Biometrical Journal **46** (2004) 5, 503–511 DOI: 10.1002/bimj.200310060

An Independence Test for Doubly Censored Failure Time Data

Jianguo Sun[*, 1], **Hee-Jeong Lim**[2], and **Xingqiu Zhao**[3]

[1] Department of Statistics, University of Missouri, Columbia, Missouri, USA 65211
[2] Department of Mathematics and Computer Science, Northern Kentucky University, Highland Heights, KY, USA 41099
[3] Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton, Alberta, Canada T6G 2G1

Received 27 March 2003, revised 14 July 2004, accepted 3 August 2004

Summary

The analysis of doubly censored failure time data has recently attracted a great deal of attention and for this, a number of methods have been proposed (De Gruttola and Lagakos, 1989; Kim et al., 1993; Pan, 2001; Sun, 2004). To simplify the analysis, most of these methods make an independence assumption: the distribution of the survival time of interest is independent of the occurrence of the initial event that defines the survival time. Although it is well-known that the assumption may not be true, there does not seem to be any existing research discussing the checking of the assumption. In this article, a Wald test is developed for testing this assumption and the method is applied to an AIDS cohort study.

Key words: Estimating equation; Interval-censoring; Proportional hazards model; Regression analysis.

1 Introduction

This paper discusses the analysis of doubly censored failure time data (De Gruttola and Lagakos, 1989; Sun, 2004). By doubly censored failure time data, we mean that the survival time of interest is defined as the elapsed time between two related events, called initial and subsequent events. Furthermore, observations on the occurrences of both events could be right- or interval-censored. A well-known example of such data arises from follow-up studies of patients who have been or are at risk of being infected by the human immunodeficiency virus (HIV) and thus are also at risk of developing the acquired immune deficiency syndrome (AIDS) (De Gruttola and Lagakos, 1989; Joly and Commenges, 1999).

Doubly censored failure time data include usual right- and interval-censored failure time data as special cases (Kalbfleisch and Prentice, 1980; Sun, 1998). For example, they reduce to interval-censored data if the occurrence of initial event can be exactly observed and observations on the occurrence of subsequent event are interval-censored (Chi and Tseng, 2002). Furthermore, if observations on the occurrence of subsequent event are right-censored, we then have usual right-censored failure time data for the survival time of interest.

The analysis of doubly censored data has recently attracted much attention, especially in the context of the analysis of AIDS incubation time, the time between the HIV infection and the diagnosis of AIDS. For example, De Gruttola and Lagakos (1989), Fang and Sun (2001), Gómez and Calle (1999) and Sun (1997) considered joint estimation of the distribution functions of HIV infection and AIDS incubation time based on AIDS cohort follow-up studies. Kim et al. (1993), Sun et al. (1999) and Pan (2001) investigated regression analysis of doubly censored data under the proportional hazards model.

[*] Corresponding author: e-mail: sunj@missouri.edu

For the analysis, most of the authors assume that the distribution of the survival time (e.g., AIDS incubation time) is independent of the occurrence of the initial event (e.g., HIV infection), which results in a simple likelihood function and greatly simplifies the analysis. However, there does not seem to be any existing test procedure for testing the assumption. In the following, a Wald test is developed for the problem. Note that data with doubly censoring could also occur in, for example, other disease progression studies and bird nesting studies (He, 2003).

We will begin with introducing notation and assumptions in Section 2. To formulate the problem, we will assume that the survival time of interest depends on the occurrence time of initial event through the proportional hazards model (Kalbfleisch and Prentice, 1980). Some comments on this will be given below. Note that this will yield an inference problem about the proportional hazards model with interval-censored continuous covariates, for which no established method is available. For the problem, Section 3 presents an estimating equation approach for estimating regression parameters, which naturally gives a Wald test procedure. Asymptotic properties of the proposed estimates are given. Section 4 reports some results from a simulation study conducted for evaluating the finite sample properties of the proposed estimate. In Section 5, we apply the method to an AIDS cohort study and Section 6 contains some concluding remarks.

2 Notation and assumptions

Consider a survival study that involves n independent subjects experiencing two related events. For subject i, let X_i and S_i^* denote the times of occurrences of initial and subsequent events, respectively, $i = 1, \ldots, n$. Define $T_i = S_i^* - X_i$, the survival time of interest. In an AIDS cohort study, X_i and S_i^* represent ages at HIV infection and AIDS diagnosis, respectively, and T_i corresponds to AIDS incubation time. For the relationship between X_i and T_i, we will assume that given $X_i = x_i$, the hazard process of T_i is given by the proportional hazards model and has the form

$$\lambda_i(t \mid x_i) = Y_i(t \mid x_i)\, \lambda_0(t) \exp(x_i \beta) \tag{1}$$

(Andersen and Gill, 1982). In the above, $\lambda_0(t)$ is an unknown baseline hazard function, $Y_i(t \mid X_i)$ is a predictable process defined below, and β denotes the regression parameter characterizing the dependence of T_i on X_i.

Note that the model (1) specifies that in AIDS context, AIDS incubation time depends on HIV infection time in a multiplicative fashion and patients infected by HIV in early days could have longer or shorter AIDS incubation time depending on situations. This seems to be a reasonable assumption since, for example, the patients infected by HIV later could in general benefit from more available and efficient treatments and thus have longer AIDS incubation time. Also it is well-known that the proportional hazards model provides good approximations to the majority of survival problems. Under model (1), the independence assumption between X_i and T_i is equivalent to the hypothesis $H_0 : \beta = 0$. It should be noted that if the X_i's can be exactly observed, this would become a standard testing problem. However, this is not the case here as we describe below.

To describe observed data, let $[L_i, R_i]$ denote the interval to which the occurrence time X_i of initial event is observed to belong, $i = 1, \ldots, n$. That is, we have interval-censored data on the X_i's. For the S_i^*'s, suppose that right-censored data are observed and given by $\{ (S_i = \min(S_i^*, C_i), \delta_i = I(S_i = S_i^*)); i = 1, \ldots, n \}$, where C_i is the censoring time associated with subject i and assumed to be independent of S_i^*. If $L_i = R_i$ or $X_i = 0$ for all i, the observed data on the T_i's reduce to usual right-censored data.

For inference, in the following, we will assume as others (Sun, 1998) that the mechanism yielding right- and interval-censoring is independent of occurrences of initial and subsequent events. More specifically, it will be assumed that for each i, L_i and R_i are independent of (X_i, C_i) and S_i^* and C_i satisfy the usual random right-censorship model. That is, the censoring is non-informative and the distributions of censoring variables $\{L_i, R_i, C_i\}$ are independent of parameters of interest. The distributions of L_i, R_i and X_i are all assumed to be of the discrete type with finite support.

For given $X_i = x_i$, define $Y_i(t \mid x_i) = I(S_i - x_i \geq t)$ and $N_i(t \mid x_i) = I(S_i - x_i \leq t, \delta_i = 1)$. Also define $\boldsymbol{X} = (X_1, \ldots, X_n)$ and

$$S^{(j)}(\beta, t, \boldsymbol{x}) = n^{-1} \sum_{i=1}^{n} Y_i(t \mid x_i)\, x_i^j \, e^{x_i \beta},$$

$j = 0, 1$, where $x_i^j = 1$ and $x_i^j = x_i$ for $j = 0$ and $j = 1$, respectively, and $\boldsymbol{x} = (x_1, \ldots, x_n)$. For estimation of β, note that if $L_i = R_i$, one can easily estimate it by the maximum partial likelihood estimator defined as the solution to the partial likelihood score equation

$$U(\beta \mid \boldsymbol{x}) = \int_0^{\tau} \sum_{i=1}^{n} \left\{ x_i - \frac{S^{(1)}(\beta, t, \boldsymbol{x})}{S^{(0)}(\beta, t, \boldsymbol{x})} \right\} dN_i(t \mid x_i) = 0 \tag{2}$$

(Andersen and Gill, 1982), where τ denotes the longest follow-up time. One advantage of the above estimation approach is that it does not involve the unknown baseline hazard function $\lambda_0(t)$ and can easily be implemented.

For current situations, however, it is obvious that the maximum partial likelihood estimator is not available. In next section, we propose an estimating equation approach for the problem.

3 Test Procedure

To test the hypothesis H_0, we first consider estimation of β in model (1). For this purpose, note that the score function $U(\beta \mid \boldsymbol{x})$ given in (2) can be regarded as a conditional score function given the X_i's or as a score function about both parameters β and the X_i's if we treat the X_i's as nuisance parameters. Thus by using the idea of the marginal likelihood method, it is natural to integrate out the unknown X_i's conditional on observed data. This motivates the following estimating equation

$$U(\beta, \hat{H}) = \left(\prod_{i=1}^{n} \hat{a}_i^{-1} \right) \int_{L_1}^{R_1} \ldots \int_{L_n}^{R_n} U(\beta \mid \boldsymbol{x}) \prod_{i=1}^{n} \{ d\hat{H}(x_i) \} = 0 \tag{3}$$

for the estimation of β. In the above, $\hat{a}_i = \int_{L_i}^{R_i} d\hat{H}(x)$, $i = 1, \ldots, n$, and $\hat{H}$ denotes the nonparametric maximum likelihood estimator of the cumulative distribution function H of the X_i's based on interval-censored data $\{ [L_i, R_i]\,;\, i = 1, \ldots, n \}$. Some comments on $\hat{H}$ will be given below.

Note that the idea behind equation (3) is similar to the one used in both the profile likelihood approach and random effects model methods. The same idea was used in Sun et al. (1999) for regression analysis of doubly censored failure time data, where unlike here covariates were assumed to be exactly observed. As based on equation (2), the method based on equation (3) has the same advantage that it does not require estimation of $\lambda_0(t)$, which makes the study of the asymptotic properties of the method possible.

In equation (3), we need to determine $\hat{H}$. For this, we will use the self-consistency algorithm proposed by Turnbull (1976) for both simulation study and the example. Note that $\hat{H}$ is a step function and not completely and uniquely defined on $\mathcal{R}^+$ (Turnbull, 1976). However, $d\hat{H}$ is completely defined and thus so does the equation (3). More discussion on $\hat{H}$ can be found in Ng (2002), Sun (1998) and Turnbull (1976) among others.

Let $\hat{\beta}$ denote the estimator of β given by the solution to equation (3) and β_0 the true value of β. It is apparent that $\hat{\beta}$ reduces to the maximum partial likelihood estimator of β if exact observations on the X_i's are available. We will show in the Appendix that $\hat{\beta}$ is a consistent estimate of β_0. It will also be shown in the Appendix that under mild regularity conditions, $n^{1/2}(\hat{\beta} - \beta_0)$ has an asymptotic normal distribution with mean zero and variance that can be consistently estimated by $\Gamma(\hat{\beta}) / A^2(\hat{\beta})$, where $A(\beta) = -n^{-1}\, \partial U(\beta, \hat{H}) / \partial\beta$ and $\Gamma(\beta) = n^{-1} \sum_{i=1}^{n} \hat{b}_i^2(\beta)$, where

$$\hat{b}_i(\beta) = \int_0^{\tau} \int_{L_1}^{R_1} \ldots \int_{L_n}^{R_n} \left\{ x_i - \frac{S^{(1)}(\beta, t, \boldsymbol{x})}{S^{(0)}(\beta, t, \boldsymbol{x})} \right\} \left\{ dN_i(t \mid x_i) - \frac{Y_i(t \mid x_i) \exp(x_i \beta)\, d\bar{N}(t \mid \boldsymbol{x})}{n S^{(0)}(\beta, t, \boldsymbol{x})} \right\} \prod_{l=1}^{n} \frac{d\hat{H}(x_l)}{\hat{a}_l},$$

$$\bar{N}(t \mid \boldsymbol{x}) = \sum_{i=1}^{n} N_i(t \mid x_i).$$

Once $\hat{\beta}$ is obtained, one can test the hypothesis H_0 by employing the Wald statistic $n\hat{\beta}^2 A^2(\hat{\beta}) / \Gamma(\hat{\beta})$ based on the χ^2 distribution with degree of freedom one. For calculation of $\hat{\beta}$, one way is to directly solve equation (3) using existing optimization algorithms, which are available in many statistical softwares. This is usually feasible for small data sets and some large data sets for which the resulting estimator $\hat{H}$ does not have many jumps. For general situations, we propose to use the following simple Monte Carlo method (Sun et al., 1999). Let K be a given integer.

Step 1. For each $k = 1, \ldots, K$ and $i = 1, \ldots, n$, sample a $X_i^{(k)}$ from $\hat{H}$ conditional on observed interval $[L_i, R_i]$, that is, $X_i^{(k)} \sim \hat{H}(\cdot \mid X \in [L_i, R_i])$.

Step 2. Let $\boldsymbol{X}^{(k)} = (X_1^{(k)}, \ldots, X_n^{(k)})$ and calculate $U(\beta \mid \boldsymbol{X}^{(k)} = \boldsymbol{x}^{(k)})$ in equation (2).

Step 3. Solve the equation $K^{-1} \sum_{k=1}^{K} U(\beta \mid \boldsymbol{x}^{(k)}) = 0$ to obtain $\hat{\beta}$.

If K is large, we should expect that the left side of the above equation will give a good approximation to $U(\beta, \hat{H})$. Note that the above Monte Carlo method is used only to compute $\hat{\beta}$, not the variance estimate. Once $\hat{\beta}$ is obtained, its variance can be estimated by the closed formula given above. We remark that an alternative to the above Monte Carlo method is to use the multiple imputation (Pan, 2001). The multiple imputation may be simpler, but the theoretical result given above about $\hat{\beta}$ may not be appropriate for finite sample situations.

4 Simulation Studies

A simulation study was conducted to evaluate finite sample properties of the proposed estimator $\hat{\beta}$. In the study, for simplicity, we assumed that all concerned variables are discrete and take integer values. The doubly censored data were generated as arising from AIDS cohort follow-up studies such as the example discussed in the next section. First we generated X_i from the uniform distribution $U\{1, \ldots, 7\}$ and T_i from a discretized exponential distribution with the hazard function given in (1) and $\lambda_0(t) = 0.1$. Then S_i^* was defined as $S_i^* = X_i + T_i$ with the common censoring time $C_i = 15$ for all $i = 1, \ldots, n$. The observed interval for X_i was generated by letting $L_i = \max\{1, X_i - a_i^{(1)}\}$ and $R_i = \min\{X_i + a_i^{(2)}, 7\}$, where $a_i^{(1)}$ and $a_i^{(2)}$ were generated from the uniform distribution $U\{0, 1, \ldots, b\}$, where b is a constant and used to control the extent of interval-censoring. The results reported below are based on $n = 100$ and $K = 100$ with 1000 replications.

In the study, we mainly focused on the comparison of the proposed point estimate and the maximum partial likelihood estimate, $\hat{\beta}_p$ say, of β that would be obtained if the X_i's were exactly observed. Also we were interested in investigating the approximation of the asymptotic normal distribution given in the previous section to the estimate. Table 1 presents the means of $\hat{\beta}$ and their MSE (values in brackets) based on simulated data for different true values of β with $b = 1$ and $b = 2$, respectively. For the comparison, for each situation, assuming that the exact time of occurrence of initial event was known, we also obtained the corresponding $\hat{\beta}_p$ and MSE and included their means in Table 1. It can be seen from the table that the results from the two methods are quite close to each other for most cases considered, indicating that the proposed method works reasonably well.

For the above simulation set-ups, we also calculated powers for testing $\beta = 0$ based on the proposed Wald statistic and $\hat{\beta}_p$, respectively. The procedure based on $\hat{\beta}$ gave powers of 0.320 and 0.224 for $\beta = 0.1$ and -0.1, respectively, for the case of $b = 1$ with the significance level of 0.05, while the procedure based on $\hat{\beta}_p$ yielded powers of 0.346 and 0.238. These suggest that the proposed method has reasonable power. To assess the approximation of the asymptotic normal distribution to the distribution of $\hat{\beta}$, the probability plots of the standardized $\hat{\beta}$ against the standard normal distribution were studied and compared to the corresponding plots of the standardized $\hat{\beta}_p$. Figures 1 and 2 display such plots of the standardized $\hat{\beta}_p$ and $\hat{\beta}$, respectively, for the situation $\beta = 0$ considered in Table 1. They suggest that the approximation is quite satisfactory and similar plots were obtained for other situations.

Note that for a given situation or particular problem, one needs to choose K for the Monte Carlo method for determining $\hat{\beta}$. A general and simple rule is to try several values of K or to increase K until a stable $\hat{\beta}$ is obtained. In the study here, we also tried $K = 200$ and $K = 500$ and no significant differences were observed.

Table 1 Means of regression parameter estimates and their MSE.

	$b = 1$		$b = 2$	
True β	$\hat{\beta}$	$\hat{\beta}_p$	$\hat{\beta}$	$\hat{\beta}_p$
0.00	−0.0026	−0.0002	−0.0047	−0.0002
	(0.0081)	(0.0082)	(0.0089)	(0.0082)
0.10	0.0885	0.0941	0.0825	0.0941
	(0.0141)	(0.0154)	(0.0138)	(0.0154)
−0.10	−0.0966	−0.0979	−0.0937	−0.0979
	(0.0205)	(0.0206)	(0.0211)	(0.0206)

5 Illustration: Application to an AIDS Study

In this section, we apply the proposed methodology to the AIDS cohort study discussed by De Gruttola and Lagakos (1989) and Sun et al. (1999) among other authors. The study consists of patients with Type A or B hemophilia who were at risk for HIV infection through the contaminated blood factor they received for their treatments. One of the main objectives of the study is to estimate the distribution of AIDS incubation time (T_i), the time between HIV infection (X_i) and AIDS diagnosis (S_i^*). For the analysis, most authors assumed that the AIDS incubation time is independent of HIV infection time. The goal here is to assess this independence assumption.

The observed data can be found in Table 1 of Sun et al. (1999) and include observed intervals for X_i and right-censored AIDS diagnosis times. Here interval-censored observations occurred due to the fact that HIV infection was detected through periodic blood tests and the reason for right-censoring on AIDS diagnosis time is that the study ended before the development of AIDS for many subjects. Assume that model (1) holds for AIDS incubation time. The application of the proposed method yielded $\hat{\beta} = -0.0639$ with the estimated standard deviation of 0.0621 with $K = 100$. This corresponds to a p-value of 0.3035 according to the χ^2 distribution and suggests that the independence assumption seems to hold for the case considered here. We also tried some larger values of K as in simulation and got similar estimates.

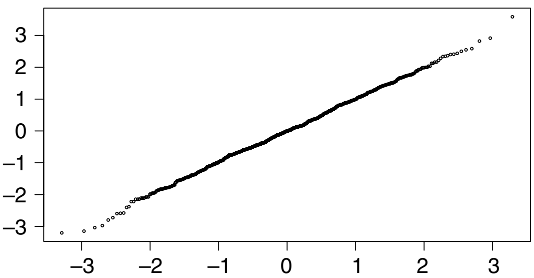

Figure 1 Quantile plot for exactly observed HIV infection time.

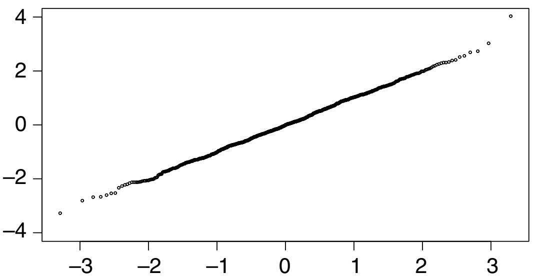

Figure 2 Quantile plot for interval-censored HIV infection time.

In the original study, the subjects were classified into two groups, lightly and heavily treated groups, according to the amount of blood they received. Corresponding to this, we also applied the method separately to each of the two groups and similar results were obtained.

6 Concluding Remarks

This paper considered the problem of testing independence assumption between HIV infection and AIDS incubation time. The assumption has been commonly used in the analysis of AIDS cohort studies and has not been addressed before. For the problem, we proposed an estimating equation-based approach assuming that the proportional hazards model can be used to describe the relationship between HIV infection and AIDS incubation time, and a Wald test statistic was presented. The simulation results suggested that the method works reasonably well.

It should be noted that the approach given here depends on model (1), the use of which makes the problem tractable. In other words, the hypothesis H_0 is equivalent to the independence assumption under the model (1) and the approach may not be able to detect the dependence of T_i on X_i if the true relationship between them does not satisfy the model (1). For example, in some situations, rather than model (1), their relationship may be better described by the additive hazards model or the proportional odds model. Another possibility is that their relationship may fit better the Cox model with time-varying covariate effect. This of course leads to a related question: the checking of the validity of model (1), which is beyond this investigation and for which some extra information or data sets may be needed. If there is evidence against model (1), then a different approach would be needed to test the independence assumption.

For the problem considered here, an alternative method is to use the full likelihood approach based on the likelihood function

$$\prod_{i=1}^{n} \int_{L_i}^{R_i} \{ \lambda_0(t_i) \exp(x_i\beta) \}^{\delta_i} \exp\left\{ -e^{x_i\beta} \int_0^{t_i} \lambda_0(u)\, du \right\} dH(x_i) \,. \tag{5}$$

Several authors have used this approach to study AIDS incubation time (De Gruttola and Lagakos, 1989; Frydman, 1995; Goggins et al., 1999). In comparison to the method given here, although it could be more efficient, the full likelihood approach has the disadvantage that it is complicated in computation since some iterative algorithms such as self-consistency algorithm or Monte Carlo EM algorithm have to be used. More importantly, the asymptotic properties are unknown and thus no rigorous test can be constructed.

In the preceding sections, we have focused on testing H_0 assuming that there do not exist covariates. However, this may not be the case in many situations. Suppose that for subject i, there exist a vector of covariates z_i whose effects on the T_i's are of interest. In this case, the model (1) can be generalized to

$$\lambda_i(t \mid x_i) = Y_i(t \mid x_i) \lambda_0(t) \exp(x_i\beta + z_i\gamma)$$

and it is straightforward to develop methods similar to the proposed approach for testing H_0 and making inference about γ, where γ represents the effect of z_i on T_i. Note that in the above model we use $Y_i(t \mid x_i)$ instead of $Y_i(t \mid x_i, z_i)$ because Y_i is an indicator function that may directly depend on x_i, but not on z_i (Andersen and Gill, 1982). More specifically, estimating equations for both β and $\boldsymbol{\gamma}$ together can be similarly developed and all derivations and properties of $\hat{\beta}$ given above should hold for the estimates of β and γ defined by these equations. Note that if the test of the hypothesis H_0 is not significant, the above problem reduces that discussed in Sun et al. (1999).

A more general situation, whose discussion is beyond this paper and will be reported elsewhere, that may occur in practice is that in addition to interval-censoring on X_i, some or all components of z_i also suffer interval-censoring. In this case, we have an inference problem about the proportional hazards model with right-censored failure time data and interval-censored covariates if the test of the hypothesis H_0 is not significant, or that with doubly censored data and interval-censored covariates otherwise. For either case, there does not seem to exist well-established methods.

Appendix: Proofs

Let β_0, $\hat{\beta}$, $U(\beta \mid \boldsymbol{x})$, and $U(\beta, \hat{H})$ be defined as before and use the notation given in the previous sections. Assume that the regularity conditions given in Andersen and Gill (1982) for the case of right-censored failure time data and in Groeneboom and Wellner (1992) and Yu et al. (1998) for the strong consistency of $\hat{H}$ hold.

Consistency of $\hat{\beta}$. Define for fixed $\boldsymbol{x} = (x_1, \ldots, x_n)$

$$X_n(\beta, t \mid \boldsymbol{x}) = n^{-1} \sum_{i=1}^{n} \int_0^t (\beta - \beta_0) x_i \mathrm{d}N_i(u) - \int_0^t \log \left\{ \frac{\sum_{i=1}^{n} Y_i(u \mid x_i) e^{x_i \beta}}{\sum_{i=1}^{n} Y_i(u \mid x_i) e^{x_i \beta_0}} \right\} \mathrm{d}\bar{N}(u \mid \boldsymbol{x}) ,$$

$$A_n(\beta, t \mid \boldsymbol{x}) = \int_0^t \left[(\beta - \beta_0) S^{(1)}(\beta_0, u, \boldsymbol{x}) - \log \left\{ \frac{S^{(0)}(\beta, u, \boldsymbol{x})}{S^{(0)}(\beta_0, u, \boldsymbol{x})} \right\} S^{(0)}(\beta_0, u, \boldsymbol{x}) \right] \lambda_0(u) \, \mathrm{d}u ,$$

and $M_i(t \mid x_i) = N_i(t \mid x_i) - \int_0^t \lambda_0(s) \, Y_i(s \mid x_i) \exp(x_i \beta_0) \, \mathrm{d}s$, $i = 1, 2, \cdots, n$. Then we have

$$X_n(\beta, t \mid \boldsymbol{x}) - A_n(\beta, t \mid \boldsymbol{x}) = n^{-1} \sum_{i=1}^{n} \int_0^t \left\{ (\beta - \beta_0) x_i - \log \frac{S^{(0)}(\beta, u, \boldsymbol{x})}{S^{(0)}(\beta_0, u, \boldsymbol{x})} \right\} \mathrm{d}M_i(u \mid x_i) ,$$

which is a locally square integrable martingale. Using arguments similar to those in Anderson and Gill (1982), we see that $X_n(\beta, \tau \mid \boldsymbol{x}) - A_n(\beta, \tau \mid \boldsymbol{x}) \xrightarrow{P} 0$ as $n \to \infty$.

Also note that $A_n(\beta, \tau \mid \boldsymbol{x}) \xrightarrow{P} A(\beta, \tau)$, where

$$A(\beta, \tau) = \int_0^\tau \left[(\beta - \beta_0) \, s^{(1)}(\beta_0, u) - \log \left\{ \frac{s^{(0)}(\beta, u)}{s^{(0)}(\beta_0, u)} \right\} s^{(0)}(\beta_0, u) \right] \lambda_0(u) \, \mathrm{d}u$$

and $s^{(0)}$ and $s^{(1)}$ are the stochastic limits of $S^{(0)}$ and $S^{(1)}$, respectively. It thus follows that $X_n(\beta, \tau \mid \boldsymbol{x}) \xrightarrow{P} A(\beta, \tau)$ and

$$X_n(\beta, \tau) = \int_{L_1}^{R_1} \cdots \int_{L_n}^{R_n} X_n(\beta, \tau \mid \boldsymbol{x}) \prod_{l=1}^{n} \hat{a}_l^{-1} \, \mathrm{d}\hat{H}(x_l) \to A(\beta, \tau)$$

both in probability due to the strong uniform consistency of $\hat{H}$ (Groeneboom and Wellner, 1992; Yu et al., 1998). The above convergence statement and the fact that both $X_n(\beta, \tau)$ and $A(\beta, \tau)$ are concave functions of β with a unique maximum at $\beta = \hat{\beta}$ and $\beta = \beta_0$, respectively, show that asymptotically, $\hat{\beta} \to \beta_0$ in probability. That is, $\hat{\beta}$ is (weakly) consistent.

Asymptotic Normality of $\hat{\beta}$. To prove the asymptotic normality, first note that the application of Taylor series expansion to $U(\beta, \hat{H})$ yields, asymptotically,

$$n^{-\frac{1}{2}} U(\beta_0, \hat{H}) = \left\{ -n^{-1} \frac{\partial U(\beta, \hat{H})}{\partial \beta} \Big|_{\beta = \beta^*} \right\} \{ n^{\frac{1}{2}} (\hat{\beta} - \beta_0) \} ,$$

where β^* is on the segment between β_0 and $\hat{\beta}$. Following Anderson and Gill (1982), we can easily show that as $n \to \infty$, $A(\beta^*) = -n^{-1} \, \partial U(\beta, \hat{H}) / \partial \beta \big|_{\beta = \beta^*}$ converges in probability to

$$\int_0^\tau \left[\frac{s^{(2)}(\beta_0, t)}{s^{(0)}(\beta_0, t)} - \left(\frac{s^{(1)}(\beta_0, t)}{s^{(0)}(\beta_0, t)} \right)^2 \right] s^{(0)}(\beta_0, t) \, \lambda_0(t) \, \mathrm{d}t ,$$

where $s^{(2)}(\beta, t) = \partial^2 s^{(0)}(\beta, t)/\partial\beta^2$. Thus, for the proof, it is sufficient to show that $n^{-1/2}U(\beta_0, \hat{H})$ is asymptotically normally distributed with mean zero and variance that can be estimated by $\Gamma(\hat{\beta})$ given in Section 3.

For this, as in Sun et al. (1999), we can show that

$$n^{-\frac{1}{2}}U(\beta_0, \hat{H}) = n^{-\frac{1}{2}} \sum_{i=1}^{n} \hat{a}_i^{-1} \int_{L_i}^{R_i} \int_0^{\tau} u_i(\beta_0, t \,|\, x_i)\, \mathrm{d}M_i(t \,|\, x_i)\, \mathrm{d}\hat{H}(x_i) \; + \; o_P(1) \,,$$

where $u_i(\beta_0, t \,|\, x_i) = x_i - s^{(1)}(\beta_0, t)/s^{(0)}(\beta_0, t)$. Let D_n^* denote the quantity at the right side of the above equation. Define

$$D_n = n^{-\frac{1}{2}} \sum_{i=1}^{n} a_i^{-1} \int_{L_i}^{R_i} \int_0^{\tau} u_i(\beta_0, t \,|\, x_i)\, \mathrm{d}M_i(t \,|\, x_i)\, \mathrm{d}H(x_i)$$

with $a_i = \int_{L_i}^{R_i} \mathrm{d}H(x_i)$. D_n can be easily shown to converge in law to a normally distributed variate with mean zero and variance that can be consistently estimated by $\Gamma(\hat{\beta})$ as defined in Section 3. Thus to prove the asymptotic distribution of $n^{-1/2}U(\beta_0, \hat{H})$, we only need to prove that for any $\varepsilon > 0$,

$$P(\,|\, D_n^* - D_n \,| \, \geq \varepsilon) \to 0 \,. \tag{5}$$

For this, note that

$$\begin{aligned} P(\,|\, D_n^* - D_n \,| \, \geq \varepsilon) \leq \frac{1}{n\varepsilon^2} \sum_{i=1}^{n} E \Bigg[& \hat{a}_i^{-1} \int_{L_i}^{R_i} \int_0^{\tau} u_i(\beta_0, t \,|\, x_i)\, \mathrm{d}M_i(t \,|\, x_i)\, \mathrm{d}\hat{H}_n(x_i) \\ & - a_i^{-1} \int_{L_i}^{R_i} \int_0^{\tau} u_i(\beta_0, t \,|\, x_i)\, \mathrm{d}M_i(t \,|\, x_i)\, \mathrm{d}H(x_i) \Bigg]^2 \,. \end{aligned} \tag{6}$$

Let $G(X_i) = \int_0^{\tau} u_i(\beta_0, t \,|\, X_i)\, \mathrm{d}M_i(t \,|\, X_i)$, which is a a right-continuous function with bounded variation on any finite interval. Then we have

$$\int_{L_i}^{R_i} \int_0^{\tau} u_i(\beta_0, t \,|\, x_i)\, \mathrm{d}M_i(t \,|\, x_i)\, \mathrm{d}\hat{H}_n(x_i) = G(R_i)\, \hat{H}_n(R_i) - G(L_i)\, \hat{H}_n(L_i) - \int_{L_i}^{R_i} \hat{H}_n(x_i)\, \mathrm{d}G(x_i) \tag{7}$$

and

$$\int_{L_i}^{R_i} \int_0^{\tau} u_i(\beta_0, t \,|\, x_i)\, \mathrm{d}M_i(t \,|\, x_i)\, \mathrm{d}H(x_i) = G(R_i)\, H(R_i) - G(L_i)\, H(L_i) - \int_{L_i}^{R_i} H(x_i)\, \mathrm{d}G(x_i) \,. \tag{8}$$

Therefore the equation (5) follows by plugging in (7) and (8) into (6), which completes the proof.

Acknowledgement The authors wish to thank the Editor, Professor Bauer, and two referees for their many helpful comments and suggestions, which greatly improved the paper. The research of the first author was supported in part by a grant from the U.S. National Institutes of Health.

References

Andersen, P. K. and Gill, R. D. (1982). Cox's regression model for counting processes: a large sample study. *Annual of Statistcs* **10**, 1100–1120.

Chi, Y. and Tseng, C. H. (2002). Comparison of several relative risk estimators with interval-censored data. *Biometrical Journal* **44**, 197–212.

De Gruttola, V. and Lagakos, S. W. (1989). Analysis of doubly-censored survival data, with application to AIDS. *Biometrics* **45**, 1–12.

Fang, H. and Sun, J. (2001). Consistency of nonparametric maximum likelihood estimation of a distribution function based on doubly interval-censored failure time data. *Statistics and Probability Letters* **55**, 311–318.

Frydman, H. (1995). Semiparametric estimation in a three-state duration-dependent Markov model from interval-censored observations with application to AIDS data. *Biometrics* **51**, 502–511.

Goggins, W. B., Finkelstein, D. M. and Zaslavsky, A. M. (1999). Applying the Cox proportional hazards model for analysis of latency data with interval censoring. *Statistics in Medicine* **18**, 2737–2747.

Gómez, G. and Calle, M. L. (1999). Nonparametric estimation with doubly censored data. *Journal of Applied Statistics* **26**, 45–58.

Groeneboom, P. and Wellner, J. A. (1992). *Information Bounds and Nonparametric Maximum Likelihood Estimation*. Birkhauser Verlag, Basel.

He, C. Z. (2003). Bayesian modeling of age-specific survival in bird nesting studies under irregular visits. *Biometrics* **59**, 962–973.

Joly, P. and Commenges, D. (1999). A penalized likelihood approach for a progressive three-state model with censored and truncated data: application to AIDS. *Biometrics* **55**, 887–890.

Kalbfleisch, J. D. and Prentice, R. L. (1980). *The Statistical Analysis of Failure Time Data.* New York: Wiley.

Kim, M. Y., De Gruttola, V. G. and Lagakos, S. W. (1993) Analyzing doubly censored data with covariates, with application to AIDS. *Biometrics* **49**, 13–22.

Ng, M. P. (2002). A modification of Peto's nonparametric estimation of survival curves for interval-censored data. *Biometrics* **58**, 439–442.

Pan, W. (2001). A multiple imputation approach to regression analysis for doubly censored data with application to AIDS studies. *Biometrics* **57**, 1245–1250.

Sun, J. (1998). Interval censoring. *Encyclopedia of Biostatistics*, John Wiley & Sons Ltd., 2090–2095.

Sun, J. (2004). Statistical analysis of doubly interval-censored failure time data. In: (eds.) Balakrishnan, N. and Rao, C. R., *Handbook of Statistics: Survival Analysis*, Elsevier, North Holland, pp. 105–122.

Sun, J., Liao, Q. and Pagano, M. (1999). Regression analysis of doubly censored failure time data with applications to AIDS studies. *Biometrics* **55**, 909–914.

Turnbull, B. W. (1976). The empirical distribution function with arbitrarily grouped censored and truncated data. *Journal of Royal Statistical Society* B **38**, 290–295.

Yu, Q., Schick, A., Li, L. and Wong, G. Y. C. (1998). Asymptotic properties of the GMLE with case 2 interval-censored data. *Statistics and Probability Letters* **37**, 223–228.