(0.5 \leq z^* \leq 1.5) + \frac{x_i}{2.5}I(1.5 \leq z^* \leq 4).$$

Conditioning on z_i, Y_i is designed to range from 1 to 10 and follow a truncated exponential distribution with a parameter $0.1z_i$. The second recurrent events are simulated with the subject-specific rate function $\lambda_i(t) = Z_i\lambda_0(t)$. The latent variable Z_i and the rate function

Table 1: The 0.95 quantile intervals (Q.I.) of estimates, the averages of the corresponding 0.95 random weighted bootstrap quantile confidence intervals (R.W.B.Q.C.I.) and random weighted bootstrap normal approximated confidence intervals (R.W.B.N.C.I.), and the empirical coverage probabilities.

Parameter	Q.I. of estimates	R.W.B.Q.C.I.	Coverage prob.
β_0	(4.2632,6.5859)	(4.1514,6.2713)	0.912
β	(0.8095,1.2009)	(0.8028,1.1950)	0.948

Parameter	Q.I. of estimates	R.W.B.N.C.I.	Coverage prob.
β_0	(4.2632,6.5859)	(4.2185,6.3567)	0.916
β	(0.8095,1.2009)	(0.8021,1.1972)	0.948

$\lambda_0(t)$ are separately set to be a uniform random variable with parameters $(0.4, 1.6)$ and

$$\lambda_0(t) = 2.5 + \frac{5(t-3)^3}{108}, \qquad t \in [0, 5].$$

The conditional distribution of Y_i, which ranges from 1 to 5, is also assumed to be a truncated exponential distribution with a parameter $0.1z_i$.

Based on 500 simulated data of the first informative censoring model, the averages of the estimates $\widehat{\beta}_0$ and $\widehat{\beta}_1$ are computed to be 5.2876 and 0.9996 with standard errors of 0.5883 and 0.1037. Same with the former study, the estimated values are close to the true values $\beta_0 = 5.278$ and $\beta = 1$. Table 1 provides the 0.95 quantile intervals of 500 simulated estimates, the 500 averages of the corresponding 0.95 weighted bootstrap confidence intervals, and the empirical coverage probabilities. From this table, we find that the 500 averages of the 0.95 random weighted bootstrap confidence intervals and the empirical coverage probabilities are all close to the expected intervals and the nominal level.

Under the second simulated model, the kernel estimator $\tilde{\lambda}_{h_t,2}(t)$ is computed by using the Gaussian kernel for $K(\cdot)$ with the weighted bootstrap local bandwidth h_t^{rwb}, which is the minimizer of $MSE^*(\tilde{\lambda}_{h_t,2}(t))$ in (20). In the construction of confidence intervals for $\lambda_0(t)$, three weighted bootstrap methods suggested in Section 4 are used. Figures (1a)-(1c) present the true occurrence rate function $\lambda_0(t)$, the 500 simulation averages of the estimated occurrence rate functions and the approximated 0.95 random weighted bootstrap confidence intervals of three methods for $\lambda_0(t)$, and the 0.95 percentile confidence intervals of the 500 simulation estimates at the corresponding time points.

From these figures, it is found that the 500 averages of the 0.95 percentile random weighted bootstrap confidence intervals are close to the 0.95 percentile confidence intervals of the 500 simulation estimates. Moreover, figure (1d) reveals that the coverage percentages of 500 simulations for our suggested approaches are very nearly the assigned nominal level. One can see that these three procedures have very similar results. In longitudinal studies, the sample size of collected recurrent events is usually large enough. For the sake of computational efficiency, Method 3 will be a good suggestion in the practical implementation.

5.2 A Data Example

The recurrent event data analyzed here are same with that used in Wang and Chiang (2002), which are collected from the AIDS Link to Intravenous Experiences (ALIVE) cohort study. Details of this study and the summary of repeated hospitalization records of intravenous drug users can be found in Vlahov, et al. (1991) and Wang and Chiang (2002).

Measurements for each patient in the following analysis include the dates of repeated impatient admissions from August 1, 1993, the last visit time, the human immune deficiency virus status, the racial indicator for black and non-black people, and the entering age group, which is defined to be 0 if he or she is younger than the median entering age and 1 otherwise. Patients who are HIV-negative and become HIV-positive during the study period are excluded from the analysis. Furthermore, since there is an extremely low proportion of non-black people among the patients in this data, the racial indicator will not be used in the informative censoring regression model. It indicates from the data that no significant association exists between entering age and HIV status.

Based on the considered informative censoring model, we implement our random weighted bootstrap procedure to detect the effects of entering age and HIV status on the repeated hospitalization process and to estimate the hospitalization rate curves. In the construction of confidence intervals, the approximated 0.95 random weighted bootstrap confidence intervals are established based on the quantile method. Same with the simulation study, the kernel estimator is computed by using the Gaussian kernel density and local weighted bootstrap bandwidths. The estimated regression effects for the older age group and HIV-positive patients are 0.051 and 0.4950 with the corresponding approximated 0.95 confidence intervals (-0.1384,0.3002) and (0.2869,0.6912). This result shows that no significant effect is detected for the entering age on the hospitalization process, while the hospitalization occurrence rate of HIV-positive drug users is significantly higher than that of HIV-negative ones. Finally, we provide the estimated occurrence rate curves with the approximated 0.95 random weighted bootstrap confidence intervals for the different entering age group and HIV status in Figures (2a)-(2b). It reveals from these figures that the frequencies of hospitalization tend to increase during the study.

REFERENCES

Chiang, C. T. and Wang, M. C. (2004). Smoothing Estimation of Rate Function for Recurrent Event Data with Informative Censoring. *Annals of the Institute of Statistical Mathematics* **56**, 87-100.

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.

Härdle, W. (1990). Applied nonparametric regression. Cambridge University Press, Cambridge, England.

James, L. F. (1997). A study of a class of weighted bootstraps for censored data. *Ann. Statist.* **25**, 1595-1621. *J. Amer. Statist. Assoc.* **91**, 401-407.

Lo, A. Y. (1987). A large sample study of the Bayesian bootstrap. *Ann. Statist.* **15**, 360-375.

Lo, A. Y. (1991). Bayesian bootstrap clones and a biometry function. *Sankhya Ser. A.* **53**, 320-333.

Lo, A. Y. (1993). A Bayesian method for weighted sampling. *Ann. Statist.* **21**, 2138-2148.

Mason, D. M. and Newton, M. (1992). A rank statistic approach to the consistency of a general bootstrap. *Ann. Statist.* **20**, 1611-1624.

Praestgaard, J. T. and Wellner, J. A. (1993). Exchangeably weighted bootstraps of the general empirical process. *Ann. Probab.* **21**, 2053-2086.

Rubin, D. (1981). The Bayesian bootstrap. *Ann. Statist.* **9**, 130-134.

Schucany, W. R. (1995). Adaptive bandwidth choice for kernel regression. *J. Amer. Statist. Assoc.* **90**, 535-540.

Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis.* London: Chapman and Hall.

Singh, K. (1981). On the asymptotic accuracy of Efron's bootstrap. *Ann. Statist.* **9**, 1187-1195.

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.

Wang, M. C. and Chiang, C. T. (2002). Nonparametric Methods for Recurrent Event Data with Informative and Non-informative Censorings. *Statistics in Medicine* **21**, 445-456.

Wang, M.-C., Qin, J. and Chiang, C.-T. (2001). Analyzing Recurrent Event Data with Informative Censoring. *Journal of the American Statistical Association* **96**, 1057-1065.

Weng, C. S. (1989). On a second-order property of the Bayesian bootstrap mean. *Ann. Statist.* **17**, 705-710.

APPENDIX

Proof of Theorem 1:

Let $\boldsymbol{\gamma}^T = (\ln(\beta_0), \boldsymbol{\beta}^T)^T$ and $\tilde{\mathbf{X}} = (1, \mathbf{X}^T)^T$. By using the fact that the random weighted bootstrap estimators $\boldsymbol{\gamma}^{rwb}$ satisfy the estimating equations (5) and the Taylor expansion, we have that

$$\sum_{i=1}^{n} W_{i:n}\tilde{\mathbf{X}}_i \left(\frac{m_i}{\widehat{F}^{rwb}(Y_i)} - \exp(\widehat{\boldsymbol{\gamma}}^T\tilde{\mathbf{X}}_i)\right)\exp(\widehat{\boldsymbol{\gamma}}^T\tilde{\mathbf{X}}_i) \tag{A.1}$$

$$= \quad -\sum_{i=1}^{n} W_{i:n}\tilde{\mathbf{X}}_i\tilde{\mathbf{X}}_i^T \left(\frac{m_i}{\widehat{F}^{rwb}(Y_i)} - 2\exp(\widehat{\boldsymbol{\gamma}}^T\tilde{\mathbf{X}}_i)\right)\exp(\widehat{\boldsymbol{\gamma}}^T\tilde{\mathbf{X}}_i)\left(\widehat{\boldsymbol{\gamma}}^{rwbT} - \widehat{\boldsymbol{\gamma}}\right) + o_{p^*}(n^{-1/2})$$

From (2), (6), the left side of (A.1) can be re-expressed as

$$\sum_{i=1}^{n} W_{i:n}\tilde{\mathbf{X}}_i \left(\frac{m_i}{\widehat{F}^{rwb}(Y_i)} - \exp(\widehat{\boldsymbol{\gamma}}^T\tilde{\mathbf{X}}_i)\right)\exp(\widehat{\boldsymbol{\gamma}}^T\tilde{\mathbf{X}}_i)$$

$$= \quad \sum_{i=1}^{n}(W_{i:n} - \frac{1}{n})\left(\widehat{\mathbf{c}}_i + \tilde{\mathbf{X}}_i\left(\frac{m_i}{\widehat{F}(Y_i)} - \exp(\widehat{\boldsymbol{\gamma}}^T\tilde{\mathbf{X}}_i)\right)\exp(\widehat{\boldsymbol{\gamma}}^T\tilde{\mathbf{X}}_i)\right), \tag{A.2}$$

where $\widehat{\mathbf{c}}_i = -\sum_{l=1}^{n} W_{l:n}\tilde{\mathbf{X}}_l m_l \widehat{b}_i(Y_l)\exp(\widehat{\boldsymbol{\gamma}}^T\tilde{\mathbf{X}}_l)/\widehat{F}(Y_l)$. Since the vector of $W_{i:n}$'s follows a Dirichlet distribution with parameters $(4, \cdots, 4)$, it implies that

$$\sum_{i=1}^{n}(W_{i:n} - \frac{1}{n})\widehat{\mathbf{c}}_i = \sum_{i=1}^{n}(W_{i:n} - \frac{1}{n})\left(\frac{-1}{n}\sum_{l=1}^{n}\frac{\tilde{\mathbf{X}}_l m_l \widehat{b}_i(Y_l)\exp(\widehat{\boldsymbol{\gamma}}^T\tilde{\mathbf{X}}_l)}{\widehat{F}(Y_l)}\right) + o_{p^*}(n^{-1/2}), \tag{A.3}$$

and

$$\sum_{i=1}^{n} W_{i:n}\tilde{\mathbf{X}}_i \left(\frac{m_i}{\widehat{F}^{rwb}(Y_i)} - \exp(\widehat{\boldsymbol{\gamma}}^T\tilde{\mathbf{X}}_i)\right)\exp(\widehat{\boldsymbol{\gamma}}^T\tilde{\mathbf{X}}_i)$$

$$= \quad \sum_{i=1}^{n} W_{i:n}\widehat{\mathbf{e}}_i + o_{p^*}(n^{-1/2}) = (\frac{1}{n}\sum_{i=1}^{n}\frac{\Gamma_i\widehat{\mathbf{e}}_i}{2})(1 + O_{p^*}(n^{-1/2})), \tag{A.4}$$

where

$$\widehat{\mathbf{e}}_i = \frac{-1}{n}\sum_{l=1}^{n}\left(\frac{\tilde{\mathbf{X}}_l m_l \widehat{b}_i(Y_l)\exp(\widehat{\boldsymbol{\gamma}}^T\tilde{\mathbf{X}}_l)}{\widehat{F}(Y_l)}\right) + \tilde{\mathbf{X}}_i\left(\frac{m_i}{\widehat{F}(t)} - \exp(\widehat{\boldsymbol{\gamma}}^T\tilde{\mathbf{X}}_i)\right)\exp(\widehat{\boldsymbol{\gamma}}^T\tilde{\mathbf{X}}_i)$$

with $\sum_{i=1}^{n} \widehat{\mathbf{e}}_i = 0$. From (A.1), (A.4), and

$$\sum_{i=1}^{n} W_{i:n} \tilde{\mathbf{X}}_i \tilde{\mathbf{X}}_i^T \left(\frac{m_i}{\widehat{F}^{rwb}(t)} - 2\exp(\widehat{\boldsymbol{\gamma}}^T \tilde{\mathbf{X}}_i) \right) \exp(\widehat{\boldsymbol{\gamma}}^T \tilde{\mathbf{X}}_i) = \widehat{\Psi}(1 + O_{p^*}(n^{-1/2})), \qquad (A.5)$$

where $\widehat{\Psi} = n^{-1} \sum_{i=1}^{n} \tilde{\mathbf{X}}_i \tilde{\mathbf{X}}_i^T (m_i/\widehat{F}(t) - 2\exp(\widehat{\boldsymbol{\gamma}}^T \tilde{\mathbf{X}}_i)) \exp(\widehat{\boldsymbol{\gamma}}^T \tilde{\mathbf{X}}_i)$, we can derive that

$$n^{1/2} \left(\widehat{\boldsymbol{\gamma}}^{rwbT} - \widehat{\boldsymbol{\gamma}} \right) = (\frac{1}{n^{1/2}} \sum_{i=1}^{n} \frac{\Gamma_i \widehat{\Psi}^{-1} \widehat{\mathbf{e}}_i}{2})(1 + O_{p^*}(n^{-1/2})). \qquad (A.6)$$

Paralleling the proof of Wang, Qin and Chiang (2001), the estimators $\widehat{\boldsymbol{\gamma}}$ can be shown to converge to $\boldsymbol{\gamma}$ in probability. By assumption (A8), $\widehat{\Psi} \xrightarrow{p} \Psi$, and $n^{-1} \sum_{i=1}^{n} \widehat{\mathbf{e}}_i \widehat{\mathbf{e}}_i^T \xrightarrow{p} \Sigma$, it can be derived by using the Berry-Esséen theorem that, for any $(p+1) \times 1$ constant vector $\mathbf{a}$,

$$\sup_u |P^* \left(\frac{\frac{1}{2^{1/2}} \sum_{i=1}^{n} \left(\frac{\Gamma_i}{2}\right) \left(\mathbf{a}^T \widehat{\mathbf{e}}_i\right)}{2(\mathbf{a}^T \Psi^{-1} \Sigma (\Psi^T)^{-1} \mathbf{a})^{1/2}} \leq u \right) - \Phi(u)| \xrightarrow{p} 0, \qquad as\ n \to \infty. \qquad (A.7)$$

From (A.6), (A.7), and the Cramér-Wold device, Theorem 1 is then obtained.

Proof of Theorem 2:

From (3) and (7), we obtain that

$$(nh_t)^{1/2} \left(\tilde{\lambda}_{h_t,2}^{rwb}(t) - \tilde{\lambda}_{h_t,2}(t) \right) \qquad (A.8)$$

$$= \frac{(nh_t)^{1/2}}{\bar{\delta}^{rwb}(t)} \sum_{i=1}^{n} W_{i:n} \delta_i(t) D_{n,h_t,2}(\mathbf{U}_i, \widehat{\Lambda}_0, t) + \frac{(nh_t)^{1/2}}{\bar{\delta}(t)} \sum_{i=1}^{n} \frac{1}{n} \delta_i(t) D_{n,h_t,2}(\mathbf{U}_i, \widehat{\Lambda}_0, t),$$

where $\bar{\delta}^{rwb}(t) = \sum_{i=1}^{n} W_{i:n} \delta_i(t)$ and $\bar{\delta}(t) = \frac{1}{n} \sum_{i=1}^{n} \delta_i(t)$. By the law of large numbers, one gets

$$\bar{\delta}(t) \xrightarrow{p} \mu_\delta(t). \qquad (A.9)$$

Moreover, using the properties $n \sum_{i=1}^{n} (W_{i:n} - 1/n)^2 \xrightarrow{p} 1/4$ and $\delta_i(t) \leq 1$, it implies that

$$|\sum_{i=1}^{n} (W_{i:n} - \frac{1}{n}) \delta_i(t)| \leq \frac{1}{n^{1/2}} \left(n \sum_{i=1}^{n} (W_{i:n} - \frac{1}{n})^2 \right)^{1/2} \xrightarrow{p} 0. \qquad (A.10)$$

Thus, from (A.9) and (A.10), we obtain that

$$\bar{\delta}^{rwb}(t) \xrightarrow{p} \mu_\delta(t). \qquad (A.11)$$

By the representation $\widehat{\Lambda}_0(t) = \Lambda_0(t)(1 + n^{-1}\sum_{i=1}^n d_i(t) + o_p(n^{-1/2}))$ of Wang, Qin and Chiang (2001), (A.9) and (A.10), (A.8) can be re-expressed as

$$(nh_t)^{1/2}\left(\tilde{\lambda}^{rwb}_{h_t,2}(t) - \tilde{\lambda}_{h_t,2}(t)\right) \tag{A.12}$$

$$= \frac{(nh_t)^{1/2}}{\mu_\delta(t)}(\sum_{i=1}^n (W_{i:n} - \frac{1}{n})\delta_i(t) D_{n,h_t,2}(\mathbf{U}_i, \Lambda_0, t)(1 + n^{-1}\bar{d}(Y_i) + o_p(n^{-1/2}))(1 + o_p(1)),$$

where $\bar{d}(t) = n^{-1}\sum_i^n d_i(t)$. By Lemma 1 of Chiang and Wang (2001) and the finite second moment of $d_i(t)$, one can derive that

$$(nh_t)^{1/2}\sum_{i=1}^n (W_{i:n} - \frac{1}{n})\delta_i(t) D_{n,h_t,2}(\mathbf{U}_i, \Lambda_0, t) n^{-1}\bar{d}(Y_i)$$

$$= (n^{-1}h_t)^{1/2}\sum_{i=1}^n (\frac{\Gamma_i}{2} - 1)\delta_i(t) D_{n,h_t,2}(\mathbf{U}_i, \Lambda_0, t)\bar{d}(Y_i)$$

$$\leq n^{-1/2}\left(\frac{1}{n}\sum_{i=1}^n (\frac{\Gamma_i}{2} - 1)^2 \delta_i(t) h_t D^2_{n,h_t,2}(\mathbf{U}_i, \Lambda_0, t)\right)^{-1/2}\left(n\sum_{i=1}^n \bar{d}^2(Y_i)\right)^{-1/2}$$

$$\xrightarrow{p} 0. \tag{A.13}$$

From (A.12) and (A.13), we have

$$(nh_t)^{1/2}\left(\frac{1}{n}\sum_{i=1}^n \xi_i(t) - \tilde{\lambda}_{h_t,2}(t)\right)$$

$$= \frac{(nh_t)^{1/2}}{\mu_\delta(t)}\left(\sum_{i=1}^n (W_{i:n} - \frac{1}{n})\delta_i(t) D_{n,h_t,2}(\mathbf{U}_i, \Lambda_0, t)\right)(1 + o_p(1)). \tag{A.14}$$

It follows immediately from Mason and Newton (1992) or Praestgaard and Wellner (1993) that, conditioning on $\{\mathbf{U}_1, \cdots, \mathbf{U}_n\}$,

$$\sup_u |P^*\left(\frac{(2n)^{1/2}\sum_{i=1}^n (W_{i:n} - \frac{1}{n}) h_t^{1/2}\xi_i(t)}{\sigma_\lambda(t)} \leq u\right) - \Phi(u)| \xrightarrow{p} 0, \ \ as\ n \to \infty. \tag{A.15}$$

Finally, (15) can be obtained directly from (A.14) and (A.15).